Numerical simulation of nonlinear shallow-water interactions between surface waves and a floating structure

Ali Haidar

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Mathématiques Appliquées

École doctorale I2S

Unité de recherche IMAG-ACSIOM

Numerical simulation of nonlinear shallow-water interactions between surface waves and a floating structure

Présentée par Ali Haidar
le 2 Décembre 2022

Sous la direction de Fabien Marche et François Vilar

Devant le jury composé de

David LANNES, Directeur de Recherche CNRS, Université de Bordeaux UMR 5251	Président
Rémi ABGRALL, Professeur des Universités, Université de Zurich	Rapporteur
Christophe BERTHON, Professeur des Universités, Université de Nantes	Rapporteur
Elena GABURRO, Chargée de Recherche INRIA, Centre Bordeaux-Sud-Ouest	Examinateur
Fabien MARCHE, Maître de Conférences, Université de Montpellier	Directeur de thèse
Roland MASSON, Professeur des Universités, Université Côte d'Azur	Examinateur
François VILAR, Maître de Conférences, Université de Montpellier	co-Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

À ma mère, À mon père,
À ma fleur Marwa

Résumé

Dans cette Thèse de Doctorat, nous nous intéressons à deux problématiques: (i) le développement de stratégies de stabilisation pour des méthodes de type *discontinuous Galerkin* (DG) appliquées à des écoulements *shallow-water* fortement non-linéaires, (ii) le développement d'une stratégie de modélisation et de simulation numérique des interactions non-linéaires entre les vagues et un objet flottant en surface, partiellement immergé. Les outils développés dans le cadre du premier axe de travail sont mis à profit et valorisés au cours de la deuxième partie.

Les méthodes de discrétisation de type DG d'ordre élevé présentent en général des problèmes de robustesse en présence de singularités de la solution. Ces singularités peuvent être de plusieurs natures: discontinuité de la solution, discontinuité du gradient ou encore violation de la positivité de la hauteur d'eau pour des écoulements à surface libre. Nous introduisons dans la première partie de ce manuscript deux approches de type *Finite-Volume Subcells* permettant d'apporter une réponse à ces problèmes de robustesse. La première approche repose sur une correction *a priori* du schéma DG associée à un limiteur TVB et un limiteur de positivité. La seconde approche s'appuie quant à elle sur une correction *a posteriori* permettant d'identifier avec une meilleure précision les cellules incriminée, ainsi que sur les propriétés de robustesse inhérentes au schéma Volumes-Finis limite d'ordre un. Cette seconde approche permet d'assurer la robustesse du schéma DG initial en présence de discontinuité, ainsi que la positivité de la hauteur d'eau, tout en préservant une excellente qualité d'approximation, bénéficiant d'une résolution de l'ordre de la sous-maille. De façon préliminaire, cette seconde approche est également étendue au cas de la dimension deux d'espace horizontal. De nombreux cas-test permettent de valider cette approche.

Dans la seconde partie, nous introduisons une nouvelle stratégie numérique conçue pour la modélisation et la simulation des interactions non linéaires entre les vagues en eau peu profonde et un objet flottant partiellement immergé. Au niveau continu, l'écoulement situé dans le domaine extérieur est globalement modélisé par les équations hyperboliques non-linéaires de Saint-Venant, tandis que la description de l'écoulement sous l'objet se réduit à une équation différentielle ordinaire non linéaire. Le couplage entre l'écoulement et l'objet est formulé comme un problème au bord, associé au calcul de l'évolution temporelle de la position des points d'interface air-eau-objet. Au niveau discret, la formulation proposée s'appuie sur une approximation DG d'ordre arbitraire, stabilisée à l'aide de la méthode de correction locale des sous-cellules (*a posteriori*) introduite dans la première partie. L'évolution temporelle de l'interface air-eau-objet est calculée à partir d'une description *Arbitrary Lagrangian-Eulerian* (ALE) et d'une transformation appropriée entre la configuration initiale et celle dépendant du temps. Pour n'importe quel ordre d'approximation polynomiale, l'algorithme résultant est capable de: (i) préserver la loi de conservation géométrique discrète (DGCL), (ii) garantir la préservation de la positivité de la hauteur d'eau au niveau des sous-cellules, (iii) préserver la classe des états stationnaires au repos (well-balancing), éventuellement en présence d'un objet partiellement immergé. Plusieurs validations numériques sont présentées, montrant le caractère opératoire de cette approche, et mettant en évidence que le modèle numérique proposé: (i) permet effectivement de modéliser les différents types d'interactions vague / objet flottants, (ii) calcul efficacement l' évolution temporelle des points de contact air-eau-objet et redéfinit en conséquence le nouveau maillage grâce à la méthode ALE (iii) gère avec précision et robustesse les possibles singularités de l'écoulement, (iv) préserve la haute résolution des schémas DG au niveau des sous-cellules.

Abstract

In this Ph.D., we investigate two main research problems: (i) the design of stabilization patches for higher-order discontinuous-Galerkin (DG) methods applied to highly nonlinear free-surface shallow-water flows, (ii) the construction of a new numerical approximation strategy for the simulation of nonlinear interactions between waves in a free-surface shallow flow and a partly immersed floating object. The stabilization methods developed in the first research line are used in the second part of this work.

High-order discontinuous-Galerkin (DG) methods generally suffer from a lack of nonlinear stability in the presence of singularities in the solution. Such singularities may be of various kinds, involving discontinuities, rapidly varying gradients or the occurence of dry areas in the particular case of free-surface flows. In the first part of this work, we introduce two new stabilization methods based on the use of Finite-Volume Subcells in order to alleviate these robustness issues. The first method relies on an *a priori* limitation of the DG scheme, together with the use of a TVB slope-limiter and a PL. The second one is built upon an *a posteriori* correction strategy, allowing to surgically detect the incriminated local subcells, together with the robustness properties of the corresponding lowest-order Finite-Volume scheme. This last strategy allows to ensure the nonlinear stability of the DG scheme in the vicinity of discontinuities, as well as the positivity of the discrete water-height, while preserving the subcell resolution of the initial scheme. This second strategy is also preliminary investigated in the two dimensional horizontal case. An extensive set of test-cases assess the validity of this approach.

In the second part, we introduce a new numerical strategy designed for the modeling and simulation of nonlinear interactions between surface waves in shallow-water and a partially immersed surface piercing object. At the continuous level, the flow located in the *exterior* domain is globally modeled with the nonlinear hyperbolic shallow-water equations, while the description of the flow beneath the object reduces to a nonlinear ordinary differential equation. The coupling between the flow and the object is formulated as a free-boundary problem, associated with the computation of the time evolution of the spatial locations of the air-water-body interface. At the discrete level, the proposed formulation relies on an arbitrary-order discontinuous Galerkin approximation, which is stabilized with the *a posteriori* Local Subcell Correction method through low-order finite volume scheme introduced in the first part. The time evolution of the air-water-body interface is computed from an Arbitrary-Lagrangian-Eulerian (ALE) description and a suitable smooth mapping between the original frame and the current configuration. For any order of polynomial approximation, the resulting algorithm is shown to: (i) preserves the Discrete Geometric Conservation Law, (ii) ensures the preservation of the water-height positivity at the subcell level, (iii) preserves the class of motionless steady states (well-balancing), possibly with the occurrence of a partially immersed object. Several numerical computations and test-cases are presented, highlighting that the proposed numerical model (i) effectively allows to model all types of wave / object interactions, (ii) efficiently provides the time-evolution of the air-water-body contact points and accordingly redefine the new mesh-grid thanks to ALE method (iii) accurately handles strong flow singularities without any robustness issues, (iv) retains the highly accurate subcell resolution of discontinuous Galerkin schemes.

Remerciement

Avant de remercier toute personne, je commencerai par remercier le Seigneur, sans lui je ne suis absolument rien, c'est lui le clément, le miséricordieux et le soutien, c'est lui qui m'a tenu la main et m'a guidé et aidé à traverser toutes les difficultés. Merci généreux, merci Dieu.

J'aimerais ensuite remercier sans doute mes encadrants de thèse, Fabien Marche et François Vilar, pour m'avoir transmis beaucoup de choses de leur grande expérience tout au long de cette thèse. Fabien et François m'ont donné de leurs grandes connaissances théoriques et pratiques dans le domaine de l'analyze numérique, ils étaient toujours présents pour me guider et répondre à mes questions. En plus de leur haute moralité, leur gentillesse et leur ouverture d'esprit, Fabien et François sont des personnes rares et agréables que je considère comme des grands frères. Collaborer avec eux m'a énormément apporté sur le coté scientifique et sur le coté humain.

A bien des égards, les compétences scientifiques de David Lannes ont joué un rôle décisif au cours de ma dernière année. Bien que nos entretiens n'aient pas été nombreux, ses idées lumineuses ont systématiquement permis de grandes avancées dans mes travaux. Il n'a pas non plus hésité à m'apporter son soutien et me donner de son temps lorsque c'était nécessaire. Un grand merci donc.

J'adresse un grand merci aux membres du jury pour s'être intéressé à l'étude de mon travail: merci Rémi Abgrall, Christophe Berthon, Elena Gaburro, David Lannes et Roland Masson pour votre temps et vos efforts.

Intégrer l'équipe des chercheurs et enseignants à l'université de Montpellier m'a permis de riches rencontres et connaissances avec des amis et collègues que chacun/une a un souvenir spécial dans mon esprit, merci Meriem, Antonia, Florian, Morgane, Paul, Thiziri, Franschesko et de gens qui ont quitté le labo, André, Gwenael et Robert. Merci à Baptiste Chapuisat de m'avoir sauvé plusieurs fois en réparant mon ordinateur, toujours disponible, efficace et de bonne humeur. Merci aussi à Ghislain Durif de m'avoir aussi sauvé une fois. Un grand merci à Sophie Cazanave, Nathalie Collain et Brigitte Labric, qui s'occupent aussi d'une manière admirable des membres de l'équipe. J'adresserai une pensée particulière à mes amis qui on partagé avec moi mon bureau, Tristan, Chayma et Tangui, merci pour leurs sourire, dynamisme, aide et profonde gentillesse.

Ibrahim Bouzalmat, mon cher ami et frère qui a partagé avec moi des beaux moments, il était toujours à coté de moi pour m'aider et me supporter, de plus, au niveau académique, un agréable collègue dans le labo.

Je tiens aussi à remercier chaleureusement Marien Hannot, un collègue et un ami qui a toujours été present pour des discussions et répondre à mes questions: il est hyper fort et hyper intelligent ce Marien.

Intégrer l'équipe ACSIOM à l'IMAG m'a permis de riches rencontres et activités, surtout les séminaires qui sont été organisé par François Vilar et Vanessa Lleras que je les remercie vivement.
Je voudrais dire merci à Pascal Azerad, Bijan Mohammadi, Damien Calaque, Vanessa Lleras, Thomas Hausberger, Alain Bruguières et Jérémy Nusa pour leur bonne collaboration lors des enseignements, j'ai éprouvé du plaisir à partager des enseignements avec eux, et acquérir de l'expérience dans le domaine de l'enseignement. Puisqu'on parle de l'enseignement, un merci particulier à

quelqu'un de sérieux et très gentil: Benoîte de Saporta, quelqu'un de très dynamique et actif.

Je ne pourrai jamais oublier mes professeurs à l'Université de Nantes, qui m'ont transmis des connaissance théoriques et numériquesafin d'être accepté dans cette thèse: merci à Christophe Berthon, Hélène Mathis, Anais Crestetto, Marianne Besmoulin, Guy Moebs, Vincent Franjou, Samuel Tapis, Mazen Saad. Un grand merci à Yannick Descantes qui m'a encadré dans mon stage de fin d'étude en Master 2. Yannick a un respect particuliers dans mon coeur, il m'a donné beaucoup de son expérience, avec lui j'ai renforcé mes capacité de recherches autonomes où j'ai du manipuler de gros codes pour la première fois. Avec Yannick, c'était mon premier pas dans le monde de la recherche.

Mes trois première années universitaires ont été faite à l'université Libanaise où j'ai acquis les premières connaissances mathématiques de bases, donc un très grand merci à mes professeurs de licence qui m'ont enseigné pendant cette période, Bassam Kojok, Mohammad Chaayto, Mohammad Abbas, Hassan Abbas, Ibtisam Zaiter, Nahla Noun, Elissar Nasreddine, Raafat Talhouk, Youssef Ayyad, Wael youssef, Zaynab Alsaghir, Ali Ayyad, Nasri Chaayto, Narjis Cheeb, Bassam Hamdoun, Mohammad Hamieh. Raafat Tarraf récemment décédé, peu d'occasions m'ont été offerte pour exprimer mon respect et appréciation pour ce prof que j'ai aimé, que ton âme repose en paix.

Contents

Chapter 1

Introduction

1.1 Modeling: a floating object in shallow-water

The mathematical and numerical study of the propagation and transformations of waves in the presence of a floating structure is a complex problem in which one has to model the time evolution of the mechanical system made of a solid body partially immersed in an incompressible fluid. Besides the hydrodynamic issues generally associated with the simulation of free surface flows, an important additional difficulty is that the immersed part of the structure (the *wetted surface*) generally depends on time, leading to another free boundary problem.

Linear modeling

The modeling of such a system can be traced back to the pioneering work [97], in which a linear potential model is used for the hydrodynamics (the fluid is assumed irrotational and inviscid), and the motion of the solid is assumed to be of small amplitude around a fixed mean position. Although being over-simplified for applications of interest, this approach put the light on the important (and difficult) issue of defining suitable transmission conditions between the *exterior* area (where the surface waves do not interact with the structure) and the *interior* area (the fluid under the solid, exercising a pressure on the wetted surface). Such a linear strategy has been refined, extended and adapted in several subsequent studies, let mention for instance the important domains of offshore structures [132], floating Wave Energy Converters (WEC) [108] or floating breakwaters, see for instance [173, 104]. When combined with Boundary Element Methods (linear-BEM, in frequency domain) for the computation of the hydrodynamics, the linear potential theory defines the background of several popular dedicated softwares, like WAMIT or ANSYS Aqwa. Among the descriptions and models grounded on linearity assumptions, let also mention "semi-analytic" methods, like for instance the *point absorber* method for the hydrodynamic interactions with WEC where some diffraction by the devices are neglected and the potential flow is described with a semi-analytical representation, see for instance [27], or the use of the (time dependent) *mild-slope* equation, see for instance [13]. Such linear models are particularly fast when compared to Reynolds Averaged Navier-Stokes (RANS) simulations, see for instance [182] for an application to WEC, or to potential approaches with a nonlinear surface boundary. Moreover, for small to moderate sea states, the assumptions related to linear theory generally provide some numerical previsions with reliable leading orders or approximation.

Nonlinear strategies

However, for larger sea states, the linearity assumptions do not hold anymore as nonlinear effects become important, see for instance the case for WECs operating inside the resonance domain, or floating breakwaters in the nearshore area. This is also true for the bathymetry effects, which are generally neglected within linear-BEM in intermediate sea levels, but cannot being neglected anymore in shallow-water, where nonlinear dynamics and bottom induced effects such as shoaling, refraction or energy transfers and dissipation, may be significant. Hence, several attempts to account for nonlinear effects have been reported in the literature and most of them rely on fully nonlinear potential flow models used together with nonlinear BEM, in time/physical domain, also allowing to account for varying bathymetry [79]. Another strategy, yet far less investigated, is to use some simplified asymptotic models instead of the full water-waves equations. Using such simpler asymptotics may appear as an interesting compromise between oversimplified linear models and too expensive CFD strategies. Such depth-integrated models may be used at least to describe the free surface fluid evolution like in [103], where floating breakwater are modelized using a Boussinesq model and a Finite-Difference (FD) scheme. The flow under the breakwater is regarded as a confined flow and the pressure field beneath the floating structure is determined by solving implicitly the Laplace equation. Asymptotic flow models may be applied also to describe the flow under the floating structure, see [40] where the Kadomtsev-Petviashvili (KP) equations are used to compute wave generation by ships in shallow-water, or [95, 180, 96] where Boussinesq equations are applied to model the interactions in the near-ship flows. Let also mention the recent numerical study [24], where a Boussinesq model is applied to compute the heave (vertical) motion of structures with straight-sided boundaries, which are assumed vertical at the fluid-structure contact line.
Recently, a new formulation of the fully nonlinear floating body problem has been introduced in [107], describing the flow with respect to the free surface parameterization and the horizontal discharge instead of the velocity potential, and the particular assumptions/simplifications leading to depth-integrated free surface models with floating-body are detailed, with a particular emphasize put on the simple elliptic equation solved by the pressure of the fluid under the body.

Focusing on shallow-water flows

In what follows, we focus on the *shallow-water* (or long-wave) regime:

$$(\textit{shallow-water regime}) \qquad \mu := \frac{H_0^2}{\lambda^2} \ll 1, \tag{1.1}$$

where H_0 refers to the typical water depth and λ the typical wave length of the flow. In this regime, the hyperbolic Nonlinear shallow-water (NSW) equations [50] can be derived from the full water waves equations by neglecting all the terms of order $\mathcal{O}(\mu)$ and greater, see for instance [106]. It is also worth noticing that no smallness assumption is made on the size of the surface perturbations for this derivation, and the corresponding regime is also said to be *fully nonlinear*:

$$(\textit{fully nonlinear / large amplitude regime}) \qquad \varepsilon := \frac{a}{H_0} = \mathcal{O}(1), \tag{1.2}$$

where a is the typical wave's amplitude. Numerous asymptotic models may be obtained from the water-wave equations with a free-surface boundary condition. Among them, the NSW equations are one of the most popular model for the description of shallow-water flows when the dispersive effects can be neglected and they are extensively used for the study of geophysical flows and coastal engineering problems related to wave propagation and transformations. Given a smooth parametrization

of the topography variations $b : \mathbb{R} \to \mathbb{R}$, and denoting by H the water-height, u the horizontal (depth-averaged) velocity, $q = Hu$ the horizontal discharge, see Fig.3.1, the NSW equations may be written as follows:

$$\partial_t H + \partial_x q = 0, \tag{1.3a}$$

$$\partial_t q + \partial_x \big(uq + \frac{1}{2} g H^2 \big) = -gH \partial_x b. \tag{1.3b}$$

A derivation of this system is provided in Appendix A for the sake of completeness. As H stands for the water-height, the flow main variables (H, q) should belong to the following convex set of physical admissible states Θ, defined as follows:

$$\Theta = \big\{ (H, q) \in \mathbb{R}^2;\ H \geq 0 \big\}. \tag{1.4}$$

For practical applications, these NSW equations may be supplemented with several additional source terms, depending on the leading physical processes at stake, modeling wind forcing, rainfall contributions, bottom friction, Coriolis effect or eddy viscosity. The interested reader may find several supplemented models in [51, 52, 93, 175, 41, 122, 6, 21, 25, 82, 120]. Also, the NSW equations do not account for dispersive effects and as a consequence, they can't describe the weakly dispersive processes that generally occur in nearshore areas, like wave shoaling for instance. To achieve this, more accurate models like the Boussinesq-type (BT) equations [26] or the Green-Naghdi (GN) equations [76] should be derived, by keeping the terms of order μ (and even beyond for higher-order models). However, those "augmented" NSW equations, which are also called *weakly dispersive shallow-water asymptotics* are much more complex than the classical NSW equations and they generally require some considerable efforts on both theoretical and numerical sides to produce reliable forecasting. Hence, in what follows, we choose to focus on the classical (hyperbolic) NSW system with a topography source term as a valuable "starting' model to describe the flow evolution.

There are however very few studies in the literature which are devoted to the possible extension of the NSW equations to embed a floating structure. Yet, such an approach appears as very promising, allowing to overcome the heavy computational cost of RANS, while accounting for the nonlinearity of the physical processes at stake. We can mention the recent studies [72, 73] for the computation of congested shallow-water flows with a compressible/incompressible projection scheme, and also [24] as the dispersive effects of the chosen BT equations are actually neglected in the vicinity of the floating structure.

Hence, the main purpose of this Ph.D. is to develop a new numerical strategy for the modeling and simulation of the evolution of a floating object in a shallow-water flow described by the NSW equations.

In this respect, we choose to start from the very recent theoretical work [87], in which a general theory for a class of quasi-linear hyperbolic IBVP with free boundaries is introduced. This theory is further applied to the modeling of a partially immersed floating structure in a shallow-water flow described by the NSW equations, providing a convenient theoretical ground for our purpose.

1.2 Numerical ingredients

The design of our new numerical strategy relies on several ingredients and methods which are briefly described in the remainder of this section. We only provide some insights of the salient features of these numerical aspects, together with some references, and we leave the details of the recipe for the next Chapters.

Numerical methods for the NSW equations

As already mentioned, the NSW equations are one of the most widely used set of equations for simulating long wave hydrodynamics. Considering their hyperbolic (and hydrostatic) nature, they generally provide a reliable description of steep-fronted flows, such as dam-breaks or flood-waves. This model is also extensively used in coastal engineering, for the study of nearshore flows involving bores propagation in the surf zone, run-up and run-down on sloping beaches or coastal structures and to forecast coastal inundations. To allow a proper description of such phenomena, accurate and robust numerical methods should be considered. Great efforts have been made since the sixties in order to produce accurate approximations of weak solutions of the NSW equations and a large variety of numerical methods have been developed, including Finite-Volumes (FV) [3, 67, 7, 14, 171, 63], Finite-Elements (FE) [129, 157, 136, 11], spectral methods [89, 121, 133] or residual distribution methods [144, 143, 8]. Among these numerical strategies, the Godunov-type FV methods are particularly praised, thanks to their low computational cost and their shock-capturing ability, which allows to preserve the discontinuous or steeply varying gradients that may occur in sharp-fronted and trans-critical shallow-water flows, see for instance [145, 7, 156, 52, 131, 18, 116] among others and also some references herein. Many of them particularly focus on the issue of balancing the flux gradient and the topography source term [9, 130, 68, 119, 116, 32, 100, 35, 125]. However, FV methods usually offer low accuracy and one generally needs to use some reconstruction methods to offset the low order of convergence and the diffusive losses, see for instance [94, 109, 138, 126, 19].

Discontinuous Galerkin methods

In what follows, we use the *discontinuous Galerkin* (DG) method to approximate the solutions of the NSW equations. This choice is mainly motivated by the numerous assets of this family of methods. Indeed, the possibility of reaching an optimal (possibly high-) order of accuracy where the solution is smooth enough is a major concern in the design of discrete formulations for transport flow problems, and the development of high-order methods and their application for solving real-world problems is a very active research topic in computational mechanics. In this context, although DG methods have existed in various forms for more than 45 years, they have experienced a vigorous development over last 25 years. The first DG method to approximate first-order PDEs has been introduced by Reed and Hill in 1973 [142] in the framework of steady neutron transport (i.e. a time independent linear hyperbolic equation), while the first analysis for steady first-order PDEs was presented by Lesaint and Raviart in 1974 [110, 111]. The error estimate was improved by Johnson and Pitkäranta in 1986 [98] who set up an order of convergence of $k + \frac{1}{2}$ in the L^2 norm for a $\mathbb{P}^k$ polynomial approximation of degree k with a smooth enough exact solution. In 1990, the method was further developed by Caussignac and Touzani [36, 37] to approximate the three-dimensional boundary-layer equations for incompressible steady fluid flows. During this period (1989-1991), DG methods were extended to time-dependent hyperbolic PDEs by Chavent and Cockburn [38] using the forward Euler scheme for time discretization together with limiters. In the same years, an improvement for time discretization

schemes was introduced by Cockburn and Shu [46, 47], by using explicit Runge-Kutta (RK) schemes, improving finally the order of accuracy. A convergence proof to the entropy solution was obtained by Jaffré, Johnson, and Szepessy [92] in 1995. Extensions where presented in a series of papers by Cockburn, Shu, and coworkers [43, 46, 44]. Nowadays, DG methods are widely used in several large classes of problems, in fluid dynamics, geophysical flows, aero-acoustics or electromagnetism for instance.
Such a success may be explained by the fact that DG methods combine the background of FE methods, FV methods and Riemann solvers, allowing to take into account the physic of the problem, and they have been successfully validated in many domains of application. Indeed, on one hand a DG method can be seen as a FE method allowing for discontinuities in the discrete trial and test spaces, localizing test functions to single mesh elements and introducing numerical fluxes at cells boundaries. On the other hand, the sought solution is only smooth inside each element and like in FV methods, the solution at cells interfaces is not uniquely defined. The interface fluxes are thus approached by some suitable numerical fluxes involving the jumps of the solution at the interface, and allowing to weakly enforce the coupling conditions of the discrete solution. Hence, DG methods can also be seen as FV methods in which the approximate solution is represented by piecewise-polynomial functions and not only by piecewise-constant values. Of course, the design of the interface fluxes is not trivial, since it is closely related to the consistency, conservation, stability and accuracy of the resulting scheme.
The generally acknowledged assets of DG schemes are the following: (i) any order of polynomial approximation can be used within elements, allowing to possibly reach some high-order of accuracy depending on the solution's regularity, (ii) the decoupling of the system of equations: the matrices involved in the linear system to be solved are sparse and structured by blocks which are dimensioned by the number of degrees of freedom in each mesh element, (iii) the size of the stencil is independent of the order of precision: the computation of the discrete residual depends only on the solution in the element and its first neighbors. This appealing feature also allows for some easier parallel computation, (iv) the boundary conditions are weakly enforced, through the numerical fluxes, without modifying the definition of the approximation space like in conforming FE methods, allowing a simplified implementation, (v) working with discontinuous discrete spaces offers a substantial amount of flexibility, making the approach appealing for multi-domain and multi-physics simulations, (vi) the sensitivity to the regularity of the mesh is weak, thanks to the discontinuity of the solution between elements. This point allows, not only the adequacy with unstructured and non-conforming meshes to represent the industrial geometries, but also the development of refinement, coarsening or moving grid strategies. For instance, it is possible to combine a mesh refinement (h-adaptation) in the areas of low regularity of the solution with an increase in the order of approximation (p-adaptation) where the solution is regular enough.
Several DG methods have been designed for the NSW equations since the early 2000s, see for instance [147, 154, 112, 174, 2, 5, 99, 16, 128, 147, 71, 65, 177, 176, 101] and some references hereafter.
In particular, the choice of using DG methods was also motivated by some of the previous works of my Ph.D. advisors [62], and the availability of a locally developed and maintained arbitrary-order DG C++ solver for the NSW equations (called WaveBox), which was provided as a starting computational code for this Ph.D.

DG suffers: *a priori* and *a posteriori* stabilization

However, while DG methods may be mature enough to accurately handle some realistic nonlinear problems in various applications, they originally suffer from the lack of nonlinear stability. In partic-

ular, high-order DG methods may produce spurious oscillations in the presence of discontinuities or steeply varying gradients (*i.e.* Gibbs phenomenon), potentially leading to overshoots and unphysical solutions. Also, focusing on the NSW equations, another challenging issue is the preservation of the set of admissible states (1.4) at the discrete level, which is closely related to the issue of the occurrence and propagation of wet/dry fronts that may occur in dam-breaks, flood-waves or run-up over coastal shores. Hence, while a minimal nonlinear stability requirement is to preserve the water-height positivity at the discrete level, this is clearly a challenging purpose when high-order polynomials are used within mesh elements and standard (non-stabilized) DG methods may produce negative values for the water-height H in the vicinity of dry areas.

Generally speaking, robustness issues may be among the main remaining challenges for the use of high-order methods in realistic problems for many domains of applications, and in recent years, several approaches have been proposed to stabilize high-order approximations. These techniques mainly rely on two different paradigms that we referred to as *a priori* and *a posteriori*. In the so-called *a priori* framework, the correction procedure is applied before advancing the numerical piecewise polynomial solution further in time. So first, a troubled zone indicator is used to find where a correction is required (see [140] for a review of such *troubled elements* sensors). Then, sufficient efforts are made on the numerical solution or on the numerical scheme to be sure that one is able to carry the computation out to the next time-step. Among others *a priori* correction techniques, we could mention *artificial viscosity* methods [134, 159, 66, 77, 102], where some dissipative mechanism is added in shock regions, borrowing ideas from the streamline upwind Petrov Galerkin (SUPG) and Galerkin least-squares methods. Some other very popular limiting techniques can be gathered and referred to as slope and moment limiters [44, 20, 29, 105, 181, 91, 57, 113]. In the former ones, as in [46, 44], the polynomial approximated solution is flattened around its mean-value to control the solution jumps at cell interfaces. A smooth extrema detector is then generally used to prevent the limitation technique to spoil the accuracy in regions where no limiting is required. Moments limiters, mainly based on [20] and further developed in [29], can be seen as the extension of the aforementioned slope-limiters to the case of very high-orders of accuracy. In those limiting strategies, the different moments of the polynomial solution are successively scaled in a decreasing sequence, from the higher degree to the lower one, allowing the preservation of the solution accuracy, as well as ensuring the solution boundedness near discontinuities. The high-order DG limiter [105], generalized moment limiter [181], hierarchical Multi-dimensional Limiting Process (MLP) [91, 90] and vertex-based hierarchical slope-limiters [57, 113] all derive from [46, 20, 29], and thus fall into this category. Now, another limiting strategy that deserves to be mentioned is the (H)WENO limiting procedure [141, 10, 187, 114, 188], where the DG polynomial is substituted in troubled regions by a reconstructed (H)WENO polynomial. An alternative way to treat this spurious oscillations issue may be to use a *solution filtering* method, see for instance [160, 153, 127, 135], which aim at removing high wave-number oscillations. Those filtering procedures are generally done in an *ad hoc* fashion, filtering being applied "as little as possible, but as much as needed". Last but not least, some original FV-Subcell shock capturing techniques in the frame of DG schemes [84, 34, 151, 49] have recently gained in popularity. In [84], the authors use a convex combination between high-order DG schemes and first-order FV on a sub-grid, allowing them to retain the very high accurate resolution of DG in smooth areas and ensuring the scheme robustness in the presence of shocks. Similarly, in [151, 49], after having detected the troubled zones, cells are then subdivided into subcells, and a robust first-order FV scheme is performed on the sub-grid in troubled cells.

The *a priori* paradigm has already and extensively proved in the past its high capability and feasibility, as in the aforementioned articles. Those techniques are *a priori* in the sense that only the

data at time t^n are needed to perform the limitation procedure. Then, the limited solution is used to advance the numerical scheme in time to t^{n+1}. The "worst case scenario" has to be generally considered as a precautionary principle. Furthermore, let us emphasize that most of the *a priori* correction procedures previously quoted do not ensure a maximum principle or the positivity-preservation of the solution. Generally, additional effort has to been made specifically on that matter, as for example by means of positivity-preserving limiters [183, 185]. Specifically concerning the issue of positivity preservation in DG methods for the NSW equations, various *a priori* strategies have been introduced recently: a free-boundary treatment in mixed Eulerian-Lagrangian elements is introduced in [22] to locate the wet/dry interface, a fixed mesh method with a local conservative slope modification technique based on a redistribution of the fluid and cut-off in discharge is presented in [64], local first moment limitations without mass adding in [28, 100, 99, 161], high-order accuracy *a priori* polynomial reconstruction and limitation to enforce a strict maximum principle on mean-values in [179, 178], in [62] for the so-called *pre-balanced* formulation of the NSW equations, or in [123] for a formulation with implicit time-stepping. Let us finally mention [124] where an *a priori* FV-Subcell approach has been adopted.

Now, the paradigm of *a posteriori* correction is different in the way that first an uncorrected candidate solution is computed at the new time-step. The candidate solution is then checked according to some criteria (for instance positivity, discrete maximum principle, ...). If the solution is considered admissible, we go further in time. Otherwise, we return to the previous time-step and correct locally the numerical solution by making use of a more robust scheme. Because the troubled zone detection is performed *a posteriori*, the correction can be done only where it is absolutely necessary. Furthermore, let us emphasize that in *a posteriori* correction procedures, the maximum principle preservation or positivity preservation is included without any additional effort. Indeed, the whole procedure is positivity-preserving as soon as the numerical scheme used as a correction procedure is. Consequently, all the *a posteriori* techniques that make use of FV scheme as correction method is then positivity-preserving. Recently, some new *a posteriori* limitations have arisen. Let us mention the so-called MOOD technique, [42, 55, 56]. Through this procedure, the order of approximation of the numerical scheme is locally reduced in an *a posteriori* sequence until the solution becomes admissible. In [61, 59, 88], a FV-Subcell technique similar to the one presented in [151] has been applied to the *a posteriori* paradigm. Practically, if the numerical solution in a cell is detected as bad, the cell is then subdivided into subcells and a first-order FV, or alternatively other robust scheme (second-order TVD FV scheme, WENO scheme, ...), is applied on each subcell. Then, through these new subcell mean-values, a high-order polynomial is reconstructed on the primal cell. Related strategies applied to dispersive and turbulent shallow-water flows have been introduced in [30, 31].

Nonetheless, in all the aforementioned limitation techniques, *a priori* and *a posteriori*, in the troubled cells the high-order DG polynomial is either globally modified in the cell, or even discard as it is in the (H)WENO limiter or any *a posteriori* correction procedure. One of the main advantage of high-order scheme is to be able to use coarse grids while still being very precise. But even in the case where the troubled zone, as the vicinity a shock for instance, is very small regarding the characteristic length of a cell, the DG polynomial is globally modified. In [164], a new conservative technique is introduced to overcome this issue, by modifying the DG numerical solution only locally at the subcell scale. This correction procedure has been designed first to avoid the occurrence of non-admissible solution, to be maximum principle preserving and to prevent the code from crashing.

Another goal of this Ph.D. is to extend and adapt this new subcell correction strategy to the NSW

equations and equip the existing arbitrary-order DG formulation for the NSW equations with a new, robust and accurate stabilization operator with the additional constraint that the well-balancing property for motionless steady states should be ensured.

We choose to call this method *a posteriori* Local Subcell Correction (LSC). The purpose of the LSC operator is of course to enforce the water-height positivity and avoid spurious oscillations in the vicinity of solution's singularity. But the resulting corrected scheme is also conservative at the subcell level. Additionally, it allows us to retain as much as possible the high accuracy and subcell resolution of DG schemes, by minimizing the number of subcells in which the solution has to be recomputed. Practically, the correction procedure only modifies the DG solution in troubled subcell regions without impacting the solution elsewhere in the cell. It is also worth mentioning that the whole procedure is totally parameter free, and behaves properly from 2nd order to any order of accuracy.

Free-boundary problems and Arbitrary Lagrangian-Eulerian description

From a numerical point of view, the study of flows with a free moving boundary is a very difficult problem, which may be encountered in engineering domains like aeroelasticity or fluid-structure interactions. On one hand, Immersed or Embedded Boundary Methods (IBM or EBM), which are classically constructed in the Eulerian setting, are particularly attractive for complex fluid–structure interaction problems characterized by large structural motions and deformations and for flow problems with topological changes and / or with cracking, as they allow to directly embed some material boundaries into the computational domain. We refer the reader to [139] for a general description of the IBM and to [150] for a recent work concerning the closely related *shifted boundary* method.
On the other hand, to avoid interface-tracking methods, one may require the formulation to handle moving or deforming domains and related meshes, following the time-evolution of the fluid-structure interface. In such a context, the Arbitrary Lagrangian-Eulerian (ALE) description appears as a popular and convenient choice for flow problems involving time-varying boundaries. Additionally, it is important to highlight that the interactions of flows with moving boundaries may also result in additional unsteady phenomena, generally coming with the need of high-order accurate approximations to resolve the unsteadiness of flows at various scales and correctly predict and model, for instance, the conditions at which some kind of instabilities may occur.
Initially developed in combination with a FD discretization in [83], and later extended to FE and FV methods for both fluids and structures, see for instance [58] for a review, the ALE description is generally put forward as combining the best of both Lagrangian (-*material* domain) and Eulerian (-*spatial* domain) worlds. In the Eulerian framework, the conservation laws governing the physical phenomenon under consideration are developed on a fixed referential, while in the Lagrangian formalism the referential is attached to the material. Thus, in the case of hydrodynamic problem for instance, the mesh move and get deformed as the fluid flows [168, 169, 166]. The ALE methods lies in between, where the computational mesh can move with an arbitrary velocity, which may be chosen independently from the material (fluid in our case) velocity. This provides some great flexibility in handling moving domains, avoiding the issues usually associated with the tracking of interfaces in the Eulerian approach, as well as the large distortion generally encountered in the Lagrangian framework when (not so) large time evolutions are considered. It may also be important to distinguish between indirect and direct ALE method. The indirect ALE methods consist in a purely Lagrangian phase, followed by re-meshing and projection phases [117, 17]. The direct ALE methods are different in the way that they take directly into account the mesh displacement in the

flux definition of the discretized system of equations, see for instance [60, 23], which make them particularly appealing for the problem at stake here, by potentially allowing to propagate the grid deformation induced by the displacement of the moving floating object at the water surface.
Within ALE simulations of flow problems with moving boundaries, it is also important to ensure that a numerical scheme reproduces exactly a constant solution. The Geometric Conservation Law (GCL) is a relation between the ALE mapping's Jacobian and the mesh velocity, which simply states that a uniform flow field should not be influenced by any arbitrary grid's motion. The notion of GCL was first introduced in [155] and is also discussed for instance in [118, 137] and [78] where relations between GCL and time stability are investigated. Of course, ALE descriptions may be also conveniently applied together with a DG discretization method and such DG-ALE numerical strategies for the study of moving boundaries in fluid-structure interactions or free-surface flows have been introduced for instance in [162, 118, 137], see also the related *space-time* DG methods of [159, 158].

Wave-structure interactions: a DG-ALE-LSC discrete formulation

In the second part of this work, we design a new robust high-order DG-ALE discrete formulation which is directly modeled from the class of IBVP introduced and analyzed in [87]. This provides a new way of simulating adaptive solutions for floating structures in nonlinear shallow-water flows. The wetted surface and contact points are of course expected to vary over time following the body motion, and the computational grid is expected to move accordingly. To achieve this, a continuous explicit mapping between a fixed reference configuration and a time-varying domain is constructed, taking inspiration from the analysis of the continuous problem in [87], and the NSW equations are recast in the reference domain with the introduction of an additional geometric term, before being approximated through high-order discontinuous polynomial interpolation. We extend the *a posteriori* LSC method to the proposed DG-ALE framework and enforce some nonlinear stability and monotonicity, that are minimal requirements for the high-order approximations of nonlinear flows with floating objects, thus leading to what we call a DG-ALE-LSC formulation. This stabilization procedure through corrected fluxes is successfully combined with some suitable local conservative variables reconstructions borrowed from [116], and a definition of the Lax-Friedrichs interface flux adapted to moving meshes, ensuring that robustness and well-balancing (for motionless steady states) are embedded properties of the limit lowest-order scheme. It is also proved that the global stabilized DG-ALE formulation proposed in this thesis naturally ensures such a property, both at the semi-discrete level (GCL) and the fully discrete level with the Discrete GCL (DGCL), hence successfully combining well-balancing with geometric conservation.

The resulting modeling strategy is then applied and validated through a battery of benchmarks involving different kind of configurations: (i) stationary partially immersed structures: the position of the structures is fixed, with no motion over time, leading to a simpler model in the interior domain, under the structure, (ii) prescribed structures's motion: an operator enforces the body to move according to some prescribed motion's laws, impacting the flow motion around the structure and leading to a more complex computational problem to be solved beneath the structure, (iii) free floating structure: the structure moves according to the forces and torque deduced from the flow and reversely, the structure's motion impacts the flow configuration. In this last case, additional equations describing the dynamics of the floating structure should be considered. We emphasize that the structure is allowed for a vertical (heaving), horizontal (surging) and rotational (pitching) motion and its shape's lateral boundaries are not necessarily vertical.

1.3 Discrete settings and basic formulations

In this section, we introduce and define several mathematical objects, tools and notations related to discretization, that are extensively used in the next Chapters. Then, we state a straightforward (non-stabilized) DG formulation for the NSW equations, as a starting point for the upcoming new materials.

Discrete setting for DG methods

Let $\Omega \subset \mathbb{R}$ denote an open segment with boundary $\partial\Omega$, which serves as the computational domain. We consider a partition $\mathscr{T}_h = \{\omega_1, \ldots, \omega_{n_{\text{el}}}\}$ of Ω in open disjoint segments ω of boundary $\partial\omega$ such that $\overline{\Omega} = \bigcup_{\omega\in\mathscr{T}_h} \overline{\omega}$. The partition is characterized by the mesh size $h := \max_{\omega\in\mathscr{T}_h} h_\omega$, where h_ω is the length of element ω. For a given mesh element $\omega_i \in \mathscr{T}_h$, we also note $\omega_i = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$ and by x_i its barycenter.

Given an integer polynomial degree $k \geq 1$, we consider the broken polynomial space

$$\mathbb{P}^k(\mathscr{T}_h) := \left\{ v \in L^2(\Omega),\ \ v_{|\omega} \in \mathbb{P}^k(\omega),\ \ \forall\, \omega \in \mathscr{T}_h \right\},$$

where $\mathbb{P}^k(\omega)$ denotes the space of polynomials in ω of total degree at most k, with $\dim(\mathbb{P}^k(\omega)) = k+1$. Piecewise polynomial functions belonging to $\mathbb{P}^k(\mathscr{T}_h)$ are denoted with a subscript h in the following, and for any $\omega \in \mathscr{T}_h$ and $v_h \in \mathbb{P}^k(\mathscr{T}_h)$, we may use the convenient shortcut: $v_h^\omega := v_{h|\omega}$ when no confusion is possible.

For any mesh element $\omega \in \mathscr{T}_h$ and any integer $k \geq 0$, we consider a basis for $\mathbb{P}^k(\omega)$ denoted by

$$\Psi_\omega = \left\{\psi_j^\omega\right\}_{j\in[\![1,\,k+1]\!]}.$$

A basis for the global space $\mathbb{P}^k(\mathscr{T}_h)$ is obtained by taking the Cartesian product of the basis for the local polynomial spaces:

$$\Psi_h = \bigtimes_{\omega\in\mathscr{T}_h} \Psi_\omega = \left\{ \left\{\psi_j^\omega\right\}_{j\in[\![1,\,k+1]\!]} \right\}_{\omega\in\mathscr{T}_h}.$$

Note that we have:

$$\operatorname{supp}(\psi_j^\omega) \subset \overline{\omega}, \quad \forall\omega \in \mathscr{T}_h, \quad \forall j \in [\![1,\, k+1]\!].$$

We introduce the following shortcut notations for smooth enough scalar-valued functions v, w:

$$\int_{\mathscr{T}_h} v(x)w(x)dx := \sum_{\omega\in\mathscr{T}_h} \int_\omega v(x)w(x)dx,$$

$$\left[v\right]_{\partial\omega_i} := v(x_{i+\frac{1}{2}}) - v(x_{i-\frac{1}{2}}), \quad \forall\omega_i \in \mathscr{T}_h.$$

For $\omega \in \mathscr{T}_h$, we denote p_ω^k the L^2-orthogonal projector onto $\mathbb{P}^k(\omega)$ and $p_{\mathscr{T}_h}^k$ the global L^2-orthogonal projector onto $\mathbb{P}^k(\mathscr{T}_h)$ that gather all the local L^2-projectors p_ω^k on each element. Similarly, we denote i_ω^k the element nodal interpolator into $\mathbb{P}^k(\omega)$. The corresponding nodal distributions in elements are chosen to be the approximate optimal nodes introduced in [39], which have better approximation

properties than equidistant distributions, and include, for each element, the elements boundaries into the interpolation nodes. The global $i^k_{\mathscr{T}_h}$ interpolator into $\mathbb{P}^k(\mathscr{T}_h)$ is obtained by gathering the local interpolating polynomials defined on each element.

We also define the broken gradient operator $\partial_x : \mathbb{P}^k(\mathscr{T}_h) \to \mathbb{P}^k(\mathscr{T}_h)$ such that, for all $v_h \in \mathbb{P}^k(\mathscr{T}_h)$,

$$(\partial_x v_h)_{|\omega} := \partial_x(v_{h|\omega}) = \partial_x v_h^\omega, \quad \forall\, \omega \in \mathscr{T}_h.$$

Remark 1. The degrees of freedom are classically chosen to be the functionals that map a given discrete unknown belonging to $\mathbb{P}^k(\mathscr{T}_h)$ to the coefficients of its expansion in the selected basis. Specifically, the degrees of freedom applied to a given function $v_h \in \mathbb{P}^k(\mathscr{T}_h)$ return the real numbers

$$\underline{v}_j^\omega \quad \text{with} \quad j \in [\![1, k+1]\!] \text{ and } \omega \in \mathscr{T}_h, \tag{1.5}$$

such that

$$v_h^\omega = \sum_{j=1}^{k+1} \underline{v}_j^\omega \psi_j^\omega, \quad \forall \omega \in \mathscr{T}_h.$$

With a little abuse in terminology, we refer hereafter to the real numbers (1.5) as the *degrees of freedom* associated with v_h and we note $\underline{v}_\omega \in \mathbb{R}^{k+1}$ the vector gathering the degrees of freedom associated with v_h^ω.

Discrete setting for FV-Subcell Correction methods

For any mesh element $\omega_i \in \mathscr{T}_h$, we introduce a sub-partition $\mathscr{T}_{\omega_i}$ into $k+1$ open disjoint subcells:

$$\overline{\omega_i} = \bigcup_{m=1}^{k+1} \overline{S}_m^{\omega_i},$$

where the subcell $S_m^{\omega_i} = \left[\widetilde{x}^{\omega_i}_{m-\frac{1}{2}}, \widetilde{x}^{\omega_i}_{m+\frac{1}{2}}\right]$ is of size $\left|S_m^{\omega_i}\right| = \left|\widetilde{x}^{\omega_i}_{m+\frac{1}{2}} - \widetilde{x}^{\omega_i}_{m-\frac{1}{2}}\right|$, with the convention $\widetilde{x}^{\omega_i}_{\frac{1}{2}} = x_{i-\frac{1}{2}}$ and $\widetilde{x}^{\omega_i}_{k+\frac{3}{2}} = x_{i+\frac{1}{2}}$, see Fig. 1.1. When considering a sequence of neighboring mesh elements $\omega_{i-1}, \omega_i, \omega_{i+1}$, the convenient convention $S_0^{\omega_i} := S_{k+1}^{\omega_{i-1}}$ and $S_{k+2}^{\omega_i} := S_1^{\omega_{i+1}}$ may be used.

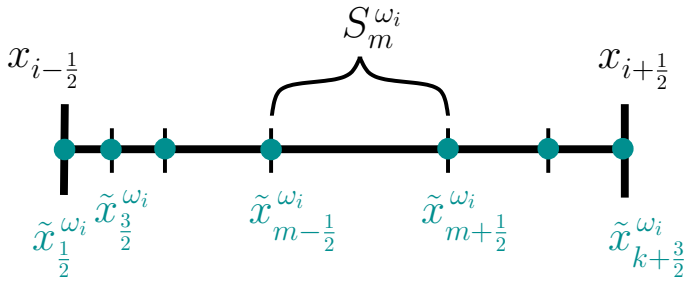

Figure 1.1: Partition of a mesh element ω_i in $k+1$ subcells

To define the *sub-resolution* basis functions, required in § 3.2 and initially introduced in [164], we introduce for a given mesh element $\omega \in \mathscr{T}_h$ the following set of *subcell indicator* functions $\left\{\mathbb{1}_m^\omega,\ \ m \in [\![1,\, k+1]\!]\right\}$, with:

$$\mathbb{1}_m^\omega(x) = \left\{ \begin{array}{ll} 1 & \text{if } x \in S_m^\omega, \\ 0 & \text{if } x \notin S_m^\omega, \end{array} \right. \qquad \forall m \in [\![1,\, k+1]\!].$$

Then, the set of *sub-resolution* basis functions $\left\{\phi_m^\omega \in \mathbb{P}^k(\omega),\ \ m \in [\![1,\, k+1]\!]\right\}$ are defined as follows:

$$\phi_m^\omega = p_\omega^k(\mathbb{1}_m^\omega), \quad \forall m \in [\![1,\, k+1]\!], \tag{1.6}$$

$$\int_\omega \phi_m^\omega \varphi dx = \int_\omega \mathbb{1}_m^\omega \varphi dx = \int_{S_m^\omega} \varphi dx, \quad \forall m \in [\![1,\, k+1]\!], \quad \forall\ \varphi \in \mathbb{P}^k(\omega). \tag{1.7}$$

For all $\omega \in \mathscr{T}_h$ we also introduce the set of piecewise constant functions on the sub-grid:

$$\mathbb{P}^0(\mathscr{T}_\omega) := \left\{ v \in L^2(\omega),\ \ v_{|S_m^\omega} \in \mathbb{P}^0(S_m^\omega),\ \ \forall\, S_m^\omega \in \mathscr{T}_\omega \right\}.$$

For any $\omega \in \mathscr{T}_h$, and any $v_h^\omega \in \mathbb{P}^k(\omega)$, let denote

$$\overline{v}_m^\omega \quad \text{with} \quad m \in [\![1,\, k+1]\!],$$

the low-order piecewise constant components defined as the mean-values of v_h^ω on the subcells belonging to the subdivision $\mathscr{T}_\omega$, called *sub-mean-values* in the following, which may be gathered in a vector $\overline{v}_\omega \in \mathbb{R}^{k+1}$. Whenever a sequence of neighboring mesh elements $\omega_{i-1}, \omega_i, \omega_{i+1}$ and associated neighboring approximations is considered, the following convenient convention may be used: $\overline{v}_0^{\omega_i} := \overline{v}_{k+1}^{\omega_{i-1}}$ and $\overline{v}_{k+2}^{\omega_i} := \overline{v}_1^{\omega_{i+1}}$.

Remark 2. We observe that the degrees of freedom $\left\{\underline{v}_m^\omega, \quad m \in [\![1,\, k+1]\!]\right\}$ are uniquely defined through the sub-mean-values $\left\{\overline{v}_m^\omega, \quad m \in [\![1,\, k+1]\!]\right\}$, and reversely. Specifically, considering the local transformation matrix $\Pi_\omega = \left(\pi_{m,p}^\omega\right)_{m,p}$ defined as:

$$\pi_{m,p}^\omega = \frac{1}{|S_m^\omega|} \int_{S_m^\omega} \psi_p^\omega \mathrm{dx}, \qquad \forall\, (m,p) \in [\![1,\, k+1]\!]^2, \tag{1.8}$$

we have the following relation:

$$\mathbf{\Pi}_\omega\ \underline{v}_\omega = \overline{v}_\omega \quad \text{and} \quad \mathbf{\Pi}_\omega^{-1}\ \overline{v}_\omega = \underline{v}_\omega. \tag{1.9}$$

As a consequence, any polynomial function $v_h^\omega \in \mathbb{P}^k(\omega)$ can be expressed equivalently either in terms of the degrees of freedom $\underline{v}_\omega$, or the sub-means values $\overline{v}_\omega$.

Finally, let introduce the (one-to-one) following projector onto the piecewise constant sub-grid space:

$$\pi_{\mathscr{T}_\omega}^k : \mathbb{P}^k(\omega) \longrightarrow \mathbb{P}^0(\mathscr{T}_\omega) \tag{1.10}$$

$$v_h^\omega \longmapsto \pi_{\mathscr{T}_\omega}^k(v_h^\omega) := \overline{v}_\omega. \tag{1.11}$$

In practice, once the transformation matrices $\mathbf{\Pi}_\omega$ are initialized in a preprocessing step, it is straightforward and computationally inexpensive to switch from one representation to another, see Fig. 1.2.

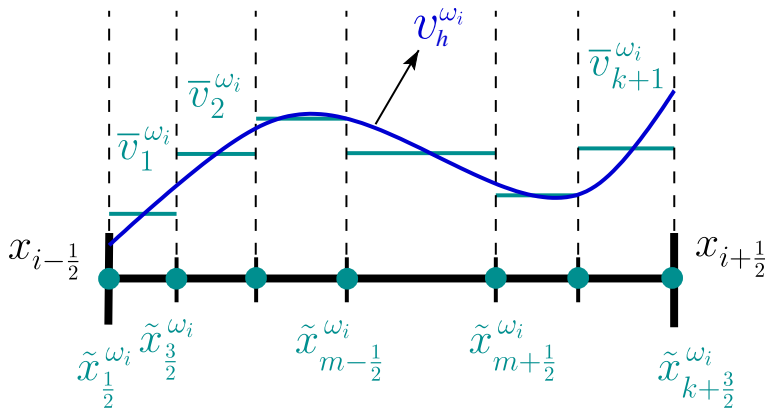

Figure 1.2: Piecewise polynomial function and associated sub-mean-values

Remark 3. The local projection matrices (1.8) are obviously non-singular. This property would be straightforwardly extended to the multi-dimensional case with Cartesian grids.

Time discretization

Concerning time discretization, for a given final computational time $t_{\max} > 0$, we consider a partition $(t^n)_{0\leq n\leq N}$ of the time interval $[0, t_{\max}]$ with $t^0 = 0$, $t^N = t_{\max}$ and $t^{n+1} - t^n =: \Delta t^n$. More details on the computation of the time-step Δt^n and on the time marching algorithms are given in § 3.3. For any sufficiently regular scalar-valued function of time w, we let $w^n := w(t^n)$.

A straightforward DG formulation for the NSW equations

The NSW equations may be written in a compact form:

$$\partial_t v + \partial_x F(v) = B(v, \partial_x b). \tag{1.12}$$

where $v : \mathbb{R}\times\mathbb{R}_+ \to \Theta$ is the vector of conservative variables, $F : \Theta \to \mathbb{R}^2$ is the flux function and $B : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the topography source term, defined as follows:

$$v = \begin{pmatrix} H \\ q \end{pmatrix}, \; F(v) = \begin{pmatrix} q \\ uq + \frac{1}{2}gH^2 \end{pmatrix}, \quad B(v, \partial_x b) = \begin{pmatrix} 0 \\ -g\, H\, \partial_x b \end{pmatrix}. \tag{1.13}$$

We multiply (1.12) by an arbitrary function $\psi \in \mathbb{P}^k(\mathscr{T}_h)$ and we integrate locally on the cell ω_i:

$$\int_{\omega_i} \psi \partial_t v \mathrm{d}x + \int_{\omega_i} \psi \partial_x F(v) \mathrm{d}x = \int_{\omega_i} \psi B(v, \partial_x b) \mathrm{d}x. \tag{1.14}$$

The basis function ψ is not time-dependent, thus we can rewrite equation (1.14) as:

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i} v\psi \mathrm{d}x + \int_{\omega_i} \psi \partial_x F(v) \mathrm{d}x = \int_{\omega_i} \psi B(v, \partial_x b) \mathrm{d}x, \tag{1.15}$$

an integration by parts leads to:

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\omega_i} v\psi \mathrm{d}x - \int_{\omega_i} F(v)\partial_x\psi \mathrm{d}x + \left[\psi F(v)\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \int_{\omega_i} \psi B(v,\partial_x b)\mathrm{d}x. \tag{1.16}$$

Let consider the restrictions of the sought solution and the bathymetry to the mesh element ω_i, which writes:

$$v_h^{\omega_i} = \sum_{p=1}^{k+1} \underline{v}_p^{\omega_i}\psi_p^{\omega_i} \quad \text{and} \quad b_h^{\omega_i} = \sum_{p=1}^{k+1} \underline{b}_p^{\omega_i}\psi_p^{\omega_i}, \tag{1.17}$$

then the DG local formulation writes:

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\omega_i} v_h^{\omega_i}\psi_l^{\omega_i}\mathrm{d}x - \int_{\omega_i} F(v_h^{\omega_i})\partial_x\psi_l^{\omega_i}\mathrm{d}x + \left[\psi_l^{\omega_i}\mathcal{F}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \int_{\omega_i} \psi_l^{\omega_i} B(v_h^{\omega_i},\partial_x b_h^{\omega_i})\mathrm{d}x, \quad l = 1,...,k+1. \tag{1.18}$$

In this last formulation, $\mathcal{F}$ is a numerical flux which should be consistent with the physical flux function F.

Remark 4. The terms $\int_{\omega_i} F(v_h^{\omega_i})\partial_x\psi_l^{\omega_i}\mathrm{d}x$ and $\left[\psi_l^{\omega_i}\mathcal{F}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}}$ are respectively referred to as volume and surface integrals. The term $\int_{\omega_i} \psi_l^{\omega_i} B(v_h^{\omega_i},\partial_x b_h^{\omega_i})\mathrm{d}x$ is referred to as source term.

In the context of DG schemes, the numerical flux is defined as a function of the left and right traces of the local polynomial approximation coming from each side of the interface:

$$\mathcal{F}_{i+\frac{1}{2}} = \mathcal{F}\left(v_h^{\omega_i}\left(x_{i+\frac{1}{2}},t\right), v_h^{\omega_{i+1}}\left(x_{i+\frac{1}{2}},t\right)\right).$$

This function is generally obtained through the resolution of an exact or approximated Riemann solver. In the remainder of this work, we use the global Lax-Friedrichs (LF) numerical flux which reads:

$$\mathcal{F}(v_L;v_R) = \frac{1}{2}\left(F(v_R) - F(v_L) - \sigma(v_R - v_L)\right), \tag{1.19}$$

with,

$$\sigma = \max_{\omega\in\mathscr{T}_h}\left(|u| + \sqrt{gH}\right), \tag{1.20}$$

and the maximum is taken over the whole region. In what follows, we use the notations $v_{i-\frac{1}{2}}^{\pm}$, $v_{i+\frac{1}{2}}^{\pm}$ for the left and right traces of v_h in the boundaries $x_{i-\frac{1}{2}}$ and $x_{i+\frac{1}{2}}$ of ω_i. Hence the formulation (1.18) becomes:

$$\sum_{p=1}^{k+1}\partial_t\underline{v}_p^{\omega_i}\int_{\omega_i}\psi_p^{\omega_i}\psi_l^{\omega_i}\mathrm{d}x - \int_{\omega_i} F(v_h^{\omega_i})\partial_x\psi_l^{\omega_i}\mathrm{d}x + \left[\psi_l^{\omega_i}\mathcal{F}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \int_{\omega_i}\psi_l^{\omega_i}B(v_h^{\omega_i},\partial_x b_h^{\omega_i})\mathrm{d}x, \quad l = 1,...,k+1. \tag{1.21}$$

The global DG solution is obtained by the gathering all the local solutions.

Remark 5. The local semi-discrete system (1.21) can be expressed in matrix form:

$$\mathbf{M}_{\omega_i}\partial_t \underline{v}_{\omega_i} = \widetilde{\mathbf{L}}_{\omega_i} \tag{1.22}$$

where $\mathbf{M}_{\omega_i} = \left(\mathbf{m}_{l,p}^{\omega_i}\right)_{l,p=1,\dots,k+1}$ is the local mass matrix defined by:

$$\mathbf{m}_{l,p}^{\omega_i} = \int_{\omega_i} \psi_l^{\omega_i}\psi_p^{\omega_i}\mathrm{d}x,$$

and the local residual vector $\widetilde{\mathbf{L}}_{\omega_i} = \left(\widetilde{\mathbf{l}}_l^{\omega_i}\right)_{l=1,\dots,k+1}$ gathers the volume integrals, surface integrals and source terms as follows:

$$\widetilde{\mathbf{l}}_l^{\omega_i} = \int_{\omega_i} F(v_h^{\omega_i})\partial_x\psi_l^{\omega_i}\mathrm{d}x - \left[\psi_l^{\omega_i}\mathcal{F}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} + \int_{\omega_i}\psi_l^{\omega_i}B(v_h^{\omega_i},\partial_x b_h^{\omega_i})\mathrm{d}x$$

Remark 6. The global semi-discrete system may be written as follows in matrix form:

$$\mathbf{M}\partial_t\underline{v} = \widetilde{\mathbf{L}} \tag{1.23}$$

where $\mathbf{M}$ is the block-diagonal matrix and $\widetilde{\mathbf{L}}$ the global residual vector are defined by:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{\omega_1} & 0 & \cdots & 0 \\ 0 & \mathbf{M}_{\omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{M}_{\omega_{n_{\mathrm{el}}}} \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{L}} = \left[\widetilde{\mathbf{L}}_{\omega_1}, \widetilde{\mathbf{L}}_{\omega_2}, \cdots, \widetilde{\mathbf{L}}_{\omega_{n_{\mathrm{el}}}}\right]^T$$

and $\underline{v}$ the global solution on the whole domain:

$$\underline{v} = \left[\underline{v}_{\omega_1}, \underline{v}_{\omega_2}, \cdots, \underline{v}_{\omega_{n_{\mathrm{el}}}}\right]^T.$$

Outline of this tapescript

We are now ready to present the work performed during this Ph.D. In this respect, the presentation is divided into two parts.

The first part of this work is devoted to the construction and analysis of new FV-Subcell correction methods for the NSW equations with varying topography. This part is divided into three chapters (Chapter 2 to Chapter 4).

In Chapter 2, we introduce a FV-Subcell correction strategy within an *a priori* approach. In particular, we introduce a novel numerical method that efficiently combines FV-Subcell correction within an *a priori* treatment, together with the high-order positivity-preserving limiter of [179] and a classical TVB slope-limiter [45]. This work has not been published yet, as we believe that the *a posteriori* strategy (introduced in Chapter 3) results in a globally more satisfying parameter-free algorithm. Though, *a priori* strategies may be of interest for other applications and this is the reason why we choose to provide some details in this chapter.

In Chapter 3, we detail how one can reformulate a straightforward DG formulation as a FV method operating on a sub-partition, with particular subcell's interfaces fluxes, which are called *reconstructed fluxes*. Then, we show how it is possible to surgically replace these interface fluxes by the lowest-order FV fluxes only in troubled subcells. This approach is then embedded into a new *a posteriori* correction strategy, which is extensively validated to highlight its robustness. It is also shown that, provided some interface states reconstructions are performed, it is possible to obtain a globally well-balanced discrete formulation. This work is already published in Journal of Computational Physics, see [80]. We also take advantage of this *a posteriori* strategy is the second part of this work to stabilize the computations associated with the floating object.

In Chapter 4, we introduce some preliminary works concerning the 2d extension of the *a posteriori* LSC method method for the NSW equations on unstructured simplicial meshes. Several new difficulties associated with the construction of the sub-partitions are identified, and some numerical validations involving wet/dry interfaces and well-balancing are shown. This part is still an ongoing project, and will be the topic of a near future article.

The second part is dedicated to the study of nonlinear interactions between free-surface shallow-water flows and a partially immersed floating object. Specifically, we design and analyze a new robust high-order DG-ALE discrete formulation for such adaptive simulations. This second part is splitted into two chapters (Chapter 5 and Chapter 6). The content of this second part has been submitted for publication in Journal of Computational Physics under the form of two full-length research papers.

Chapter 5 is devoted to the introduction of the continuous models associated with the presence of a partly immersed floating object in shallow-water. Some local existence results coming from [87] are also recalled.

In Chapter 6, we finally design and analyze the corresponding discrete formulation, which is then extensively validated through several benchmarks. This formulation depends on the object's motion, which may be either prescribed, or computed as a response to the hydrodynamic forcing associated

with the flow (with heave, surge and pitch motions allowed in the horizontal one-dimensional case). These assets are numerically illustrated through an extensive set of manufactured benchmarks validating the water-body interaction model.

We provide a brief conclusion that summarizes the new materials obtained during these three years and described here. This conclusion is supplemented with several insights of the upcoming works.

Finally, several Appendices are provided in order to specify several technical issues without making the tapescript more cumbersome.

Part I

Stabilization of DG through FV-Subcell correction

Chapter 2

An *a priori* hybrid DG / FV-Subcell method for the NSW equations

In this chapter, we develop an *a priori* positivity-preserving high-order accurate, well-balanced DG method for the NSW equations with topography source term, using the high-order Positivity-Limiter (PL) of [183, 185]. Most existing numerical solutions to deal with wet/dry fronts in NSW equations are developped within an *a priori* reconstructions or limitations. While such approaches may potentially ensure the water-height positivity at a given time-step, nothing generally ensures that it remains non-negative after the upcoming time-step. Borrowing some ideas from [183, 185], we use a positivity-preserving limiter which preserves the accuracy and ensures the local mass conservation. We also introduce some modifications on the numerical fluxes in order to further ensure the well-balanced property. Assuming that the water-height is initially positive, we show that the water-height remains positive at the next time-step under a suitable CFL condition.

However, in practice, additional trouble may generally be experienced as the water-height is close to zero. In such almost-dry areas, the fluid horizontal velocity $u = q/H$ is not computed accurately and very large values can be produced, even with a small numerical error on the discharge, leading to over-restricted time-steps and an inaccurate description of the flow. To alleviate this limitation, by taking inspiration from the NSW DG/FV scheme of A. Meister and S. Ortleb, see [124], we introduce in this chapter an hybrid approach, in which we locally replace the high-order DG scheme by some lowest-order FV scheme acting on a dedicated sub-grid, only in the vicinity of dry areas. Hence, the high-order DG scheme with positivity-preserving limiter approach is used only in wet areas. The precise definition of the vicinity of such dry areas is specified in an *ad hoc* fashion with an arbitrary threshold value ε_d. Incidentally, this method also leads to some sensible decreasing of the computational cost, thanks to the use of a less expensive lowest-order FV scheme in all the dry and almost dry regions. Also, some local modifications of the numerical fluxes are performed for the lowest-order FV-Subcell scheme to enforce the well-balanced property. A proof of the well-balanced property for each ingredient (pure DG, pure FV-Subcell and hybrid DG / FV-Subcell) is provided in § 2.5.1 and § 2.5.2. This positivity-preserving hybrid approach is denoted by $PL\ DG\text{-}FV_{subcell}$ in the following. Several numerical computations that highlight the resulting properties of this new *a priori* hybrid strategy, are presented in this chapter.

We also show that a TVB slope-limiter may be efficiently embedded within the high-order DG schemes, in combination with the PL [183, 185]. The characteristic-wise total variation bounded

(TVB) limiter is used, see [48, 148]. This combination was introduced by [179], and is denoted by PL/TVB in the following. Several numerical computations for the resulting properties of the PL/TVB method, are presented and compared with the *a posteriori* LSC method in chapter 3.

2.1 Well-balanced numerical fluxes

Remark 7. The notations used in this chapter are defined either whenever needed, or previously in Chapter 1, §1.3.

Several well-balanced DG methods for the NSW equations have been developed, see for example [146] for a list of references. In this work, we consider the approach of [177], (see also [179]), where it is shown that the straightforward DG method is able to exactly preserve the motionless steady state, provided some modifications of the interface fluxes. This is one of the simplest approaches to obtain a high-order well-balanced DG scheme, and the computational cost is basically the same as the straightforward DG method. In this section, we briefly recall this well-balanced approach. We first define the reconstructed interface values $v^{*,\pm}_{i+\frac{1}{2}}$ as follows:

$$v^{*,\pm}_{i+\frac{1}{2}} = \begin{pmatrix} H^{*,\pm}_{i+\frac{1}{2}} \\ H^{*,\pm}_{i+\frac{1}{2}} u^{\pm}_{i+\frac{1}{2}} \end{pmatrix}, \tag{2.1}$$

with the cell's interface reconstruction for the water-height:

$$H^{*,\pm}_{i+\frac{1}{2}} = \max\left(0, H^{\pm}_{i+\frac{1}{2}} + b^{\pm}_{i+\frac{1}{2}} - \max\left(b^{-}_{i+\frac{1}{2}}, b^{+}_{i+\frac{1}{2}}\right)\right), \tag{2.2}$$

the numerical flux $\mathcal{F}$ are replaced by the redefined numerical flux $\mathcal{F}^*$ as follows:

$$\mathcal{F}^{*,l}_{i+\frac{1}{2}} = \mathcal{F}\left(v^{*,-}_{i+\frac{1}{2}}; v^{*,+}_{i+\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H^{-}_{i+\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(H^{*,-}_{i+\frac{1}{2}})^2 \end{pmatrix}, \tag{2.3}$$

$$\mathcal{F}^{*,r}_{i-\frac{1}{2}} = \mathcal{F}\left(v^{*,-}_{i-\frac{1}{2}}; v^{*,+}_{i-\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H^{+}_{i-\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(H^{*,+}_{i-\frac{1}{2}})^2 \end{pmatrix}. \tag{2.4}$$

Finally, the DG scheme (1.18) is replaced by the well-balanced DG scheme (2.5):

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\omega_i} v_h^{\omega_i}\psi_l^{\omega_i}\mathrm{d}x - \int_{\omega_i} F(v_h^{\omega_i})\partial_x\psi_l^{\omega_i}\mathrm{d}x + \left[\psi_l^{\omega_i}\mathcal{F}^*\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \int_{\omega_i}\psi_l^{\omega_i} B(v_h^{\omega_i}, \partial_x b_h^{\omega_i})\mathrm{d}x, \quad \forall\, l = 0, ..., k. \tag{2.5}$$

2.2 TVB slope-limiter

Another stabilization process for the DG methods may generally be needed to deal with the occurence of discontinuities in the underlying solution. To this end, one of the simplest approach is to use a

slope-limiter, as in the FV methods, which should be applied after each inner stage of the RK time-stepping. Such a slope-limiter, like the TVB limiter of [44], may be used on top of the positivity-preserving limiter § 2.4. This double limitation process (positivity-preserving limiter and TVB slope-limiter) is used for instance in [179] and referred to as PL/TVB method in what follows.
Usually, for the shallow-water system, we perform the TVB limiting in the local characteristic variables (Riemann invariants). However, this limiter procedure might destroy the preservation of the still water steady state $H+b=cst$. Therefore, following the idea presented in [9, 186], we apply the limiter procedure on the function $(H+b,q)^T$ instead. The modified RK-DG solution is then defined by $\widetilde{H}=\widetilde{H+b}-b$. We denote by $\eta=H+b$ the total water elevation and the modified solution writes $\widetilde{H}=\widetilde{\eta}-b$.

Remark 8. We observe that this procedure does not destroy the conservation of H, which should be maintained during the limiter process. In fact $\overline{\widetilde{H}}=\overline{\widetilde{\eta}-b}=\overline{\widetilde{\eta}}-\overline{b}\overbrace{=}^{(2.7)}\overline{\eta}-\overline{b}=\overline{\eta-b}=\overline{H}$.

We start by introducing the minmod function defined by:

$$\text{minmod}\left(a_1,\cdots,a_m\right)=\begin{cases} s\min_i|a_i|, & \text{if } s=\text{sign}\left(a_1\right)=\cdots=\text{sign}\left(a_m\right),\\ 0, & \text{otherwise.}\end{cases}$$

If the following inequalities

$$|\eta^-_{i+\frac{1}{2}}-\overline{\eta^{\omega_i}_h}|\leq Mh^2_{\omega_i}\quad\text{and}\quad|\eta^+_{i-\frac{1}{2}}-\overline{\eta^{\omega_i}_h}|\leq Mh^2_{\omega_i}\tag{2.6}$$

are satisfied, then the solution in cell ω_i is assumed to be smooth, and thus the cell is not considered problematic. Otherwise, we compute the following quantities:

$$\Delta\widetilde{\eta}^-_{\omega_i}=\text{minmod}\left(\overline{\eta}_{\omega_i}-\eta^+_{i-\frac{1}{2}},\overline{\eta}_{\omega_{i+1}}-\overline{\eta}_{\omega_i},\overline{\eta}_{\omega_i}-\overline{\eta}_{\omega_{i-1}}\right),$$

$$\Delta\widetilde{\eta}^+_{\omega_i}=\text{minmod}\left(\eta^-_{i+\frac{1}{2}}-\overline{\eta}_{\omega_i},\overline{\eta}_{\omega_{i+1}}-\overline{\eta}_{\omega_i},\overline{\eta}_{\omega_i}-\overline{\eta}_{\omega_{i-1}}\right).$$

Then we set

$$\Delta\widetilde{\eta}_{\omega_i}=\min\left(\Delta\widetilde{\eta}^-_{\omega_i},\Delta\widetilde{\eta}^+_{\omega_i}\right).$$

Finally, the modified $\widetilde{\eta}$ is defined by

$$\widetilde{\eta}=\overline{\eta}+\frac{2}{h_{\omega_i}}(x-x_i)\Delta\widetilde{\eta}_{\omega_i}.\tag{2.7}$$

We remind the reader that x_i is the center of the cell ω_i and $\overline{\eta}$ is the mean-value of η on ω_i:

$$\overline{\eta}=\frac{1}{h_{\omega_i}}\int_{\omega_i}\eta dx.$$

We refer the reader to [179] for more information about the choice of the parameter M in (2.6). Finally, the modified water-height H is simply defined by $\widetilde{H}=\widetilde{\eta}-b$. We do exactly the same procedure for the modified water discharge $\widetilde{q}$. Whenever the condition (2.6) is not satisfied the solution $(H,q)^T$ is replaced by $(\widetilde{H},\widetilde{q})^T$.
The discrete initial data $\mathbf{v}^0_h$ is defined as in Remark 24

2.3 Time marching algorithm

Supplementing (2.5) with an initial datum $v(0,\cdot) = v_0 = (H_0, q_0)^{\mathrm{t}}$, the time-stepping may be carried out using explicit SSP-RK schemes, [74, 149]. For instance, writing the semi-discrete equation (2.5) in the operator form

$$\partial_t v_h + \mathcal{A}_h(v_h) = 0,$$

we advance from time level n to $(n+1)$ with the third-order scheme as follows:

$$\begin{aligned}
v_h^{n,1} &= v_h^n - \Delta t^n \mathcal{A}_h(v_h^n), \\
v_h^{n,2} &= \frac{1}{4}(3v_h^n + v_h^{n,1}) - \frac{1}{4}\Delta t^n \mathcal{A}_h(v_h^{n,1}), \\
v_h^{n+1} &= \frac{1}{3}(v_h^n + 2v_h^{n,2}) - \frac{2}{3}\Delta t^n \mathcal{A}_h(v_h^{n,2}),
\end{aligned}$$

where $v_h^{n,i}$, $1 \leq i \leq 2$, are the solutions obtained at intermediate stages. As the correction described in the following section make use of both DG scheme on the primal cells $\omega \in \mathscr{T}_h$ and FV scheme on the subcells $S_m^\omega \in \mathscr{T}_\omega$, the time-step Δt^n is computed adaptively using the following CFL condition:

$$\Delta t^n = \frac{\min\limits_{\omega \in \mathscr{T}_h} \left(\dfrac{h_\omega}{2k+1}, \min\limits_{S_m^\omega \in \mathscr{T}_\omega} |S_m^\omega| \right)}{\sigma}, \tag{2.8}$$

where σ is the constant defined in (2.33).

2.4 Positivity-preserving limiter

Recalling that high-order RK SSP time marching algorithms [74, 149] may be regarded as convex combinations of first-order forward Euler schemes, for sake of simplicity we will consider the Euler first-order scheme. By taking the test function $\psi_l^{\omega_i} = 1$ in (2.5), we obtain the governing equation satisfied by the cell averages. Dropping the "n" notation for the numerical fluxes and source term for conciseness, we get:

$$\overline{v}_{\omega_i}^{n+1} = \overline{v}_{\omega_i}^n - \frac{\Delta t}{h_{\omega_i}} \left(\mathcal{F}_{i+\frac{1}{2}}^{*,l} - \mathcal{F}_{i-\frac{1}{2}}^{*,r} \right) + \frac{\Delta t}{h_{\omega_i}} \int_{\omega_i} B(v_h^{\omega_i}, \partial_x b_h^{\omega_i}) dx. \tag{2.9}$$

We remind the reader that $h_{\omega_i} = |\omega_i|$ is the size of the cell ω_i and $\overline{v}_{\omega_i}^n$ the average value of $v_h^{\omega_i}$ on ω_i:

$$\overline{v}_{\omega_i}^n = \frac{1}{h_{\omega_i}} \int_{\omega_i} v_h^{\omega_i, n} dx.$$

By plugging (2.3)-(2.4) and (2.1) into (2.9), the governing equations of the cell averages of the water-height in the well-balanced DG scheme (2.5) can be written as:

$$\overline{H}_{\omega_i}^{n+1} = \overline{H}_{\omega_i}^n - \frac{\Delta t}{h_{\omega_i}} \left[\mathcal{F}_1 \left(H_{i+\frac{1}{2}}^{*,-}, u_{i+\frac{1}{2}}^-; H_{i+\frac{1}{2}}^{*,+}, u_{i+\frac{1}{2}}^+ \right) - \mathcal{F}_1 \left(H_{i-\frac{1}{2}}^{*,-}, u_{i-\frac{1}{2}}^-; H_{i-\frac{1}{2}}^{*,+}, u_{i-\frac{1}{2}}^+ \right) \right], \tag{2.10}$$

where the $H_{i\pm 1/2}^{*,\pm}$ are defined in (2.2) and $\mathcal{F}_1$ is the first component of the numerical flux:

$$\mathcal{F}_1\left(H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}; H^{*,+}_{i+\frac{1}{2}}, u^+_{i+\frac{1}{2}}\right) = \frac{1}{2}\left(H^{*,-}_{i+\frac{1}{2}} u^-_{i+\frac{1}{2}} + H^{*,+}_{i+\frac{1}{2}} u^+_{i+\frac{1}{2}} - \sigma\left(H^{*,+}_{i+\frac{1}{2}} - H^{*,-}_{i+\frac{1}{2}}\right)\right). \tag{2.11}$$

Before going further, the essential ingredient to ensure a high-order positivity-preserving scheme is to guarantee that the first-order version of it does indeed produce of positive solution. To this end, let us define the cell reconstructed first-order interface values for the topography:

$$\overline{b}_{i+\frac{1}{2}} := \max(\overline{b}_{\omega_i}, \overline{b}_{\omega_{i+1}}) \quad \text{and} \quad \overline{b}_{i-\frac{1}{2}} := \max(\overline{b}_{\omega_{i-1}}, \overline{b}_{\omega_i}), \tag{2.12}$$

where $\overline{b}_{\omega_i}$ and $\overline{b}_{\omega_{i\pm1}}$ are the average values of $b_h^{\omega_i}$ on ω_i and $\omega_{i\pm1}$ respectively.

Lemma 9. Under the CFL condition $\frac{\Delta t}{h_{\omega_i}}\alpha \leq 1$, where $\alpha = \max(|u| + \sqrt{gH})$, we consider the following first order scheme (we drop the "n" notation)

$$\overline{H}^{n+1}_{\omega_i} = \overline{H}_{\omega_i} - \frac{\Delta t}{h_{\omega_i}}\left[\mathcal{F}_1\left(\overline{H}^+_{\omega_i}, \overline{u}_{\omega_i}; \overline{H}^-_{\omega_{i+1}}, \overline{u}_{\omega_{i+1}}\right) - \mathcal{F}_1\left(\overline{H}^+_{\omega_{i-1}}, \overline{u}_{\omega_{i-1}}; \overline{H}^-_{\omega_i}, \overline{u}_{\omega_i}\right)\right], \tag{2.13}$$

where $\mathcal{F}_1$ is defined as in (2.11). We also consider

$$\overline{H}^{\pm}_{\omega_i} = \max\left(0, \overline{H}_{\omega_i} + \overline{b}_{\omega_i} - \overline{b}_{i\pm\frac{1}{2}}\right). \tag{2.14}$$

If $\overline{H}_{\omega_i}$ and $\overline{H}_{\omega_{i\pm1}}$ are non-negative, then $\overline{H}^{n+1}_{\omega_i}$ is non-negative.

Proof. We suppose first that $\overline{H}_{\omega_i}$ and $\overline{H}_{\omega_{i\pm1}}$ are strictly positive. Denoting by $\lambda = \frac{\Delta t}{h_{\omega_i}}$, the scheme (2.13) is equivalent to:

$$\begin{aligned}\overline{H}^{n+1}_{\omega_i} =& \overline{H}^n_{\omega_i} - \lambda\left[\frac{1}{2}\left(\overline{H}^+_{\omega_i}\overline{u}_{\omega_i} + \overline{H}^-_{\omega_{i+1}}\overline{u}_{\omega_{i+1}} - \sigma\left(\overline{H}^-_{\omega_{i+1}} - \overline{H}^+_{\omega_i}\right)\right)\right] \\ &+ \lambda\left[\frac{1}{2}\left(\overline{H}^+_{\omega_{i-1}}\overline{u}_{\omega_{i-1}} + \overline{H}^-_{\omega_i}\overline{u}_{\omega_i} - \sigma\left(\overline{H}^-_{\omega_i} - \overline{H}^+_{\omega_{i-1}}\right)\right)\right].\end{aligned}$$

Finally, the scheme (2.13) can be written as:

$$\begin{aligned}\overline{H}^{n+1}_{\omega_i} =& \left[1 - \frac{1}{2}\lambda\left(\sigma + \overline{u}_{\omega_i}\right)\frac{\overline{H}^-_{\omega_i}}{\overline{H}_{\omega_i}} - \frac{1}{2}\lambda\left(\sigma - \overline{u}_{\omega_i}\right)\frac{\overline{H}^+_{\omega_i}}{\overline{H}_{\omega_i}}\right]\overline{H}_{\omega_i} \\ &+ \left[\frac{1}{2}\lambda\left(\sigma + \overline{u}_{\omega_{i-1}}\right)\frac{\overline{H}^+_{\omega_{i-1}}}{\overline{H}_{\omega_{i-1}}}\right]\overline{H}_{\omega_{i-1}} + \left[\frac{1}{2}\lambda\left(\sigma - \overline{u}_{\omega_{i+1}}\right)\frac{\overline{H}^-_{\omega_{i+1}}}{\overline{H}_{\omega_{i+1}}}\right]\overline{H}_{\omega_{i+1}}.\end{aligned}$$

Therefore, $\overline{H}^{n+1}_{\omega_i}$ is a convex combination of $\overline{H}_{\omega_{i-1}}$, $\overline{H}_{\omega_i}$ and $\overline{H}_{\omega_{i+1}}$. Moreover the coefficients are non-negative since $\lambda\alpha \leq 1$ and $0 \leq \overline{H}^{\pm}_{\omega_i} \leq \overline{H}_{\omega_i}$ for all ω_i. Thus, we get $\overline{H}^{n+1}_{\omega_i} \geq 0$.

Let us suppose now that $\overline{H}_{\omega_i} = 0$. Since $\overline{b}_{\omega_i} - \overline{b}_{i\pm\frac{1}{2}} \leq 0$, we have $\overline{H}^+_{\omega_i} = \overline{H}^-_{\omega_i} = 0$. Thus, the equation (2.13) writes

$$\overline{H}^{n+1}_{\omega_i} = \left[\frac{1}{2}\lambda\left(\sigma + \overline{u}_{\omega_{i-1}}\right)\frac{\overline{H}^+_{\omega_{i-1}}}{\overline{H}_{\omega_{i-1}}}\right]\overline{H}_{\omega_{i-1}} + \left[\frac{1}{2}\lambda\left(\sigma - \overline{u}_{\omega_{i+1}}\right)\frac{\overline{H}^-_{\omega_{i+1}}}{\overline{H}_{\omega_{i+1}}}\right]\overline{H}_{\omega_{i+1}},$$

which is still a positive term. We proceed in the same way for the remaining cases ($\overline{H}_{\omega_{i\pm1}} = 0$). □

Let us introduce the N-points Legendre Gauss-Lobatto quadrature rule on the interval $\omega_i = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$. This rule is exact for the integral of polynomials of degree up to $2N-3$, where N is such that $2N-3 \geq k$. We denote these quadrature points on ω_i as

$$S_i = \left\{x_{i-1/2} = \widehat{x}^1_{\omega_i}, \widehat{x}^2_{\omega_i}, \cdots, \widehat{x}^{N-1}_{\omega_i}, \widehat{x}^N_{\omega_i} = x_{i+1/2}\right\}.$$

Let $\{\widehat{w}_t\}_{t=1,\ldots,N}$ be the quadrature weights for the interval such that $\sum_{t=1}^N \widehat{w}_t = 1$. We recall that $H_h^{\omega_i}(x)$ denotes the DG polynomial approximating the water-height in the cell ω_i. We have:

$$\overline{H}_{\omega_i} = \frac{1}{h_{\omega_i}} \int_{\omega_i} H_h^{\omega_i}(x)dx = \sum_{t=1}^{N} \widehat{w}_t H_h^{\omega_i}\left(\widehat{x}^t_{\omega_i}\right) = \sum_{t=2}^{N-1} \widehat{w}_t H_h^{\omega_i}\left(\widehat{x}^t_{\omega_i}\right) + \widehat{w}_1 H^+_{i-\frac{1}{2}} + \widehat{w}_N H^-_{i+\frac{1}{2}}, \tag{2.15}$$

where $H^+_{i-\frac{1}{2}} = H_h^{\omega_i}(\widehat{x}^1_{\omega_i})$ and $H^-_{i+\frac{1}{2}} = H_h^{\omega_i}(\widehat{x}^N_{\omega_i})$.

Proposition 1. *Let $H_h^{\omega_i}(x)$ be the DG polynomial for the water-height in the cell ω_i obtained through (2.5). If $H^{-,n}_{i-\frac{1}{2}}$, $H^{+,n}_{i+\frac{1}{2}}$ and $H_h^{\omega_i,n}\left(\widehat{x}^t_{\omega_i}\right)$ are non-negative $\forall\ t = 1, \cdots, N$, then the water-height mean value at time level $n+1$, $\overline{H}^{n+1}_{\omega_i}$, is non-negative under the CFL condition*

$$\lambda\sigma \leq \widehat{w}_1. \tag{2.16}$$

Proof. We drop the "n" notation for the sake of conciseness. By substituting (2.15) into (2.10), we can rewrite (2.10) by adding and subtracting the term $\mathcal{F}_1\left(H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}; H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}\right)$:

$$\begin{aligned}
\overline{H}^{n+1}_{\omega_i} &= \sum_{t=2}^{N-1} \widehat{w}_t H_h^{\omega_i}\left(\widehat{x}^t_{\omega_i}\right) + \widehat{w}_1 H^+_{i-\frac{1}{2}} + \widehat{w}_N H^-_{i+\frac{1}{2}} \\
&\quad - \lambda\left[\mathcal{F}_1\left(H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}; H^{*,+}_{i+\frac{1}{2}}, u^+_{i+\frac{1}{2}}\right) - \mathcal{F}_1\left(H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}; H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}\right)\right. \\
&\quad \left. + \mathcal{F}_1\left(H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}; H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}\right) - \mathcal{F}_1\left(H^{*,-}_{i-\frac{1}{2}}, u^-_{i-\frac{1}{2}}; H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}\right)\right] \\
&= \sum_{t=2}^{N-1} \widehat{w}_t H_h^{\omega_i}\left(\widehat{x}^t_{\omega_i}\right) + \widehat{w}_N H_N + \widehat{w}_1 H_1,
\end{aligned} \tag{2.17}$$

where

$$H_1 = H^+_{i-\frac{1}{2}} - \frac{\lambda}{\widehat{w}_1}\left[\mathcal{F}_1\left(H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}; H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}\right) - \mathcal{F}_1\left(H^{*,-}_{i-\frac{1}{2}}, u^-_{i-\frac{1}{2}}; H^{*+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}\right)\right] \tag{2.18}$$

and

$$H_N = H^-_{i+\frac{1}{2}} - \frac{\lambda}{\widehat{w}_N}\left[\mathcal{F}_1\left(H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}; H^{*,+}_{i+\frac{1}{2}}, u^+_{i+\frac{1}{2}}\right) - \mathcal{F}_1\left(H^{*,+}_{i-\frac{1}{2}}, u^+_{i-\frac{1}{2}}; H^{*,-}_{i+\frac{1}{2}}, u^-_{i+\frac{1}{2}}\right)\right] \quad . \tag{2.19}$$

We notice that (2.18) and (2.19) are both of the type (2.13). Hence $H_1 \geq 0$ and $H_N \geq 0$ under the CFL condition (2.16) $\lambda\sigma \leq \widehat{w}_1 = \widehat{w}_N$. Therefore, $\overline{H}_{\omega_i}^{n+1} \geq 0$ since it is a convex combination of H_1, H_N and $H_h^{\omega_i}\left(\widehat{x}_{\omega_i}^t\right)_{t=2,\cdots,N-1}$.

□

To enforce the conditions of this proposition, we need to modify $H_h^{\omega_i,n}(x)$ for it to be non-negative for all $x \in S_i$. At time level n, given that $\overline{H}_{\omega_i}^n \geq 0$, we can introduce the following limiter on the DG polynomial $v_h^{\omega_i,n}(x) = \left(H_h^{\omega_i,n}(x), (Hu)_h^{\omega_i,n}(x)\right)^T$ as introduced in [179]. This limiter is a linear scaling around its cell average:

$$\widetilde{v}_h^{\omega_i,n}(x) = \theta\left(v_h^{\omega_i,n}(x) - \overline{v}_{\omega_i}^n\right) + \overline{v}_{\omega_i}^n, \quad \theta = \min\left\{1, \frac{\overline{H}_{\omega_i}^n}{\overline{H}_{\omega_i}^n - m_{\omega_i}}\right\} \tag{2.20}$$

$$m_{\omega_i} = \min_{x\in S_i} H_h^{\omega_i,n}(x) = \min_{t=1,\dots N} H_h^{\omega_i,n}\left(\widehat{x}_{\omega_i}^t\right).$$

It is easy to observe that $\widetilde{H}_h^{\omega_i,n}\left(\widehat{x}_{\omega_i}^t\right) \geq 0\ _{t=1,\cdots,N}$ for any cell ω_i. We compute the modified polynomial $\widetilde{v}_h^{\omega_i,n}(x)$ and use it instead of $v_h^{\omega_i,n}(x)$. Hence, using the positivity-preserving limiter (2.20), we can ensure the positivity of the water-height mean-value $\overline{H}_{\omega_i}^{n+1}$ for any cell ω_i at the next time level $n+1$. Therefore (2.20) is indeed a positivity-preserving limiter for the well-balanced DG scheme (2.5).

Until now we can ensure the positivity of the height mean-values $\overline{H}_{\omega_i}^{n+1}{}_{i=1,\dots,n_{\text{el}}}$ at the next time level $n+1$ via the well-balanced DG scheme (2.5) with the previously defined positivity-preserving limiter (2.20). We already know that a first-order well-balanced FV-Subcell scheme § 2.5.1 is applied on specific cells (where $\overline{H}_{\omega_i}^n < \varepsilon_d$), see Remark 10-11. Considering a cell ω_i, to ensure the positivity of the water-height mean-value on ω_i at the next time level $n+1$, this is enough to guarantee the positivity of the water-height sub-mean-values $\overline{H}_1^{\omega_{i+1},n}$, $\overline{H}_{k+1}^{\omega_{i-1},n}$ and $\{\overline{H}_m^{\omega_i,n}\}_{m=1,\dots,k+1}$ on ω_i at the current time level n and then apply a FV-Subcell scheme on ω_i, since a first-order FV scheme conserves positivity.

The positivity limiter already introduced in (2.20) does not ensure the positivity of the water-height sub-mean values on subcells. Therefore we need to modifie the coefficient θ and replace it by θ':

$$\theta' = \min\left\{1, \frac{\overline{H}_{\omega_i}^n}{\overline{H}_{\omega_i}^n - m'_{\omega_i}}\right\}, \tag{2.21}$$

where

$$m'_{\omega_i} = \min(\min_{x\in S_i} H_h^{\omega_i,n}(x), \min_{m=1,\dots,k+1} \overline{H}_m^{\omega_i,n}). \tag{2.22}$$

The new limiter on the DG polynomial $v_h^{\omega_i,n}(x)$ is finally written as:

$$\widetilde{v}_h^{\omega_i,n}(x) = \theta'\left(v_h^{\omega_i,n}(x) - \overline{v}_{\omega_i}^n\right) + \overline{v}_{\omega_i}^n. \tag{2.23}$$

We note this new positivity limiter (2.23) by "*DG-FV$_{subcell}$ positivity limiter*". Thanks to (2.23), the water-height sub-mean-values $\widetilde{\overline{H}}^{\omega_i,n}_{m\ \ m=1,\dots,k+1}$ associated to the "limited" polynomial $\widetilde{v}_h^{\omega_i,n}$ (2.23) are non-negative for any cell ω_i. We drop the tilde notation in what follows.

Finally, using the well-balanced FV-Subcell scheme § 2.5.1, the *DG-FV$_{subcell}$ positivity limiter* ensures the positivity of the water-height mean-value $\overline{H}^{n+1}_{\omega_i}$ for any cell ω_i at the next time level $n+1$. Thus *DG-FV$_{subcell}$ positivity limiter* is a positivity-preserving limiter for both well-balanced DG method § 2.5.2 and well-balanced FV-Subcell method § 2.5.1.

Remark 10. A cell ω_i for which the water-height mean-value is smaller than the threshold ε_d (for a small enough ε_d) is considered as a dry cell. In this case, we use a first-order FV-Subcell scheme instead of the DG scheme. Also, there is no need to worry about negative water-height problem in dry regions, since we have just shown that the first-order FV-Subcell scheme ensures the positivity of H, thanks to the PL.

Remark 11. The *DG-FV$_{subcell}$ positivity limiter* (2.23) is applied in both cases, $\overline{H}^n_{\omega_i} > \varepsilon_d$ and $\overline{H}^n_{\omega_i} < \varepsilon_d$. In the first case ($\overline{H}^n_{\omega_i} > \varepsilon_d$), the well-balanced DG scheme (2.5) is applied. While in the second case ($\overline{H}^n_{\omega_i} < \varepsilon_d$), we use the well-balanced FV-Subcell scheme (see § 2.5.1). In both cases, the *DG-FV$_{subcell}$ positivity limiter* ensures the water-height positivity. This procedure is referred to as *PL DG-FV$_{subcell}$* method in what follows.

Remark 12. The TVB limiter should be used before the positivity-preserving limiter, so that the water-height positivity is not impacted. Indeed, by applying the positivity-preserving limiter on the linear solution (TVB-limited polynomial) on ω_i, we do not increase the slope associated with the piecewise-linear solution. For higher-order time-discretization, the positivity-preserving and TVB limiters are applied at each sub-step.

2.5 DG and FV-Subcell methods

Let us consider a cell ω_i. If the water-height mean-value on ω_i at time level n satisfies $\overline{H}^n_{\omega_i} < \varepsilon_d$, then a well-balanced FV-Subcell method is applied on ω_i. To this end, we subdivide ω_i into $k+1$ subcells $(S_m^{\omega_i})_{m=1,\dots,k+1}$. Then, using the transformation matrix $\mathbf{\Pi}_\omega$ (1.9), we compute the $k+1$ sub-mean-values from the high-order DG polynomial solution. The next step consists in applying a first-order FV scheme on each subcell of the cell under consideration (see Remark 13 for the different involved cases), and we obtain accordingly $k+1$ new sub-mean-values at time level $n+1$. Finally, using the inverse transformation matrix $\mathbf{\Pi}_\omega^{-1}$, we reconstruct a high-order $\mathbb{P}^k$ polynomial solution on ω_i.

2.5.1 FV-Subcell method: well-balancing and water-height positivity

The first-order FV scheme on a subcell $S_m^{\omega_i}$ is defined as follows:

$$\overline{v}_m^{\omega_i,n+1} = \overline{v}_m^{\omega_i,n} - \frac{\Delta t}{|S_m^{\omega_i}|}\left(\mathcal{F}^l_{m+\frac{1}{2}} - \mathcal{F}^r_{m-\frac{1}{2}}\right) + \frac{\Delta t}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}} B\left(\overline{v}_m^{\omega_i}, (\partial_x \overline{b})_m^{\omega_i}\right) dx, \tag{2.24}$$

where the well-balanced first-order FV numerical fluxes $\mathcal{F}^r_{m-\frac{1}{2}}$ and $\mathcal{F}^l_{m+\frac{1}{2}}$ will be defined, accordingly to the different scenarii, in the remainder of this section. For the source term discretization, we may use a simple FD approximation of term $(\partial_x \overline{b})_m^{\omega_i}$:

$$\frac{1}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}} B\left(\overline{v}_m^{\omega_i},(\partial_x\overline{b})_m^{\omega_i}\right)dx=\begin{pmatrix}0\\ \frac{1}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}}-\mathrm{g}\overline{H}_m^{\omega_i}(\partial_x\overline{b})_m^{\omega_i}dx\end{pmatrix}=\begin{pmatrix}0\\ -\mathrm{g}\overline{H}_m^{\omega_i}\dfrac{\left(\overline{b}_{m+\frac{1}{2}}-\overline{b}_{m-\frac{1}{2}}\right)}{|S_m^{\omega_i}|}\end{pmatrix},$$

where $\overline{b}_{m\pm\frac{1}{2}}$ would be the subcell version of the first-order topography interface reconstrcuted values defined in (2.12). However, let us emphasize that because we aim at designing an hybrid DG/FV scheme, we already have introduced an underlying polynomial representation of the water height and bathymetry, see (1.17). The source term

$$\frac{1}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}} B\left(v_h^{\omega_i},\partial_x b_h^{\omega_i}\right)dx=\begin{pmatrix}0\\ \frac{1}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}}-gH_h^{\omega_i}\partial_x b_h^{\omega_i}dx\end{pmatrix}$$

can then be computed exactly through an accurate enough quadrature rule. This option naturally enables us to impose a well-balanced property, see the remainder of the section.

Remark 13. For the computation of the numerical fluxes in (2.24), we have to distinguish two cases. We note *FV subcell*, the subcell in which we apply the well-balanced first-order FV scheme (2.24) and by *DG cell* the cell in which we apply the well-balanced DG scheme (2.5)-(2.3)-(2.4). Then, the two cases are the following: (i) the *FV subcell* $S_m^{\omega_i}$ is surrounded by *FV subcells*, (ii) the *FV subcell* $S_m^{\omega_i}$ is bounded from the left by a *DG cell* and from the right by a *FV subcell*. By symmetry, the situation where *FV subcell* $S_m^{\omega_i}$ is bounded from the left by a *FV subcell* and from the right by a *DG cell* falls into this latter.

Remark 14. For sake conciseness, we generally replace the notation of the sub-mean-value $\overline{v}_m^{\omega_i}$ on the cell $S_m^{\omega_i}$ by the notation $\overline{v}_m$.

Case 1: If the *FV subcell* $S_m^{\omega_i}$ is surrounded by *FV subcells*, see Fig. 2.1.

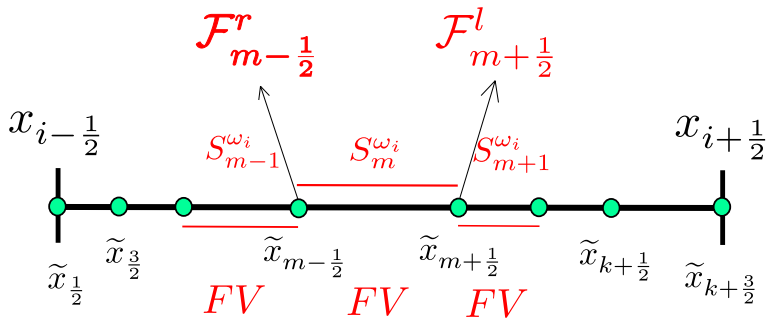

Figure 2.1: *FV subcell* $S_m^{\omega_i}$ is surroundedt by *FV subcells*

In this case, the numerical fluxes $\mathcal{F}^r_{m-\frac{1}{2}}$ and $\mathcal{F}^l_{m+\frac{1}{2}}$ are defined by:

$$\mathcal{F}^r_{m-\frac{1}{2}} = \mathcal{F}\left(\overline{H}^+_{m-1}, \overline{u}_{m-1}; \overline{H}^-_m, \overline{u}_m\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H_{m-\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(\overline{H}^-_m)^2 \end{pmatrix}, \tag{2.25}$$

$$\mathcal{F}^l_{m+\frac{1}{2}} = \mathcal{F}\left(\overline{H}^+_m, \overline{u}_m; \overline{H}^-_{m+1}, \overline{u}_{m+1}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H_{m+\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(\overline{H}^+_m)^2 \end{pmatrix}, \tag{2.26}$$

where $H_{m+\frac{1}{2}}$ and $H_{m-\frac{1}{2}}$ are respectively the interpolated values of the water-height polynomial $H_h^{\omega_i}$ on $S_m^{\omega_i}$ interfaces $\widetilde{x}_{m+\frac{1}{2}}$ and $\widetilde{x}_{m-\frac{1}{2}}$. Let us note that the $\overline{H}^{\pm}_m$ are nothing but the subcell version of the $\overline{H}^{\pm}_{\omega_i}$ defined in (2.14). By applying Lemma 9 on subcells, this first-order FV scheme does indeed preserve water-height positivity under the assumptions

$$\frac{\Delta t}{|S_m^{\omega_i}|}\,\sigma \leq 1 \quad \text{and} \quad \sigma = \max_m(|\overline{u}_m| + \sqrt{g\overline{H}_m}).$$

Remark 15. Let us note that the additional terms present in the numerical fluxes definition (2.25)-(2.26) do no impact the water height positivity.

Now, let us ensure that this scheme does enforce a well-balanced property.

Lemma 16. We consider the scheme (2.24) with the numerical fluxes defined in (2.25)-(2.26). This scheme preserves the motionless steady-states at the subcell level:

$$\forall \omega_i \in \mathscr{T}_h, \quad \forall m \in [1,\dots,k+1], \quad \overline{H}_m^{\omega_i,n} + \overline{b}_m^{\omega_i} = \eta_c, \quad \overline{u}_m^{\omega_i,n} = 0 \quad \Longrightarrow \quad \overline{H}_m^{\omega_i,n+1} + \overline{b}_m^{\omega_i} = \eta_c, \quad \overline{u}_m^{\omega_i,n+1} = 0.$$

Proof. Dropping the "ω_i" notation for submean-values, we have at time level n:

$$\overline{u}_m^n = \frac{\overline{q}_m^n}{\overline{H}_m^n} = 0 \quad \text{and} \quad \overline{H}_m^n + \overline{b}_m = \eta_c.$$

Let us show that the free-surface elevation is always equal to η_c and the water velocity is always zero in $S_m^{\omega_i}$ at the next time level $n+1$. Assuming that ω_i is a wetted cell, we have then:

$$\partial_x\left(H_h^{\omega_i,n} + b_h^{\omega_i,n}\right) = 0, \tag{2.27}$$

where $H_h^{\omega_i,n}$ and $b_h^{\omega_i,n}$ are the restriction of the polynomial solutions over the cell ω_i. It follows that:

$$\partial_x H_h^{\omega_i,n} = -\partial_x b_h^{\omega_i,n} \Rightarrow \mathrm{g}H_h^{\omega_i,n}\partial_x H_h^{\omega_i,n} = -\mathrm{g}H_h^{\omega_i,n}\partial_x b_h^{\omega_i,n}.$$

Then, by means of:

$$\partial_x\left(\frac{1}{2}\mathrm{g}(H_h^{\omega_i,n})^2\right) = -\mathrm{g}H_h^{\omega_i,n}\partial_x b_h^{\omega_i,n},$$

we are able to reformulate the source term as $F\left(v_h^{\omega_i,n}\right) = \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H_h^{\omega_i,n})^2 \end{pmatrix}$. We finally get:

$$\partial_x(F(v_h^{\omega_i,n})) - B(v_h^{\omega_i,n}, \partial_x b_h^{\omega_i,n}) = 0. \tag{2.28}$$

We drop the "n" notation in what follows. Let us first compute the global LF numerical flux with the re-defined variables (2.14) on $S_m^{\omega_i}$ left interface:

$$\mathcal{F}\left(\overline{H}_{m-1}^{+},\overline{u}_{m-1};\overline{H}_{m}^{-},\overline{u}_{m}\right)=\frac{1}{2}\left[\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(\overline{H}_{m}^{-})^2\end{pmatrix}+\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(\overline{H}_{m-1}^{+})^2\end{pmatrix}-\sigma\left[\begin{pmatrix}\overline{H}_{m}^{-}\\ 0\end{pmatrix}-\begin{pmatrix}\overline{H}_{m-1}^{+}\\ 0\end{pmatrix}\right]\right],$$

since $\overline{u}_m=\overline{u}_{m-1}=0$. Recalling the definitions of $\overline{H}_m^-$ and $\overline{H}_{m-1}^+$ from (2.14), one can easily note that $\overline{H}_m^-=\overline{H}_{m-1}^+$. Following the same development for the right interface, the numerical fluxes yield:

$$\mathcal{F}\left(\overline{H}_{m-1}^{+},\overline{u}_{m-1};\overline{H}_{m}^{-},\overline{u}_{m}\right)=\frac{1}{2}\left[\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(\overline{H}_{m}^{-})^2\end{pmatrix}+\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(\overline{H}_{m-1}^{+})^2\end{pmatrix}\right]=\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(\overline{H}_{m}^{-})^2\end{pmatrix}.$$

This enables to rewrite the subcell mean governing equation as:

$$\overline{v}_m^{n+1}=\overline{v}_m^n-\frac{\Delta t}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}}\partial_x(F(v_h^{\omega_i}))dx+\frac{\Delta t}{|S_m^{\omega_i}|}\int_{S_m^{\omega_i}}B(v_h^{\omega_i},\partial_x b_h^{\omega_i})dx.$$

It directly follows that $\overline{v}_m^{n+1}=\overline{v}_m^n$, which is equivalent to:

$$\overline{H}_m^{n+1}+\overline{b}_m=\overline{H}_m^n+\overline{b}_m=\eta_c\quad\text{ and }\quad\overline{u}_m^{n+1}=\frac{\overline{q}_m^{n+1}}{\overline{H}_m^{n+1}}=\frac{\overline{q}_m^n}{\overline{H}_m^n}=0.$$

We conclude that this scheme satisfies the well-balanced property.

□

Case 2: If the *FV subcell* $S_m^{\omega_i}$ is bounded from the left by a *DG cell* and from the right by a *FV subcell*, see Fig. 2.2.

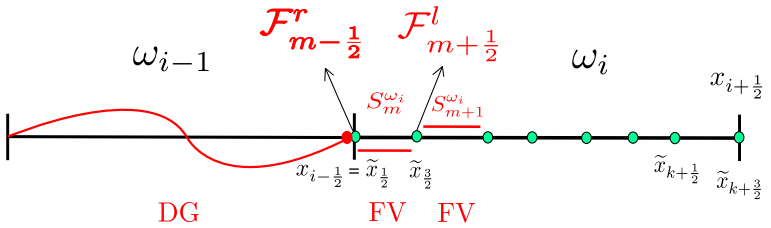

Figure 2.2: *FV subcell* $S_m^{\omega_i}$ is bounded from the left by a *DG cell* and from the right by a *FV subcell*

In this case, the numerical fluxes $\mathcal{F}^r_{m-\frac{1}{2}}$ and $\mathcal{F}^l_{m+\frac{1}{2}}$ are defined by:

$$\mathcal{F}^r_{m-\frac{1}{2}}=\mathcal{F}\left(H^{*,-}_{m-\frac{1}{2}},u^-_{m-\frac{1}{2}};\overline{H}_m^-,\overline{u}_m\right)+\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(H_{m-\frac{1}{2}})^2-\frac{1}{2}\mathrm{g}(\overline{H}_m^-)^2\end{pmatrix},\tag{2.29}$$

$$\mathcal{F}^l_{m+\frac{1}{2}}=\mathcal{F}\left(\overline{H}_m^+,\overline{u}_m;\overline{H}_{m+1}^-,\overline{u}_{m+1}\right)+\begin{pmatrix}0\\ \frac{1}{2}\mathrm{g}(H_{m+\frac{1}{2}})^2-\frac{1}{2}\mathrm{g}(\overline{H}_m^+)^2\end{pmatrix}.\tag{2.30}$$

As we see in Fig. 2.2, the left edge $\widetilde{x}_{m-\frac{1}{2}}$ coincides with $x_{i-\frac{1}{2}}$. We use on its left an interpolated value and on its right a sub-mean-value:

$$H^{*,-}_{m-\frac{1}{2}} = H^{*,-}_{i-\frac{1}{2}} = \max\left(0, H^-_{i-\frac{1}{2}} + b^-_{i-\frac{1}{2}} - \max\left(b^-_{i-\frac{1}{2}}, \overline{b}_m\right)\right), \tag{2.31}$$

$$\overline{H}^-_m = \max\left(0, \overline{H}_m + \overline{b}_m - \max\left(b^-_{i-\frac{1}{2}}, \overline{b}_m\right)\right). \tag{2.32}$$

As for the right interface, we keep the same reconstructed sub-mean values on both sides of the interface $\widetilde{x}_{m+\frac{1}{2}}$ as in the first case.

Lemma 17. The first-order FV scheme (2.24), provided with the numerical fluxes defined in (2.29)-(2.30), is indeed positivity-preserving under the CFL condition $\frac{\Delta t}{|S^{\omega_i}_m|}\sigma \leq 1$, with

$$\sigma = \max\left(\max_m(|\overline{u}_m| + \sqrt{g\overline{H}_m}) \ , \ \max_{\omega_i}(|u^{\pm}_{i+\frac{1}{2}}| + \sqrt{gH^{\pm}_{i+\frac{1}{2}}}\ )\right), \tag{2.33}$$

as if $\overline{H}_m$, $\overline{H}_{m+1}$ and $H^-_{m-\frac{1}{2}}$ are non-negative, then $\overline{H}^{n+1}_m$ is non-negative.

Proof. It is clear that $0 \leq H^{*,-}_{m-\frac{1}{2}} \leq H^-_{m-\frac{1}{2}}$, $0 \leq \overline{H}^{\pm}_m \leq \overline{H}_m$ and $0 \leq \overline{H}^-_{m+1} \leq \overline{H}_{m+1}$. Proceeding in the same way as in the proof of Lemma 9, we obtain that $\overline{H}^{n+1}_m$ is non-negative under the CFL condition $\frac{\Delta t}{|S^{\omega_i}_m|}\sigma \leq 1$. □

As said previously, the third case where the *FV subcell* $S^{\omega_i}_m$ is bounded from the left by a *FV subcell* and from the right by a *DG cell* falls in this latter case by symmetry. Let us now show the well-balanced property for this scheme.

Lemma 18. We consider the scheme (2.24) with the numerical fluxes defined in (2.29)-(2.30). This scheme preserves the motionless steady-states at the subcell level:

$$\forall \omega_i \in \mathscr{T}_h, \quad \forall m \in [1, \dots, k+1], \ \ \overline{H}^{\omega_i,n}_m + \overline{b}^{\omega_i}_m = \eta_c, \ \ \overline{u}^{\omega_i,n}_m = 0 \quad \Longrightarrow \quad \overline{H}^{\omega_i,n+1}_m + \overline{b}^{\omega_i}_m = \eta_c, \ \ \overline{u}^{\omega_i,n+1}_m = 0.$$

Proof. This is enough to notice that $H^{*,-}_{m-\frac{1}{2}} = \overline{H}^-_m$ under steady state hypothesis. We then proceed in a similar way as in the proof of lemma 16. □

2.5.2 DG method: well-balancing and water-height positivity

We now consider a cell ω_i where the water-height mean-value at time level n satisfies $\overline{H}^n_{\omega_i} > \varepsilon_d$. Then, the well-balanced DG scheme (2.5) is applied on this cell. Let us recall that the well-balanced DG scheme (2.5) on the cell ω_i, with the redefined well-balanced numerical fluxes (2.3)-(2.4):

$$\int_{\omega_i} \partial_t v^{\omega_i}_h \psi^{\omega_i}_l \mathrm{d}x = \int_{\omega_i} F\left(v^{\omega_i}_h\right) \partial_x \psi^{\omega_i}_l \mathrm{d}x - \left[\mathcal{F}^* \psi^{\omega_i}_l\right]^{i+\frac{1}{2}}_{i-\frac{1}{2}} + \int_{\omega_i} B\left(v^{\omega_i}_h, \partial_x b^{\omega_i}_h\right) \psi^{\omega_i}_l \mathrm{d}x, \quad \forall\ l = 0, ..., k. \tag{2.34}$$

Recalling once more that SSP-RK time discretization can be seen as a convex combination of first-order forward Euler time marching, let us focus on this latter for sake of simplicity. The total discrete version of DG scheme (1.22) then becomes:

$$\underline{v}_{\omega_i}^{n+1} = \underline{v}_{\omega_i}^{n} + \Delta t \mathbf{M}_{\omega_i}^{-1} \widetilde{\mathbf{L}}_{\omega_i}, \tag{2.35}$$

where $\underline{v}_{\omega_i} = \left(\underline{v}_p^{\omega_i}\right)_{p=0,\dots,k}$ stands for the local vector solution that gathers the degrees of freedom associated with $v_h^{\omega_i} = \sum_{p=0}^{k} \underline{v}_p^{\omega_i} \psi_p^{\omega_i}$, $\mathbf{M}_{\omega_i}$ the local mass matrix, and $\widetilde{\mathbf{L}}_{\omega_i} = \left(\widetilde{\mathbf{l}}_l^{\omega_i}\right)_{l=0,\dots,k}$ is local residual vector that gathers the volume integral, the surface integral and the source term. We recall that $\widetilde{\mathbf{L}}_{\omega_i}$ is defined by:

$$\widetilde{\mathbf{l}}_l^{\omega_i} = \int_{\omega_i} F(v_h^{\omega_i}) \partial_x \psi_l^{\omega_i} \mathrm{d}x - \left[\psi_l^{\omega_i} \mathcal{F}^*\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} + \int_{\omega_i} \psi_l^{\omega_i} B(v_h^{\omega_i}, \partial_x b_h^{\omega_i}) \mathrm{d}x, \qquad \forall\, l = 0, ..., k.$$

For the computation of the numerical fluxes, we have two cases to distinguish.

Case 1: If the *DG cell* ω_i is surrounded by *DG cells*, see Fig 2.3.

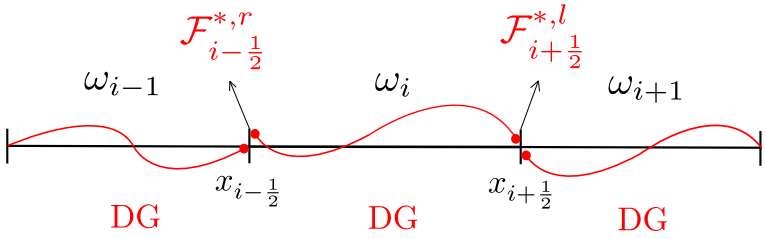

Figure 2.3: *DG cell* ω_i is surrounded by *DG cells*

In this case, the numerical fluxes $\mathcal{F}^{*,r}_{i-\frac{1}{2}}$ and $\mathcal{F}^{*,l}_{i+\frac{1}{2}}$ are defined by:

$$\mathcal{F}^{*,r}_{i-\frac{1}{2}} = \mathcal{F}\left(H^{*,-}_{i-\frac{1}{2}}, u^{-}_{i-\frac{1}{2}}; H^{*,+}_{i-\frac{1}{2}}, u^{+}_{i-\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H^{+}_{i-\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(H^{*,+}_{i-\frac{1}{2}})^2 \end{pmatrix}, \tag{2.36}$$

$$\mathcal{F}^{*,l}_{i+\frac{1}{2}} = \mathcal{F}\left(H^{*,-}_{i+\frac{1}{2}}, u^{-}_{i+\frac{1}{2}}; H^{*,+}_{i+\frac{1}{2}}, u^{+}_{i+\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}(H^{-}_{i+\frac{1}{2}})^2 - \frac{1}{2}\mathrm{g}(H^{*,-}_{i+\frac{1}{2}})^2 \end{pmatrix}. \tag{2.37}$$

As we saw in Proposition 1, the DG scheme (2.34) with the numerical fluxes (2.36)-(2.37) preserves water-height positivity. Furthermore, as mentioned in [179], this DG scheme is capable of maintaining the still solution exactly, *i.e.* it satisfies the well-balanced property.

case 2: If the *DG cell* ω_i is bounded from the left by a *DG cell* and from the right by *FV subcell*, see Fig 2.4.

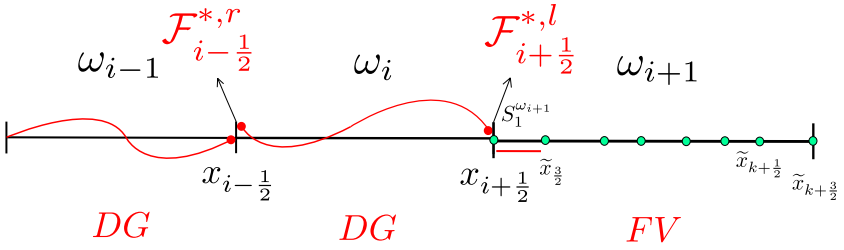

Figure 2.4: *DG cell* ω_i is bounded from the left by a *DG cell* and from the right by *FV subcell*

In this case, the numerical fluxes $\mathcal{F}^{*,r}_{i-\frac{1}{2}}$ and $\mathcal{F}^{*,l}_{i+\frac{1}{2}}$ are defined by:

$$\mathcal{F}^{*,r}_{i-\frac{1}{2}} = \mathcal{F}\left(H^{*,-}_{i-\frac{1}{2}}, u^{-}_{i-\frac{1}{2}}; H^{*,+}_{i-\frac{1}{2}}, u^{+}_{i-\frac{1}{2}}\right) + \begin{pmatrix} 0 \\ \frac{1}{2}g(H^{+}_{i-\frac{1}{2}})^2 - \frac{1}{2}g(H^{*,+}_{i-\frac{1}{2}})^2 \end{pmatrix}, \tag{2.38}$$

$$\mathcal{F}^{*,l}_{i+\frac{1}{2}} = \mathcal{F}\left(H^{*,-}_{i+\frac{1}{2}}, u^{-}_{i+\frac{1}{2}}; \overline{H}^{-}_{1}, \overline{u}_1\right) + \begin{pmatrix} 0 \\ \frac{1}{2}g(H^{-}_{i+\frac{1}{2}})^2 - \frac{1}{2}g(H^{*,-}_{i+\frac{1}{2}})^2 \end{pmatrix}, \tag{2.39}$$

where

$$H^{*,-}_{i+\frac{1}{2}} = \max\left(0, H^{-}_{i+\frac{1}{2}} + b^{-}_{i+\frac{1}{2}} - \max\left(b^{-}_{i+\frac{1}{2}}, \overline{b}_1\right)\right) \tag{2.40}$$

and

$$\overline{H}^{-}_{1} = \max\left(0, \overline{H}_1 + \overline{b}_1 - \max\left(b^{-}_{i+\frac{1}{2}}, \overline{b}_1\right)\right), \tag{2.41}$$

using the simplified notations $\overline{b}_1 = \overline{b}_1^{\omega_{i+1}}$ and $\overline{H}_1 = \overline{H}_1^{\omega_{i+1}}$.

Lemma 19. Scheme (2.34), provided with the numerical fluxes defined in (2.38)-(2.39), does preserve the motionless steady states, as:

$$\forall \omega_i \in \mathscr{T}_h,\ \ \overline{H}_h^{\omega_i,n} + \overline{b}_h^{\omega_i} = \eta_c,\ \ \overline{u}_h^{\omega_i,n} = 0 \quad \Longrightarrow \quad \overline{H}_h^{\omega_i,n+1} + \overline{b}_h^{\omega_i} = \eta_c,\ \ \overline{u}_h^{\omega_i,n+1} = 0.$$

Proof. We assume that, at time level n, we have

$$u_h^{\omega_i,n} = \frac{q_h^{\omega_i,n}}{H_h^{\omega_i,n}} = 0 \quad \text{and} \quad H_h^{\omega_i,n} + b_h^{\omega_i} = \eta_c. \tag{2.42}$$

Let us show that the surface elevation remains equal to η_c and the water velocity remains equal to zero in ω_i at the next time level $n+1$. Under the steady state assumptions (2.42), we get $H^{*,-}_{i+\frac{1}{2}} = \overline{H}^{-}_{1}$. Thus,

$$\mathcal{F}^{*,l}_{i+\frac{1}{2}} = \begin{pmatrix} 0 \\ \frac{1}{2}g(H^{-}_{i+\frac{1}{2}})^2 \end{pmatrix} = F(v_h^{\omega_i}(x_{i+\frac{1}{2}})).$$

Similarly, since $H^{*,-}_{i-\frac{1}{2}} = H^{*,+}_{i-\frac{1}{2}}$, we get

$$\mathcal{F}^{*,r}_{i-\frac{1}{2}} = \begin{pmatrix} 0 \\ \frac{1}{2}g(H^{+}_{i-\frac{1}{2}})^2 \end{pmatrix} = F(v_h^{\omega_i}(x_{i-\frac{1}{2}})).$$

This leads us to:

$$\left[\mathcal{F}^*\psi_l^{\omega_i}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \left[F(v_h^{\omega_i}(x_{i+\frac{1}{2}}))\psi_l^{\omega_i}(x_{i+\frac{1}{2}}) - F(v_h^{\omega_i}(x_{i-\frac{1}{2}}))\psi_l^{\omega_i}(x_{i-\frac{1}{2}})\right] = \left[F(v_h^{\omega_i})\psi_l^{\omega_i}\right]_{i-\frac{1}{2}}^{i+\frac{1}{2}}.$$

Finally, the residual $\widetilde{\mathbf{l}}_l^{\omega_i}$ can be recast into:

$$\begin{aligned}\widetilde{\mathbf{l}}_l^{\omega_i} &= \int_{\omega_i} \psi_l^{\omega_i} B(v_h^{\omega_i}, \partial_x b_h^{\omega_i})\mathrm{d}x - \int_{\omega_i} \psi_l^{\omega_i} \partial_x F\left(v_h^{\omega_i}\right)\mathrm{d}x, \\ &= \int_{\omega_i} \left(B\left(v_h^{\omega_i}, \partial_x b_h^{\omega_i}\right) - \partial_x F\left(v_h^{\omega_i}\right)\right)\psi_l^{\omega_i}\mathrm{d}x \overbrace{=}^{(2.28)} 0, \qquad \forall\, l = 0, ..., k.\end{aligned} \tag{2.43}$$

Replacing $\widetilde{\mathbf{L}}_{\omega_i} = 0$ in (2.35), we finally get $\underline{v}_{\omega_i}^{n+1} = \underline{v}_{\omega_i}^{n}$. □

Let us still consider scheme (2.34) with the numerical fluxes defined in (2.38)-(2.39). Assuming that $H_{i+\frac{1}{2}}^-$, $\overline{H}_1$ and $H_{i-\frac{1}{2}}^\pm$ are non-negative, it is then easy to notice that $0 \leq H_{i+\frac{1}{2}}^{*,-} \leq H_{i+\frac{1}{2}}^-$, $0 \leq \overline{H}_1^- \leq \overline{H}_1$ and $0 \leq H_{i-\frac{1}{2}}^{*,\pm} \leq H_{i-\frac{1}{2}}^\pm$. Proceeding in a similar way as for the proof of Proposition 1, we obtain that the water-height mean-value $\overline{H}_{\omega_i}^{n+1}$ at time level $n+1$ on the cell ω_i is non-negative under a suitable CFL condition.

Remark 20. We can show the well-balanced property when only wet cells are involved. Indeed, a motionless steady-state is defined through the assumption $\eta = cst$ everywhere. This assumption does not hold anymore in the vicinity of a wet/dry interface, when higher-order polynomials are used to describe the solution (see Fig. 2.5). From a practical viewpoint, it is however possible to numerically preserve a motionless steady-states even in the presence of wet/dry interfaces with a fine enough mesh, allowing to maintain a very small computational error.

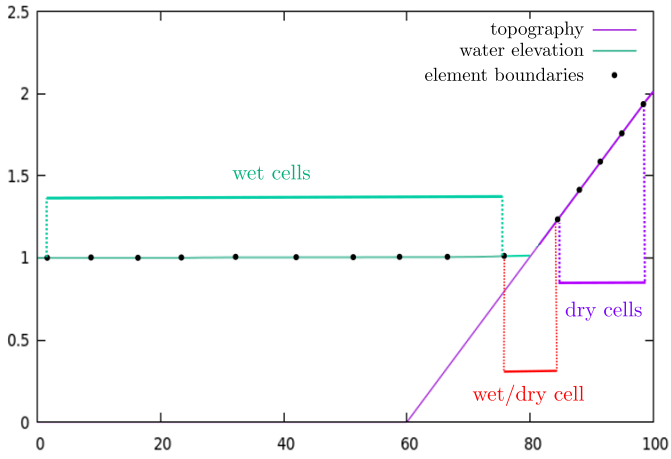

Figure 2.5: steady state

2.6 Numerical validations

This section is dedicated to the validation of the well-balanced *PL DG-FV*$_{subcell}$ method through the use of different standard test-cases.

2.6.1 Well-balancing property

In this test, we focus on the preservation of the motionless steady states in the case of a totally submerged bump. The computational domain is $\Omega = [0, 1]$. The topography profile is defined as follows

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{(x - x_1)\cdot\pi}{75}\right)\right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere,} \end{cases} \tag{2.44}$$

where $A = 4.75$, $x_1 = 0.125$ and $x_2 = 0.875$. The initial data is defined as

$$\eta_0(x) = 10 \qquad \text{and} \qquad q_0(x) = 0.$$

We evolve this initial configuration in time up to $t_{max} = 50s$, with a tenth-order approximation and 10 mesh elements. The numerical results obtained with the well-balanced *PL DG-FV*$_{subcell}$ method are shown on Fig. 2.6. In Table 2.1, we gather the global L^2-errors obtained for several orders of

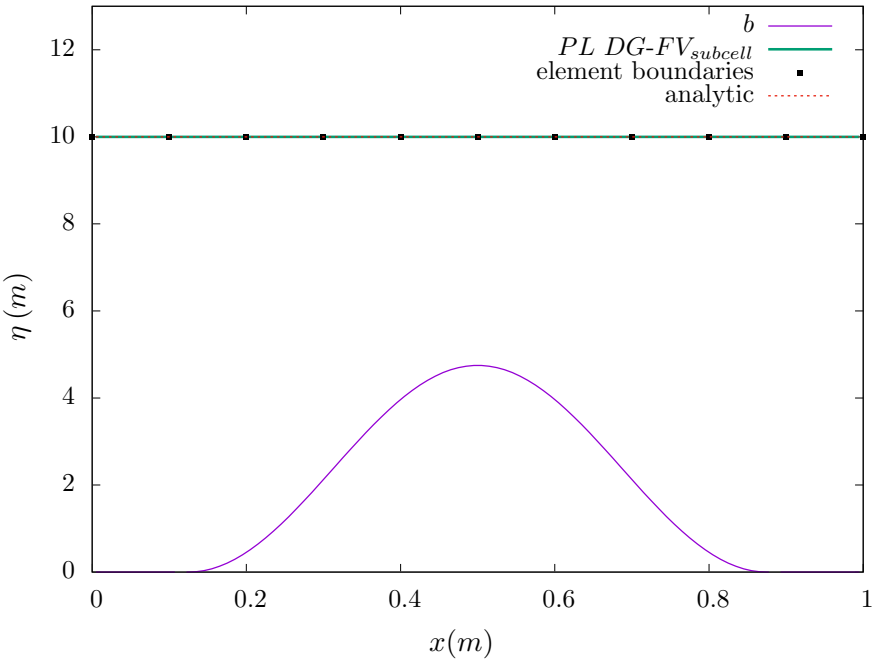

Figure 2.6: Test 1 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$, for $k = 9$ and $n_{\text{el}} = 10$.

approximation for the surface elevation at $t = 50s$. As expected, the steady state is preserved up to double precision accuracy.

k	1	2	3
h	$E^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$E^{\eta}_{L_2}$
$\frac{1}{15}$	1.26E-15	8.13E-16	9.48E-17
$\frac{1}{30}$	3.63E-16	1.77E-16	5.11E-17
$\frac{1}{60}$	1.53E-16	4.68E-17	1.01E-17
$\frac{1}{120}$	5.71E-17	1.38E-17	1.26E-18

Table 2.1: Test 1 - Preservation of a motionless steady state: L^2-errors between numerical and exact steady state solutions for η at time $t = 50s$.

Next, we slightly modify the initial condition for the water-height in order to have the bump above the water level:

$$\eta_0(x) = \max(3, b(x)) \qquad \text{and} \qquad q_0(x) = 0.$$

We evolve this initial configuration in time up to $t_{max} = 50s$, with a fourth-order approximation and 120 mesh elements. The numerical results obtained with the well-balanced *PL DG-FV$_{subcell}$* method are shown on Fig. 2.7.

Remark 21. We refine the mesh sufficiently so that the interface of the cell almost falls on the wet/dry transition point, see Fig. 2.8 (right).

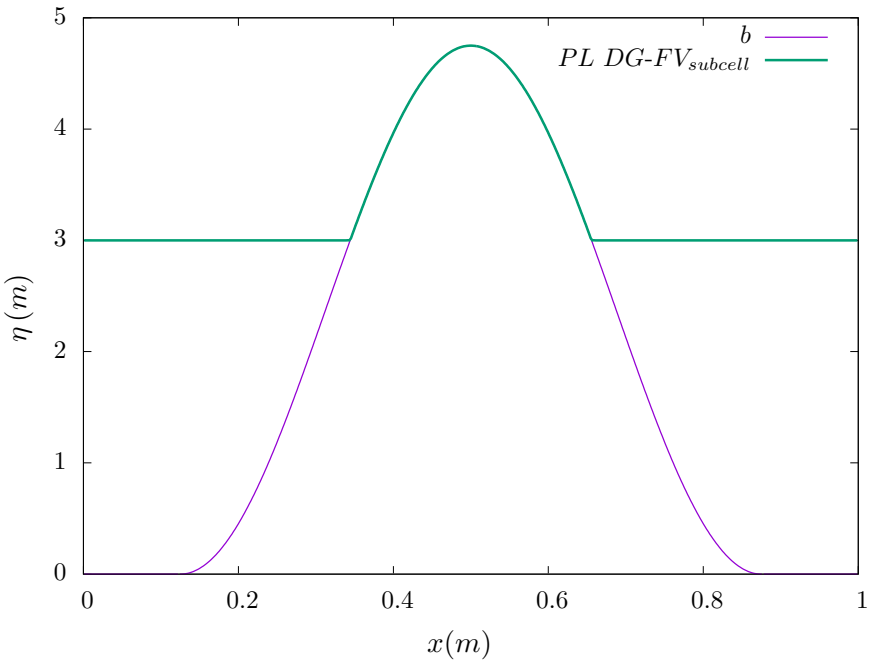

Figure 2.7: Test 2 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$, with $k = 3$ and $n_{\text{el}} = 120$.

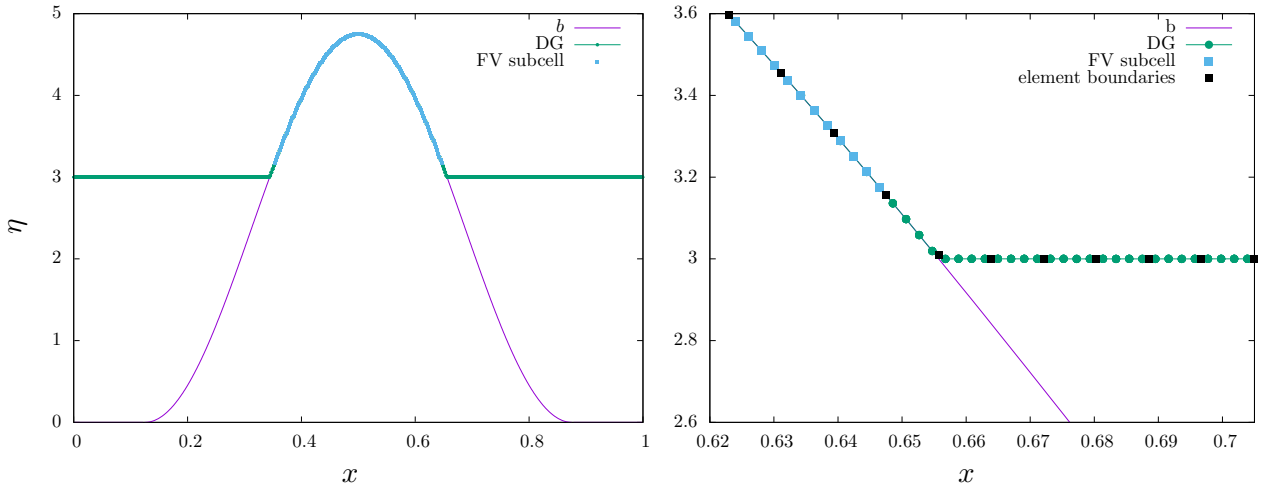

Figure 2.8: Test 2 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$, with $k = 3$ and $n_{\text{el}} = 120$ (left), with a zoom on the wet/dry interface (right).

We highlight in Fig. 2.8 the particular marked cells (blue squares), in which the FV-Subcell method has been performed. We emphasize that the steady state is effectively preserved up to the machine accuracy, validating numerically the compatibility of the *PL DG-FV$_{subcell}$* method with the well-balancing property for the wet/dry context. A similar behavior is reported for other orders of accuracy and number of cells.

2.6.2 Run-up of a solitary wave on a plane beach

The last test-case is devoted to the computation of the run-up of a solitary wave on a constant slope. Such run-up phenomena are investigated experimentally and numerically in [152]. In this test, a solitary wave traveling from the shoreward is let run-up and run-down on a plane beach, before being fully reflected and evacuated from the computational domain. The topography is made of a constant depth area juxtaposed with a plane sloping beach of constant slope α such that $\cot(\alpha) = 19.85$. The right boundary condition is transmissive. The initial condition is defined as follows:

$$\eta_0(x) = H_0 + \frac{A}{H_0}\,\text{sech}^2\left(\gamma\,(x - x_1)\right) \qquad \text{and} \qquad u_0(x) = \sqrt{\frac{g}{H_0}}\,(\eta_0(x) - H_0)\,,$$

with $\gamma = \sqrt{\dfrac{3A}{4H_0}}$, and where $x_1 = \sqrt{\frac{4H_0}{3A}}\,\text{arcosh}\left(\sqrt{\frac{1}{0.05}}\right)$ is nothing but the initial position of the center of the solitary wave. This test is run with $A = 0.019\,m$, $H_0 = 1.0\,m$, $n_{\text{el}} = 150$ and $t = 40\,s$. We show on Fig. 2.9 the free surface obtained with the well-balanced *PL DG-FV$_{subcell}$* method at several times in the range $[1\,s,\, 40\,s]$ for $k = 1$, and for $k = 3$ on Fig. 2.10, showing a good agreement with the reference solution obtained with a robust FV method on a very fine mesh $n_{\text{el}} = 10000$.

To be more precise, we see a good performance for the run-up of the solitary wave. As for the run-down, we notice small disturbances, specially for high-order of approximation ($k = 3$). In this case the solution is not robust enough, because during water run down, the wave breaks acting like discontinuities or steeply varying gradients and thus, weak spurious oscillations (or disturbances) have appeared (see Fig. 2.11). In order to solve this issue, a slope-limiter can be added (besides of the

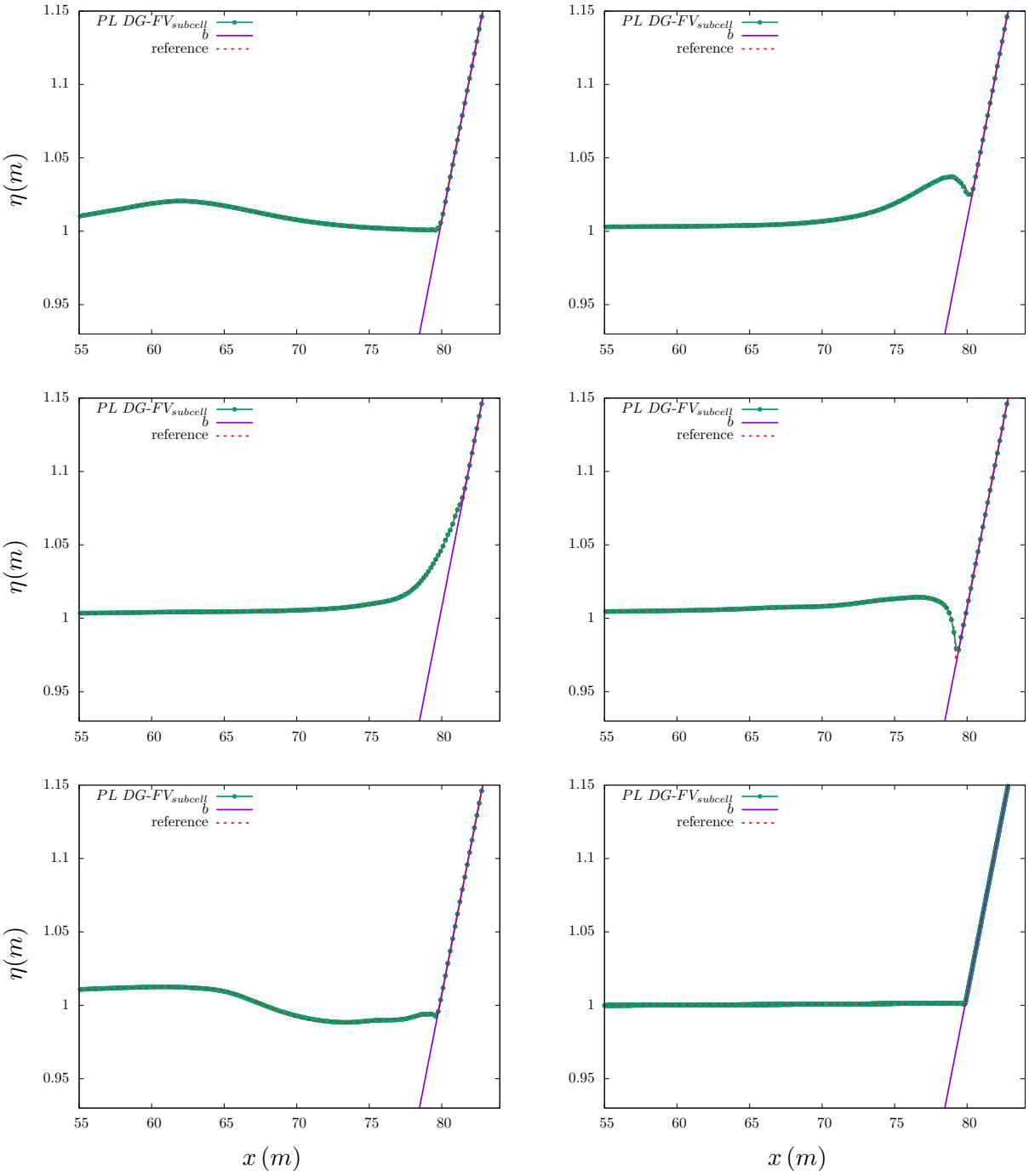

Figure 2.9: Test 3 - Run-up of a solitary wave on a plane beach - Free surface elevation computed for different values of time in the range $[1\,s, t = 40\,s]$ with the *PL DG-FV$_{subcell}$* method obtained for $k = 1$ and $n_{\text{el}} = 300$.

positivity-preserving limiter), as the TVB slope-limiter § 2.2. We are talking here about an *a priori* correction procedure. Another method that allows us to solve the problem of spurious oscillations, in addition to being a substitute for the positivity-preserving limiter, is the *a posteriori* LSC method

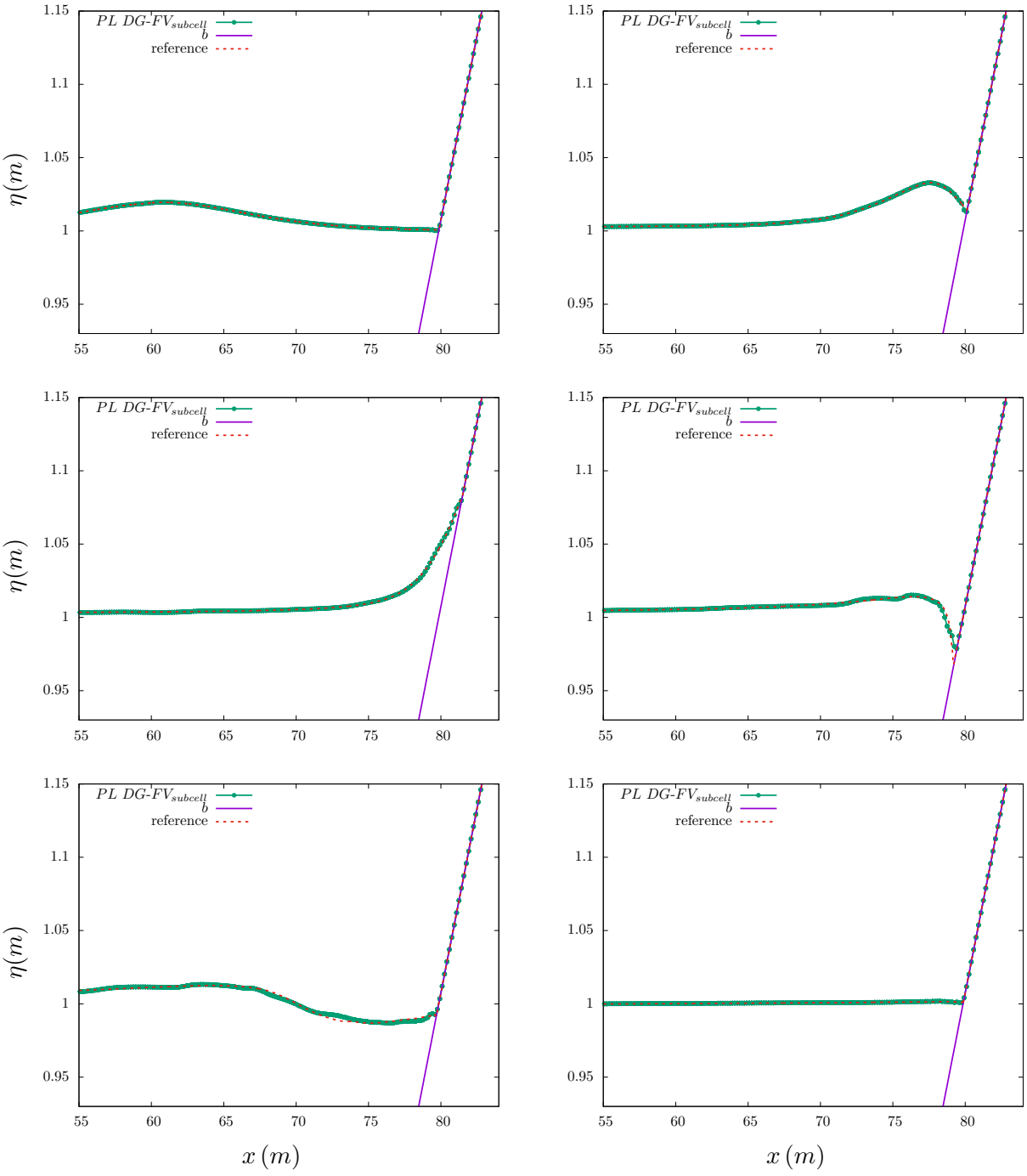

Figure 2.10: Test 3 - Run-up of a solitary wave on a plane beach - Free surface elevation computed for different values of time in the range $[1\,s, t = 40\,s]$ with the *PL DG-FV$_{subcell}$* method obtained for $k = 3$ and $n_{\text{el}} = 150$.

that will be introduced in details in the next chapter. The main tool of this correction procedure is a first-order FV scheme, which is a positivity-preserving scheme and shocks capturing (*i.e.* robust scheme). In chapter 3, we take advantage of these first-order FV robustness properties to reach our

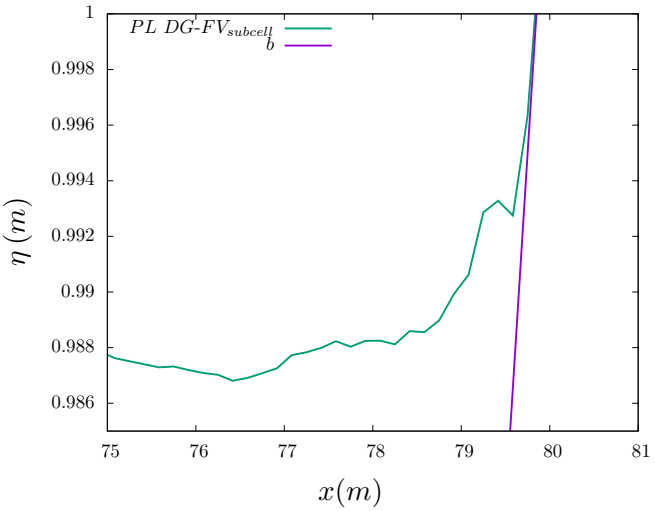

Figure 2.11: Test 3 - Run-up of a solitary wave on a plane beach - Free surface elevation computed at $t = 22\,s$ with the *PL DG-FV$_{subcell}$* method obtained for $k = 3$ and $n_{\text{el}} = 150$, with a zoom on the shoreline.

goal by solving these robustness issues while maintaining a high-order DG subcell accuracy.

Chapter 3

An *a posteriori* DG-LSC method for the NSW equations

Robustness issues may be among the main remaining challenges for the use of high-order DG methods in realistic problems for many domains of applications. In recent years, several approaches have been proposed to stabilize high-order approximations. As we mentioned in § 1.2, these techniques mainly rely on two different paradigms that we referred to as *a priori* and *a posteriori*.
In the paradigm of *a posteriori* correction, an uncorrected candidate solution is first computed at the new time-step. The candidate solution is then checked according to some admissibility criteria. If the solution is considered admissible, we go further in time. Otherwise, we return to the previous time-step and correct locally the numerical solution by making use of a more robust scheme. Recently, some new *a posteriori* limitations have arisen. Let us mention the so-called MOOD technique, see [42, 55, 56]. Through this procedure, the order of approximation of the numerical scheme is locally reduced in an *a posteriori* sequence until the solution becomes admissible. In [61, 59, 88], a FV-Subcell technique similar to the one presented in [151] has been applied to the *a posteriori* paradigm. Practically, if the numerical solution in a cell is detected as bad, the cell is then subdivided into subcells and a first-order FV, or alternatively other robust scheme (second-order TVD FV scheme or WENO scheme for instance), is applied on each subcell. Then, through these new subcell mean-values, a high-order polynomial is reconstructed on the primal cell. Related strategies applied to dispersive and turbulent shallow-water flows have been introduced in [30, 31].

Making use of the *a posteriori* LSC method introduced in [164] for general hyperbolic conservation laws, the main objective in this chapter is to develop a novel shock-capturing, positivity-preserving and well-balanced DG method for the NSW equations with topography source term by using specific local flux correction at the subcell level, with *a posteriori* numerical admissibility detectors. To be more precise, *a posteriori* correction is only applied locally at the subcell level where it is absolutely needed (*i.e.* only non-admissible subcells are marked), while not neglecting the scheme conservation property. In practice, we first reformulate DG scheme as a FV-like subcell schemes provided the use of the so-called DG reconstructed flux. Then, the correction procedure is done as follows: at each SSP-RK time-step, we compute a high-order DG candidate solution and check its admissibility (non-negative water-height and no spurious oscillations). If the solution is admissible, we go further in time. If it is not the case, we go back to the previous time-step and correct locally at the subcell level the non-admissible local numerical solution. Actually, we divide the cell into subcells, then, if the solution at a specific subcell is detected as bad, we substitute the DG reconstructed flux on

the subcell boundaries by a robust low-order FV numerical flux. Otherwise, if the solution is detected as admissible on this subcell, we keep the high-order DG reconstructed numerical flux. The purpose of applying this correction procedure is to enforce the water-height positivity and to avoid spurious oscillations in the vicinity of solution's singularity while preserving as much as possible the high accuracy and the very precise subcell resolution of high-order DG schemes, by minimizing the number of subcells in which the solution has to be recomputed. To this end, we pay attention to keep the scheme local conservation property. Not only the solutions are recomputed in the troubled subcells, but also in its first neighbors, so that we can have the same numerical fluxes on both sides of subcells interfaces. To complete the picture, it remains to ensure the well-balanced property for our *a posteriori* LSC scheme in wet-wet and wet-dry contexts.

The well-balanced property for NSW equations, first introduced in [15], has been widely studied in recent years. Following the ideas of [62], we use the so-called *pre-balanced* formulation of the NSW equations. Indeed, the alternative formulation of the NSW equations in deviatoric form, obtained by subtracting an equilibrium solution, and introduced in [145] is interesting as it leads to a balanced set of hyperbolic equations that does not require specific numerical algorithms to obtain a well-balanced property. However, such a formulation is given in terms of free surface elevation above the still water level (denoted ζ in (A.3)), and is therefore not suitable to model cases involving dry areas (the still water depth is undefined in dry areas). In [115, 116], a new formulation given in terms of the total free surface elevation $\eta = H + b$ (see Fig. 3.1) allows to alleviate this drawback, allowing to study cases involving dry areas. Such a formulation is very suitable to show the well-balanced property due to assumptions ($\eta = \eta^c$) that simplifie the calculation procedure. In addition, the *pre-balanced* formulation can be less computationally expensive than the (H, q) formulation (A.38), for example, in computing volume integrals and source terms numerically (see Remark 23).

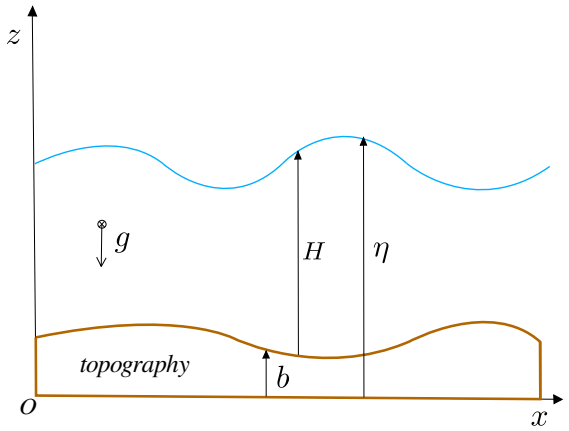

Figure 3.1: Free surface flow: main notations for the *pre-balanced* formulation

Indeed, observing that

$$\frac{1}{2}g\partial_x H^2 + gH\partial_x b = \frac{1}{2}g\partial_x(\eta^2 - 2\eta b) + g\eta\partial_x b,$$

we obtain the so-called *pre-balanced* form of the NSW equations, given in a compact form:

$$\partial_t \mathbf{v} + \partial_x \mathbf{F}(\mathbf{v}, b) = \mathbf{B}(\mathbf{v}, \partial_x b), \tag{3.1}$$

where $\mathbf{v} : \mathbb{R} \times \mathbb{R}_+ \to \Theta$ is the vector of conservative variables, $\mathbf{F} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the flux function and $\mathbf{B} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the topography source term, defined as follows:

$$\mathbf{v} = \begin{pmatrix} \eta \\ q \end{pmatrix}, \ \mathbf{F}(\mathbf{v}, b) = \begin{pmatrix} q \\ uq + \frac{1}{2} g(\eta^2 - 2\,\eta\, b) \end{pmatrix}, \quad \mathbf{B}(\mathbf{v}, \partial_x b) = \begin{pmatrix} 0 \\ -g\,\eta\,\partial_x b \end{pmatrix}. \tag{3.2}$$

Finally, the proposed strategy is investigated through an extensive set of benchmarks, including a brand new smooth solution for the computation of convergence rates, stabilization of flows with discontinuities, the preservation of motionless steady states, or moving shorelines over varying bottoms. We observe in particular that this approach provides a very accurate description of wet/dry interfaces, even with the use of very high-order schemes on coarse meshes, showing the subcell resolution ability of the resulting high-order DG scheme.

Remark 22. The notations used in this chapter are defined either in this chapter or in chapter 1, section § 1.3.

Let start by presenting the DG formulation for the *pre-balanced* NSW equations (3.1).

3.1 DG formulation

Let $b_h = i^k_{\mathscr{T}_h}(b)$ denote a globally continuous piecewise polynomial approximation of the topography parametrization. A straightforward semi-discrete in space DG approximation of (3.1) reads: find $\mathbf{v}_h = (\eta_h, q_h) \in (\mathbb{P}^k(\mathscr{T}_h))^2$ such that, for all $\varphi \in \mathbb{P}^k(\mathscr{T}_h)$,

$$\int_{\mathscr{T}_h} \partial_t \mathbf{v}_h \varphi dx + \int_{\mathscr{T}_h} \mathcal{A}_h(\mathbf{v}_h)\varphi dx = 0, \tag{3.3}$$

where the discrete nonlinear operator $\mathcal{A}_h$ is defined by

$$\int_{\mathscr{T}_h} \mathcal{A}_h(\mathbf{v}_h)\varphi dx := -\sum_{\omega\in\mathscr{T}_h} \int_\omega \mathbf{F}(\mathbf{v}_h, b_h)\partial_x\varphi dx + \sum_{\omega\in\mathscr{T}_h} \left[\varphi\,\mathcal{F}\right]_{\partial\omega} - \sum_{\omega\in\mathscr{T}_h} \int_\omega \mathbf{B}(\mathbf{v}_h, \partial_x b_h)\varphi dx, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h). \tag{3.4}$$

In (3.4), $\mathcal{F}$ stands for the interface numerical flux function. Denoting by $\mathbf{v}^-_{i+\frac{1}{2}}$ and $\mathbf{v}^+_{i+\frac{1}{2}}$, respectively the left and right traces of $\mathbf{v}_h$ on interface $x_{i+\frac{1}{2}}$, and by $b_{i+\frac{1}{2}} = b^-_{i+\frac{1}{2}} = b^+_{i+\frac{1}{2}}$ the trace of b_h, we define the numerical flux function $\mathcal{F}_{i+\frac{1}{2}}$ on interface $x_{i+\frac{1}{2}}$ as follows:

$$\mathcal{F}_{i+\frac{1}{2}} := \mathcal{F}(\mathbf{v}^-_{i+\frac{1}{2}}, \mathbf{v}^+_{i+\frac{1}{2}}, b_{i+\frac{1}{2}}), \tag{3.5}$$

where the numerical flux function chosen here is the simple global LF flux:

$$\mathcal{F}(\mathbf{v}^-, \mathbf{v}^+, b) := \frac{1}{2}\left(\mathbf{F}(\mathbf{v}^-, b) + \mathbf{F}(\mathbf{v}^+, b) - \sigma(\mathbf{v}^+ - \mathbf{v}^-)\right), \tag{3.6}$$

with $\sigma := \max_{\omega\in\mathscr{T}_h} \sigma_\omega$ and

$$\sigma_\omega := \max_m \left(|\overline{u}^\omega_m| + \sqrt{g\overline{H}^\omega_m}\right).$$

Remark 23. We require that the volume integral and source term in formula (3.4) are exactly computed at motionless steady states. This can be achieved, for the *pre-balanced* formulation (3.1)-(3.2), by using any quadrature rule exact for polynomials of degree up to $2k$, thanks to the *pre-balanced* formulation (3.1)-(3.2). On the other hand, for the classical NSW formulation (A.38), a quadrature rule exact for polynomials of degree up to $3k$ is needed.

Remark 24. The topography b is interpolated by b_h through a piecewise polynomial but globally continuous function over the mesh. To achieve this, one can simply choose the elements boundaries among the interpolation points with any corresponding interpolation method. To ensure that the scheme is indeed well-balanced, and particularly in wet/dry context, see § 3.8, we initialize the surface elevation η_h in dry areas by setting $\eta_h = b_h$. Thus, water-height positivity is also assured in dry areas since $H_h = \eta_h - b_h = 0$, by construction. We emphasize that as long as $H = \eta - b$ is non-negative, its subcell mean-values are also non-negative. However, nothing ensures that after performing a L^2 projection of the initial water-height, the associated sub-mean-values of the L^2 projection H_h^ω are positive. This is the reason why, in wet regions, for the initialization, we start by computing the positive H sub-mean-values using (3.7) and then reconstruct the associated polynomial using $\mathbf{\Pi}_\omega^{-1}$.

$$\overline{H}_m^\omega = \frac{1}{|S_m^\omega|}\int_{S_m^\omega} H(x)dx. \tag{3.7}$$

In the following, similarly to what has been done in [164], we demonstrate an equivalence relation between (3.3) and a FV-like method on a sub-mesh.

3.2 DG formulation as a FV-like scheme on a sub-grid

Let introduce the L^2-projections of the flux function $\mathbf{F}_h = p^k_{\mathscr{T}_h}(\mathbf{F}(\mathbf{v}_h, b_h))$ and of the source term $\mathbf{B}_h = p^k_{\mathscr{T}_h}(\mathbf{B}(\mathbf{v}_h, \partial_x b_h))$, such that:

$$\int_{\mathscr{T}_h} \mathbf{F}(\mathbf{v}_h, b_h)\varphi dx = \int_{\mathscr{T}_h} \mathbf{F}_h\varphi dx, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h), \tag{3.8a}$$

$$\int_{\mathscr{T}_h} \mathbf{B}(\mathbf{v}_h, \partial_x b_h)\varphi dx = \int_{\mathscr{T}_h} \mathbf{B}_h\varphi, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h). \tag{3.8b}$$

Remark 25. As we mentioned, in DG schemes, volume integral and source term contribution are computed using quadrature rule. This quadrature rule should be used to compute the left hand side of the L^2 projections (3.8a) and (3.8b).

From (3.3), we now have:

$$\int_{\mathscr{T}_h} \partial_t \mathbf{v}_h\varphi dx - \sum_{\omega\in\mathscr{T}_h}\int_\omega \mathbf{F}(\mathbf{v}_h, b_h)\partial_x\varphi dx + \sum_{\omega\in\mathscr{T}_h}\left[\varphi\,\mathcal{F}\right]_{\partial\omega} - \sum_{\omega\in\mathscr{T}_h}\int_\omega \mathbf{B}(\mathbf{v}_h, \partial_x b_h)\varphi dx = 0, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h),$$

which is the so called strong DG scheme. Using the L^2 projections (3.8a) and (3.8b):

$$\int_{\mathscr{T}_h} \partial_t \mathbf{v}_h\varphi dx - \sum_{\omega\in\mathscr{T}_h}\int_\omega \mathbf{F}_h\partial_x\varphi dx + \sum_{\omega\in\mathscr{T}_h}\left[\varphi\,\mathcal{F}\right]_{\partial\omega} - \sum_{\omega\in\mathscr{T}_h}\int_\omega \mathbf{B}_h\varphi dx = 0, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h),$$

or equivalently, using an integration by parts:

$$\int_{\mathscr{T}_h} \partial_t \mathbf{v}_h \varphi dx + \sum_{\omega\in\mathscr{T}_h} \int_\omega \partial_x \mathbf{F}_h \varphi dx - \sum_{\omega\in\mathscr{T}_h} \left[\varphi\ (\mathbf{F}_h - \mathcal{F})\ \right]_{\partial\omega} - \sum_{\omega\in\mathscr{T}_h} \int_\omega \mathbf{B}_h \varphi dx = 0, \quad \forall \varphi \in \mathbb{P}^k(\mathscr{T}_h), \tag{3.9}$$

Substituting ϕ_m^ω, defined in (1.6), into (3.9) gives the local equations on mesh element $\omega \in \mathscr{T}_h$:

$$\int_\omega \partial_t \mathbf{v}_h^\omega \phi_m^\omega dx = -\int_\omega \partial_x \mathbf{F}_h^\omega \phi_m^\omega dx + \int_\omega \mathbf{B}_h^\omega \phi_m^\omega dx + \left[(\mathbf{F}_h^\omega - \mathcal{F})\,\phi_m^\omega\right]_{\partial\omega}, \qquad \forall m \in [\![1,\, k+1]\!].$$

Since $\partial_t \mathbf{v}_h^\omega$, $\partial_x \mathbf{F}_h^\omega$ and $\mathbf{B}_h^\omega$ belong to $(\mathbb{P}^k(\omega))^2$ and considering (1.7), it follows that,

$$\int_{S_m^\omega} \partial_t \mathbf{v}_h^\omega dx = -\int_{S_m^\omega} \partial_x \mathbf{F}_h^\omega dx + \int_{S_m^\omega} \mathbf{B}_h^\omega dx + \left[(\mathbf{F}_h^\omega - \mathcal{F})\,\phi_m^\omega\right]_{\partial\omega}, \qquad \forall m \in [\![1,\, k+1]\!].$$

and then,

$$\partial_t \overline{\mathbf{v}}_m^\omega = -\frac{1}{|S_m^\omega|}\left(\left[\mathbf{F}_h^\omega\right]_{\partial S_m^\omega} - \left[\phi_m^\omega\,(\mathbf{F}_h^\omega - \mathcal{F})\,\right]_{\partial\omega}\right) + \overline{\mathbf{B}}_m^\omega, \quad \forall m \in [\![1,\, k+1]\!],$$

where $\overline{\mathbf{v}}_m^\omega$ and $\overline{\mathbf{B}}_m^\omega$ are respectively the mean-values of $\mathbf{v}_h^\omega$ and $\mathbf{B}_h^\omega$ on the subcell S_m^ω, defined by:

$$\overline{\mathbf{v}}_m^\omega = \frac{1}{S_m^\omega}\int_{S_m^\omega} \mathbf{v}_h^\omega dx \ \text{ and } \ \overline{\mathbf{B}}_m^\omega = \frac{1}{S_m^\omega}\int_{S_m^\omega} \mathbf{B}_h^\omega dx. \tag{3.10}$$

Let introduce the $k+2$ subcells interfaces fluxes $\left\{\widehat{\mathbf{F}}^{\,\omega}_{m+\frac{1}{2}}\right\}_{m\in[\![0,\,k+1]\!]}$ such that:

$$\widehat{\mathbf{F}}^{\,\omega}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}^{\,\omega}_{m-\frac{1}{2}} = \left[\mathbf{F}_h^\omega\right]_{\partial S_m^\omega} - \left[\phi_m^\omega\,(\mathbf{F}_h^\omega - \mathcal{F})\,\right]_{\partial\omega}, \quad \forall m \in [\![1,\, k+1]\!], \tag{3.11}$$

so that we have

$$\partial_t \overline{\mathbf{v}}_m^\omega = -\frac{1}{|S_m^\omega|}\left(\widehat{\mathbf{F}}^{\,\omega}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}^{\,\omega}_{m-\frac{1}{2}}\right) + \overline{\mathbf{B}}_m^\omega. \tag{3.12}$$

Formulation (3.12) can be seen as a FV-like scheme on subcell S_m^ω. The values $\left\{\widehat{\mathbf{F}}^{\,\omega}_{m+\frac{1}{2}}\right\}_{m\in[\![0,\,k+1]\!]}$ is thereafter referred to as *reconstructed fluxes*. Considering the mesh element $\omega_i = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right] \in \mathscr{T}_h$, and setting the first and last reconstructed fluxes to the DG numerical flux values at cell boundaries such as:

$$\widehat{\mathbf{F}}^{\,\omega_i}_{\frac{1}{2}} := \mathcal{F}_{i-\frac{1}{2}} \qquad \text{and} \qquad \widehat{\mathbf{F}}^{\,\omega_i}_{k+\frac{3}{2}} := \mathcal{F}_{i+\frac{1}{2}}. \tag{3.13}$$

The linear system (3.11)-(3.13) is straightforward to solve. Indeed, substituting subscript m by p in (3.11) and summing for p from 1 to m leads to

$$\begin{aligned}\widehat{\mathbf{F}}^{\,\omega_i}_{m+\frac{1}{2}} = \mathbf{F}_h^{\omega_i}\left(\widetilde{x}^{\,\omega_i}_{m+\frac{1}{2}}\right) \quad &- \quad \left(1 - \textstyle\sum_{p=1}^m \phi_p^{\omega_i}\left(x_{i-\frac{1}{2}}\right)\right)\left(\mathbf{F}_h^{\omega_i}\left(x_{i-\frac{1}{2}}\right) - \mathcal{F}_{i-\frac{1}{2}}\right) \\ &- \quad \left(\textstyle\sum_{p=1}^m \phi_p^{\omega_i}\left(x_{i+\frac{1}{2}}\right)\right) \quad \left(\mathbf{F}_h^{\omega_i}\left(x_{i+\frac{1}{2}}\right) - \mathcal{F}_{i+\frac{1}{2}}\right)\end{aligned}$$

the m interior reconstructed fluxes expression. We recast those expressions into:

$$\widehat{\mathbf{F}}^{\,\omega_i}_{m+\frac{1}{2}} = \mathbf{F}_h^{\omega_i}\left(\widetilde{x}^{\,\omega_i}_{m+\frac{1}{2}}\right) - C^{i-\frac{1}{2}}_{m+\frac{1}{2}}\left(\mathbf{F}_h^{\omega_i}(x_{i-\frac{1}{2}}) - \mathcal{F}_{i-\frac{1}{2}}\right) - C^{i+\frac{1}{2}}_{m+\frac{1}{2}}\left(\mathbf{F}_h^{\omega_i}(x_{i+\frac{1}{2}}) - \mathcal{F}_{i+\frac{1}{2}}\right), \quad \forall m \in [\![1,\, k+1]\!], \tag{3.14}$$

where the $C^{i\pm\frac{1}{2}}_{m+\frac{1}{2}}$ are defined by:

$$C^{i-\frac{1}{2}}_{m+\frac{1}{2}} = \sum_{p=m+1}^{k+1} \phi_p^{\omega_i}(x_{i-\frac{1}{2}}) \qquad \text{and} \qquad C^{i+\frac{1}{2}}_{m+\frac{1}{2}} = \sum_{p=1}^{m} \phi_p^{\omega_i}(x_{i+\frac{1}{2}}). \tag{3.15}$$

As detailled in [164], it is possible to explicitly express the $k+2$ correction coefficients $C^{i\pm\frac{1}{2}}_{m+\frac{1}{2}}$. To this end, let us set $J \in \mathbb{R}^{k+1}$ to be the vector defined as

$$J_j = (-1)^{j+1} \begin{pmatrix} k+j \\ j \end{pmatrix} \begin{pmatrix} k+1 \\ j \end{pmatrix},$$

where $\begin{pmatrix} p \\ j \end{pmatrix}$ stands for the binomial coefficient $\begin{pmatrix} p \\ j \end{pmatrix} = \frac{p!}{j!(p-j)!}$. Let us note that vector J only depends on the degree of approximation k, and not on the flux points position. By introducing $\left\{\widetilde{\xi}_{m+\frac{1}{2}}\right\}_m$ the flux points counterpart in the referential element $[0,1]$, as $\widetilde{\xi}_{m+\frac{1}{2}} = \frac{\widetilde{x}_{m+\frac{1}{2}} - x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}}$, the correction coefficients finally write

$$C^{i-\frac{1}{2}}_{m+\frac{1}{2}} = 1 - \begin{pmatrix} \widetilde{\xi}_{m+\frac{1}{2}} \\ \left(\widetilde{\xi}_{m+\frac{1}{2}}\right)^2 \\ \vdots \\ \left(\widetilde{\xi}_{m+\frac{1}{2}}\right)^{k+1} \end{pmatrix} \cdot J.$$

For further details, we refer the reader to [164].

Remark 26. One can see that the reconstructed flux is nothing but the polynomial interior flux $\mathbf{F}_\omega$, plus some correction terms taking into account the jump in fluxes at cell boundary $\partial\omega$.

Remark 27. Let us note that if this particular definition of $\{C^{i\pm\frac{1}{2}}_{m+\frac{1}{2}}\}_{m\in[\![0,\,k+1]\!]}$, (3.15), gives the equivalence with DG schemes, other choices obviously lead to other schemes. For instance, if one set these constants to zero, except for the first and last to be one, one would then recover the Spectral Volume (SV) method, [172, 81]. Indeed, even if the interior flux, here referred to as reconstructed flux, is continuous inside the cells, as for the SV methods, those two methods are still quite different. In SV methods, the interior flux is nothing but the flux function applied to the polynomial solution, *i.e.* $\mathbf{F}(\mathbf{v}_h, b_h)$. Here, the reconstructed flux can be seen as some specific approximation of the L_2 projection of such function but prescribing its boundary values to be the DG numerical fluxes. Hence, such an approach may be rather compared to the Flux Reconstruction method, also referred to as CPR method (Correction Procedure through Reconstruction), with which we share this reconstructed fluxes framework, see for instance [86, 170, 4, 69, 85] or the dedicated paragraph in [164] for more insight on the analogy between the present theory and Flux Reconstruction schemes.

Remark 28. The choice of the sub-partition points, $\{\widetilde{x}^{\,\omega}_{m+\frac{1}{2}}\}_{m\in[\![0,\,k+1]\!]}$, has already been discussed in [164]. It appeared that, regarding the reformulation of DG schemes into FV-Subcell methods, the cell decomposition into subcells does not come into account, as any choice would lead to the same piecewise polynomial solution. However, for the correction procedure introduced in § 3.4, the sub-division does has a slight impact. Indeed, the use of a non-uniform sub-partition, for instance by means of the Gauss-Lobatto points, leads to better results compared to a uniform sub-division. This is more likely the manifestation of the Runge phenomenon in the context of histopolation, as the histopolation basis functions underlying the sub-mean-value representation, are more oscillatory for a uniform cell sub-partition. Consequently, in the remainder, we make use of Gauss-Lobatto points to define the sub-partition points $\{\widetilde{x}^{\,\omega}_{m+\frac{1}{2}}\}_{m\in[\![0,\,k+1]\!]}$.

3.3 Time marching algorithm

Supplementing (3.3) with an initial datum $\mathbf{v}(0,\cdot)=\mathbf{v}_0=(\eta_0,q_0)^{\mathrm{t}}$, the time-stepping may be carried out using explicit SSP-RK schemes, [74, 149], see §2.3 for explicit discretisation example. The discrete initial data $\mathbf{v}_h^0$ is defined as in Remark 24. As the correction described in the following section make use of both DG scheme on the primal cells $\omega\in\mathscr{T}_h$ and FV scheme on the subcells $S_m^\omega\in\mathscr{T}_\omega$, the time-step Δt^n is computed adaptively using the same CFL condition introduced in (2.8):

$$\Delta t^n=\frac{\min\limits_{\omega\in\mathscr{T}_h}\left(\dfrac{h_\omega}{2k+1},\ \min\limits_{S_m^\omega\in\mathscr{T}_\omega}|S_m^\omega|\right)}{\sigma}, \tag{3.16}$$

where σ is the constant previously introduced in the global LF numerical flux definition (3.6).

3.4 *A posteriori* local subcell correction

In this section, we show how it is possible to modify the reconstructed fluxes $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ in a robust way in subcells where the *uncorrected* DG scheme (3.3) has failed, either by obtaining negative value for the water-height or by generating nonphysical oscillations due to the Gibbs phenomenon in the vicinity of discontinuities. For sake of conciseness in notations, the superscript ω_i may be avoided in the following when no confusion is possible.

Once again, as high-order RK SSP time marching algorithms may be regarded as convex combinations of first-order forward Euler schemes, we consider in the following, for sake of simplicity, a fully discrete formulation obtained from (3.3) and a first-order forward Euler scheme. We assume that at time level n the numerical solution $\mathbf{v}_h^n$ is *admissible* in a sense to be clarified later. We then compute an updated candidate solution $\mathbf{v}_h^{n+1}$ through the *uncorrected* DG scheme (3.3). If the candidate $\mathbf{v}_h^{n+1}$ is admissible, no correction is needed. Otherwise, the *uncorrected* DG scheme has produced an updated solution $\mathbf{v}_h^{n+1}$ which is not admissible on at least one particular mesh element cell ω_i. Looking at the subcell level, and assuming that $\mathbf{v}_h^{\omega,n+1}$ is not admissible in the particular subcell $S_m\in\mathscr{T}_{\omega_i}$, which is thus called a *troubled subcell* in the following, the main idea of our *a posteriori* LSC method is to replace the incriminated subcell mean-value $\overline{\mathbf{v}}_m^{n+1}$ by a new one, denoted with a $\star$ as follows $\overline{\mathbf{v}}_m^{\star,n+1}$, which is computed using a first-order FV-Subcell scheme of the form:

$$\overline{\mathbf{v}}_m^{\star,n+1} = \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|}\left(\mathcal{F}^l_{m+\frac{1}{2}} - \mathcal{F}^r_{m-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_m, \tag{3.17}$$

with some new subcell *lowest-order corrected* numerical fluxes $\mathcal{F}^l_{m+\frac{1}{2}}$, $\mathcal{F}^r_{m-\frac{1}{2}}$ which are defined hereafter. Indeed, because the *uncorrected* DG scheme (3.3) is equivalent to the subcell FV-like scheme (3.12) with high-order reconstructed fluxes (3.14), we propose to substitute, at the boundaries of S_m, the high-order reconstructed fluxes with first-order FV numerical fluxes. Finally, new degrees of freedom at discrete time t^{n+1} are computed from the modified set of sub-mean-values, now given as a blend of uncorrected values $\overline{\mathbf{v}}_m^{n+1}$ and corrected values $\overline{\mathbf{v}}_m^{\star,n+1}$. This strategy is illustrated in Fig. 3.2, where the marked subcell is identified with red color.

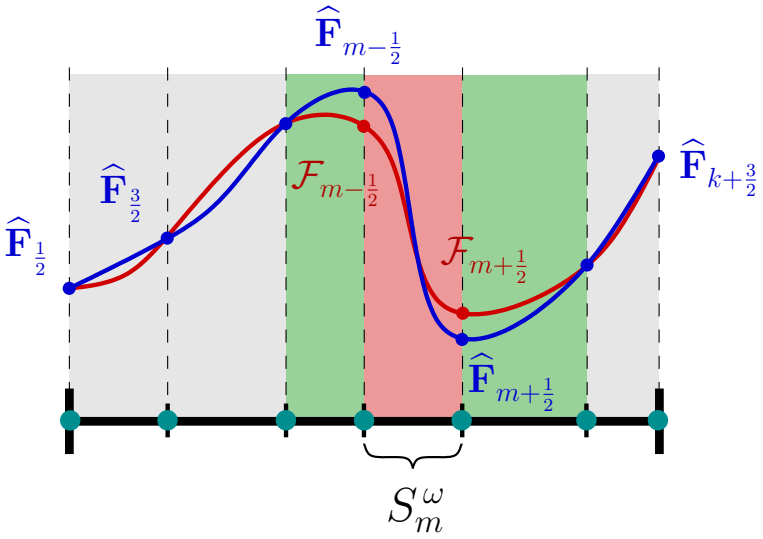

Figure 3.2: Sketch of the correction of the reconstructed fluxes at subcell boundaries

Additionally, to preserve the local conservation property of the resulting scheme, the left and right neighboring subcells, colored in green in Fig. 3.2, have to be updated too, even if they are flagged as admissible subcells, since we have substituted the reconstructed fluxes $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ with corrected ones. In the particular case depicted in Fig. 3.2 where S_{m-2} and S_{m+2} are also flagged as admissible, the sub-mean-values $\overline{\mathbf{v}}_{m+1}^{n+1}$ and $\overline{\mathbf{v}}_{m-1}^{n+1}$ are thus replaced respectively by $\overline{\mathbf{v}}_{m-1}^{\star,n+1}$ and $\overline{\mathbf{v}}_{m+1}^{\star,n+1}$ computed through a high-order reconstructed flux on one end and a first-order FV numerical flux on the other end, as follows:

$$\overline{\mathbf{v}}_{m-1}^{\star,n+1} = \overline{\mathbf{v}}_{m-1}^n - \frac{\Delta t^n}{|S_{m-1}|}\left(\mathcal{F}^l_{m-\frac{1}{2}} - \widehat{\mathbf{F}}_{m-3/2}\right) + \Delta t^n \overline{\mathbf{B}}_{m-1}, \tag{3.18}$$

$$\overline{\mathbf{v}}_{m+1}^{\star,n+1} = \overline{\mathbf{v}}_{m+1}^n - \frac{\Delta t^n}{|S_{m+1}|}\left(\widehat{\mathbf{F}}_{m+3/2} - \mathcal{F}^r_{m+\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_{m+1}. \tag{3.19}$$

For all the remaining admissible subcells (left in grey on Fig. 3.2), because the associated reconstructed fluxes are not corrected, they do not require any further computation, and the corresponding sub-mean-values are the values obtained through the *uncorrected* DG scheme, see Fig. 3.3.

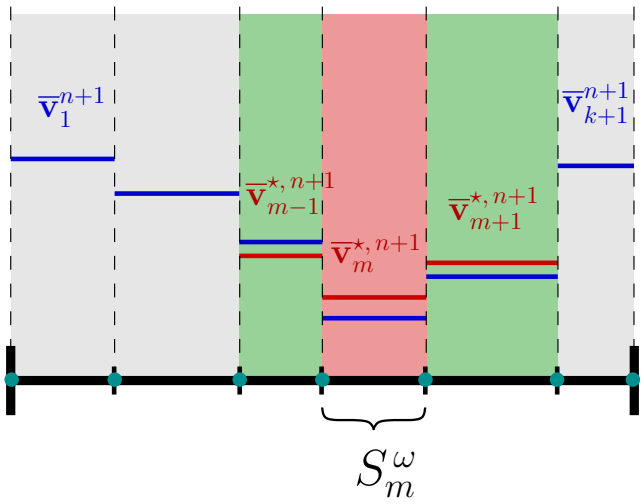

Figure 3.3: Sketch of sub-mean-values before and after correction

3.5 Subcell low-order corrected FV fluxes

In this section, we define the *corrected FV fluxes* $\mathcal{F}^{l/r}_{m\pm\frac{1}{2}}$. Such corrected fluxes are designed in order to: (i) ensure the desired robustness properties, in particular we aim at preserving the set of admissible states (3.20), see § 3.7 for the details, (ii) obtain a global discrete formulation which is well-balanced.

$$\Theta = \left\{(H,q) \in \mathbb{R}^2;\ H \geq 0\right\}. \tag{3.20}$$

To achieve this, we adapt the ideas introduced in [116, 62] to the framework of the current FV-Subcell method. For any $\omega_i \in \mathscr{T}_h$ and any marked subcell $S_m \in \mathscr{T}_{\omega_i}$, let define the sub-mesh reconstructed interface values for the topography:

$$\overline{b}_{m+\frac{1}{2}} := \max(\overline{b}_m, \overline{b}_{m+1}) \quad \text{and} \quad \overline{b}_{m-\frac{1}{2}} := \max(\overline{b}_{m-1}, \overline{b}_m),$$

and the additional subcell's interfaces (considering S_m) topography values:

$$\overline{b}^{\pm}_m := \overline{b}_{m\pm\frac{1}{2}} - \max\left(0, \overline{b}_{m\pm\frac{1}{2}} - \overline{\eta}_m\right), \tag{3.21}$$

$$\overline{b}^{-}_{m+1} := \overline{b}_{m+\frac{1}{2}} - \max\left(0, \overline{b}_{m+\frac{1}{2}} - \overline{\eta}_m\right), \qquad \overline{b}^{+}_{m-1} := \overline{b}_{m-\frac{1}{2}} - \max\left(0, \overline{b}_{m-\frac{1}{2}} - \overline{\eta}_m\right). \tag{3.22}$$

We introduce subcell's interfaces reconstructions for the water-height as follows:

$$\overline{H}^{\pm}_m := \max\left(0, \overline{\eta}_m - \overline{b}_{m\pm\frac{1}{2}}\right),$$

and for the surface elevation and discharge:

$$\overline{\eta}^{\pm}_m := \overline{H}^{\pm}_m + \overline{b}^{\pm}_m, \qquad \overline{q}^{\pm}_m := \overline{H}^{\pm}_m \frac{\overline{q}_m}{\overline{H}_m}, \tag{3.23}$$

leading to the new subcell's interfaces values:

$$\overline{\mathbf{v}}_m^{\pm} := (\overline{\eta}_m^{\pm}, \overline{q}_m^{\pm}).$$

Using these reconstructed values, we introduce some new FV numerical fluxes on subcell's S_m left and right interfaces, denoted by $\mathcal{F}^r_{m-\frac{1}{2}}$ and $\mathcal{F}^l_{m+\frac{1}{2}}$, as follows:

$$\mathcal{F}^l_{m+\frac{1}{2}} := \mathcal{F}\left(\overline{\mathbf{v}}_m^+, \overline{\mathbf{v}}_{m+1}^-, \overline{b}_m^+\right) + \begin{pmatrix} 0 \\ g\overline{\eta}_m^+\left(\overline{b}_m^+ - b_{\widetilde{x}_{m+\frac{1}{2}}}\right) \end{pmatrix}, \tag{3.24}$$

$$\mathcal{F}^r_{m-\frac{1}{2}} := \mathcal{F}\left(\overline{\mathbf{v}}_{m-1}^+, \overline{\mathbf{v}}_m^-, \overline{b}_m^-\right) + \begin{pmatrix} 0 \\ g\overline{\eta}_m^-\left(\overline{b}_m^- - b_{\widetilde{x}_{m-\frac{1}{2}}}\right) \end{pmatrix}, \tag{3.25}$$

where $b_{\widetilde{x}_{m\pm\frac{1}{2}}}$ are respectively the interpolated polynomial values of b_h at $\widetilde{x}_{m+\frac{1}{2}}$ and $\widetilde{x}_{m-\frac{1}{2}}$.

Remark 29. To compute the velocity in the vicinity of dry areas, we classically set a numerical threshold $\epsilon = 10^{-8}$ to numerically define what "a dry cell" is and set the velocity to 0 if $H < \epsilon$.

Remark 30. Considering the initialization choice mentioned in Remark 24 for the water elevation and the topography, we can ensure the well-balanced property for the *uncorrected* DG scheme simply by means of:

$$\mathcal{F}_{i+1/2} = \mathcal{F}\left(\mathbf{v}^-_{i+1/2}, \mathbf{v}^+_{i+1/2}, b_{i+1/2}\right) \quad \text{and} \quad \mathcal{F}_{i-1/2} = \mathcal{F}\left(\mathbf{v}^-_{i-1/2}, \mathbf{v}^+_{i-1/2}, b_{i-1/2}\right),$$

the standard DG numerical fluxes, where $\mathbf{v}^{\pm}_{i\pm1/2}$ are nothing but the interpolated interface values of $\mathbf{v}_h^{\omega_i}$. If one would like to use discontinuous topographies, one can then refer to the modified numerical flux strategy for DG schemes (see [62, 116]), as we did in the previous chapter §2.1. In the remainder, this stategy is used only for the first-order FV schemes.

3.6 Flowchart

We summarize the proposed new *a posteriori* LSC method of DG schemes through the following flowchart:

1. starting from an admissible piecewise polynomial approximate solution $\mathbf{v}_h^n \in (\mathbb{P}^k(\mathscr{T}_h))^2$, compute the candidate solution $\mathbf{v}_h^{n+1} \in (\mathbb{P}^k(\mathscr{T}_h))^2$ using the *uncorrected* DG scheme (3.3),

2. for any mesh element $\omega \in \mathscr{T}_h$, compute the candidate associated sub-mean-values:

$$\mathbb{P}^0(\mathscr{T}_\omega) \ni \overline{\mathbf{v}}_\omega^{n+1} = \pi_{\mathscr{T}_\omega}(\mathbf{v}_h^{\omega,n+1}),$$

3. for any mesh element $\omega \in \mathscr{T}_h$, for any subcell $S_m \in \mathscr{T}_\omega$, check admissibility of the associated sub-mean-values $\overline{\mathbf{v}}_m^{n+1}$, and identify accordingly the sub-partition $\mathscr{T}_\omega = \mathscr{T}_\omega^{\mathrm{f}} \cup \mathscr{T}_\omega^{\mathrm{u}}$, where $\mathscr{T}_\omega^{\mathrm{f}}$ and $\mathscr{T}_\omega^{\mathrm{u}}$ respectively refer to the set of flagged (non-admissible) subcells and the set of non-flagged (admissible) subcells (note that the sub-mean-values $\overline{\mathbf{v}}_m^{n+1}$ may be obtained either from Step 2. (without correction) or from Step 4. (b) (after correction)),

4. if, for all $\omega \in \mathscr{T}_h$, the identity $\mathscr{T}_\omega^{\mathrm{u}} = \mathscr{T}_\omega$ holds, then for all $\omega \in \mathscr{T}_h$, $\mathbf{v}_h^{\omega,n+1} = \pi_{\mathscr{T}_\omega}^{-1}(\overline{\mathbf{v}}_\omega^{n+1})$ is admissible, no additional correction is required and we can go further in time: go to Step 1, starting from $\mathbf{v}_h^{n+1}$ instead of $\mathbf{v}_h^n$.

   Otherwise:

   (a) for all $\omega \in \mathscr{T}_h$ such that $\mathscr{T}_\omega^{\mathrm{f}} \neq \emptyset$, and all $S_m \in \mathscr{T}_\omega^{\mathrm{f}}$, substitute the corresponding *reconstructed fluxes* with some *corrected fluxes* defined in (3.24)-(3.25), as follows:

$$\begin{cases} \widetilde{\mathbf{F}}^l_{m+\frac{1}{2}} \longleftarrow \mathcal{F}^l_{m+\frac{1}{2}} \quad \text{and} \quad \widetilde{\mathbf{F}}^r_{m+\frac{1}{2}} \longleftarrow \mathcal{F}^r_{m+\frac{1}{2}} \quad \text{if either } S_m \text{ or } S_{m+1} \text{ is marked,} \\ \widetilde{\mathbf{F}}^l_{m+\frac{1}{2}} \longleftarrow \widehat{\mathbf{F}}_{m+\frac{1}{2}} \quad \text{and} \quad \widetilde{\mathbf{F}}^r_{m+\frac{1}{2}} \longleftarrow \widehat{\mathbf{F}}_{m+\frac{1}{2}} \quad \text{otherwise,} \end{cases}$$

   (b) for all $\omega \in \mathscr{T}_h$ such that $\mathscr{T}_\omega^{\mathrm{f}} \neq \emptyset$, and all $S_m \in \mathscr{T}_\omega^{\mathrm{f}}$, compute new sub-mean-values for the marked subcells and their first neighboring subcells, respectively denoted $\overline{\mathbf{v}}_m^{*\,n+1}$, $\overline{\mathbf{v}}_{m-1}^{*\,n+1}$, $\overline{\mathbf{v}}_{m+1}^{*\,n+1}$, by means of a corrected FV-Subcell scheme as:

$$\overline{\mathbf{v}}_p^{\star,n+1} = \overline{\mathbf{v}}_p^n - \frac{\Delta t^n}{|S_p|}\left(\widetilde{\mathbf{F}}^l_{p+\frac{1}{2}} - \widetilde{\mathbf{F}}^r_{p-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_p, \tag{3.26}$$

   for $p \in [\![m-1,\, m+1]\!]$. This subcell corrected method (3.26) falls in one of the previously introduced cases (3.17), (3.18) or (3.19),

   (c) for all $\omega \in \mathscr{T}_h$ such that at least one subcell has been corrected, gather the uncorrected sub-mean-values $\overline{\mathbf{v}}_m^{n+1}$ and corrected sub-mean-values $\overline{\mathbf{v}}_m^{*\,n+1}$ in a new element of $\mathbb{P}^0(\mathscr{T}_\omega)$, which is still denoted $\overline{\mathbf{v}}_\omega^{n+1}$ for the sake of simplicity,

   (d) go to step 3,

Step 3 of the flowchart is detailed in the next section.

3.7 Admissibility criteria

A large number of sensors or detectors have been introduced in the literature, to identify the marked subcells, where some kind of stabilization is required to avoid a loss of robustness. Following [164], we use two admissibility criteria: one for the *Physical Admissibility Detection* (PAD), another addressing the occurrence of spurious oscillations, namely the *Subcell Numerical Admissibility Detection* (SubNAD). This last criterion is supplemented with a relaxation procedure to exclude the smooth extrema from the troubled cells.

Physical Admissibility Detection (PAD)

Here, we define a sensor function that :

- Check if the sub-mean-values $\overline{\mathbf{v}}_m^{n+1}$ belongs to Θ, see (3.20).
- Check if there is any NaN values.

Those are the minimum requirements to enforce the code robustness.

Subcell Numerical Admissibility Detection (SubNAD)

In order to tackle the issue of spurious oscillations near discontinuities, we enforce a local *Discrete Maximum Principle* (DMP), at the subcell level, on the surface elevation as follows:

- Check if, for $m = 1, \ldots, k+1$, the following inequalities hold:

$$\min\left(\overline{\eta}_{m-1}^{n}, \overline{\eta}_{m}^{n}, \overline{\eta}_{m+1}^{n}\right) \leq \overline{\eta}_{m}^{n+1} \leq \max\left(\overline{\eta}_{m-1}^{n}, \overline{\eta}_{m}^{n}, \overline{\eta}_{m+1}^{n}\right).$$

The SubNAD criterion relies on a DMP based on subcell mean-values, and not the whole polynomial set of values. Furthermore, as the neighboring subcells set used in the SubNAD is reduced to the first left and right subcells, and not all the subcells contained in the DG cell as well as in the the left and right first neighboring DG cells, see [164], one has to introduce a relaxation mechanism in order to preserve the scheme accuracy in the vicinity of smooth extrema.

Detection of smooth extrema.

In the relaxation procedure proposed in [164], it is assumed that the numerical solution exhibits a smooth extremon if at least the following linearized version of the surface elevation spatial derivative:

$$(\partial_x \eta)_{\omega_i}^{\text{lin}}(x) = \overline{\partial_x \eta_h^{\omega_i,n+1}} + (x - x_i)\,\overline{\partial_{xx} \eta_h^{\omega_i,n+1}},$$

has a monotonous profile, where $\overline{\partial_x \eta_h^{\omega_i,n+1}}$ and $\overline{\partial_{xx} \eta_h^{\omega_i,n+1}}$ are respectively the mean-values of $(\partial_x \eta_h)_{|\omega_i}$ and $(\partial_{xx} \eta_h)_{|\omega_i}$ on mesh element ω_i. In practice, the DMP relaxation used here works as a vertex-based limiter on $(\partial_x \eta)_{\omega_i}^{\text{lin}}$. Hence, we set $\partial_x \eta_L := \overline{\partial_x \eta_h^{\omega_i,n+1}} - \frac{h_{\omega_i}}{2}\,\overline{\partial_{xx} \eta_h^{\omega_i,n+1}}$ to be the left boundary value of $(\partial_x \eta)_{\omega_i}^{\text{lin}}$ on cell ω_i, as well as $\partial_x \eta_{\min\setminus\max}^{L} = \min\setminus\max\left(\overline{\partial_x \eta_h^{\omega_{i-1},n+1}}, \overline{\partial_x \eta_h^{\omega_i,n+1}}\right)$ respectively the minimum and maximum values of the mean derivative around $x_{i-\frac{1}{2}}$. We then define the left detection factor α_L as follows:

$$\alpha_L = \begin{cases} \min\left(1, \dfrac{\partial_x \eta_{\max}^{L} - \overline{\partial_x \eta_h^{\omega_i,n+1}}}{\partial_x \eta_L - \overline{\partial_x \eta_h^{\omega_i,n+1}}}\right), & \text{if } \partial_x \eta_L > \overline{\partial_x \eta_h^{\omega_i,n+1}}, \\ 1, & \text{if } \partial_x \eta_L = \overline{\partial_x \eta_h^{\omega_i,n+1}}, \\ \min\left(1, \dfrac{\partial_x \eta_{\min}^{L} - \overline{\partial_x \eta_h^{\omega_i,n+1}}}{\partial_x \eta_L - \overline{\partial_x \eta_h^{\omega_i,n+1}}}\right), & \text{if } \partial_x \eta_L < \overline{\partial_x \eta_h^{\omega_i,n+1}}. \end{cases}$$

Introducing the symmetric values $\partial_x \eta_{\min\setminus\max}^{R} = \min\setminus\max\left(\overline{\partial_x \eta_h^{\omega_i,n+1}}, \overline{\partial_x \eta_h^{\omega_{i+1},n+1}}\right)$ and $\partial_x \eta_R := \overline{\partial_x \eta_h^{\omega_i,n+1}} + \frac{h_{\omega_i}}{2}\,\overline{\partial_{xx} \eta_h^{\omega_i,n+1}}$, the right detection factor α_R is obtained in a similar manner. Finally, introducing $\alpha := \min(\alpha_L, \alpha_R)$, we consider that the numerical solution presents a smooth profile on the cell ω_i if $\alpha = 1$. In this particular case, the SubNAD criterion is relaxed, allowing the high-order accuracy preservation of smooth extrema.

Remark 31. One can apply the subcell numerical admissibility detection SubNAD and relaxation method detailed above on the Riemann invariants $I^{\pm} = u \pm 2\sqrt{gH}$ instead of the surface elevation η. Actually, the simplest choice, that also leads to the best results, is to perform the detection on the surface elevation η. The detection applied to the Riemann invariants produces a more diffused solution.

3.8 Well-balancing property

This section is now dedicated to the demonstration of the well-balanced property of this *a posteriori* LSC of DG schemes.

Remark 32. Let us note that under the motionless steady-state assumption $\eta_h = \eta^e$ and $q_h = 0$, the following relation holds:

$$\partial_x \mathbf{F}(\mathbf{v}_h^\omega, b_h^\omega) = \mathbf{B}(\mathbf{v}_h^\omega, \partial_x b_h^\omega), \quad \forall \omega \in \mathscr{T}_h.$$

Moreover, as we have

$$\mathbf{F}(\mathbf{v}_h, b_h) = \begin{pmatrix} 0 \\ \frac{1}{2} g \eta_h^2 - g \eta_h b_h \end{pmatrix},$$

we emphasize that, under the steady-state hypothesis, $\mathbf{F}(\mathbf{v}_h, b_h)$ belongs to $\mathbb{P}^k(\mathscr{T}_h)^2$ and not only to $\mathbb{P}^{2k}(\mathscr{T}_h)^2$, since $\eta_h = \eta^e$. Therefore, at steady state,

$$\mathbf{F}_h := p^k_{\mathscr{T}_h}(\mathbf{F}(\mathbf{v}_h, b_h)) = \mathbf{F}(\mathbf{v}_h, b_h).$$

As for $\mathbf{B}$, under the same assumptions we have:

$$\mathbf{B}(\mathbf{v}_h, \partial_x b_h) = \begin{pmatrix} 0 \\ -g \eta^e \partial_x b_h \end{pmatrix} \in \mathbb{P}^k(\mathscr{T}_h) \times \mathbb{P}^{k-1}(\mathscr{T}_h) \subset \mathbb{P}^k(\mathscr{T}_h)^2,$$

thus,

$$\mathbf{B}_h := p^k_{\mathscr{T}_h}(\mathbf{B}(\mathbf{v}_h, \partial_x b_h)) = \mathbf{B}(\mathbf{v}_h, \partial_x b_h).$$

We have then the following result:

Proposition 1. The discrete formulation obtained by gathering (3.3) and the local corrected FV schemes on subcells (3.17), (3.18) and (3.19), together with a first-order Euler time-marching algorithm, preserves the motionless steady states, providing that the integrals of (3.4) are exactly computed for the motionless steady states. Specifically, for all $n \geq 0$ and all $\eta^e \in \mathbb{R}$,

$$(\eta_h^n = \eta^e \text{ and } q_h^n = 0) \implies \left(\eta_h^{n+1} = \eta^e \text{ and } q_h^{n+1} = 0\right).$$

Proof. We consider the scheme (3.12) on uncorrected subcells, and schemes (3.17), (3.18) and (3.19) on corrected subcells. We have to distinguish three different situations: (i) uncorrected subcell, (ii) neighbor of a marked subcell, (iii) marked subcell. We show in what follows that in all those situations, the corrected DG scheme preserves the motionless steady-states at the subcell level:

$$\forall \omega \in \mathscr{T}_h, \quad \forall m \in [1, \ldots, k+1], \quad \overline{\eta}_m^{\omega,n} = \eta^e, \quad \overline{q}_m^{\omega,n} = 0 \quad \implies \quad \overline{\eta}_m^{\omega,n+1} = \eta^e, \quad \overline{q}_m^{\omega,n+1} = 0.$$

1. **Uncorrected subcell**: S_{m-1}, S_m and S_{m+1} are not marked.
   In this case, we consider the *uncorrected* DG scheme or this is equivalent FV-like scheme with reconstructed fluxes:

$$\overline{\mathbf{v}}_m^{n+1} = \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|} \left(\widehat{\mathbf{F}}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}} \right) + \Delta t^n \overline{\mathbf{B}}_m, \tag{3.27}$$

   with $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ defined in (3.14). We have, at steady state:

$$\eta^+_{i\pm\frac{1}{2}} = \eta^-_{i\pm\frac{1}{2}} = \eta^e, \qquad q^+_{i\pm\frac{1}{2}} = q^-_{i\pm\frac{1}{2}} = 0, \qquad \text{and} \qquad b^+_{i\pm\frac{1}{2}} = b^-_{i\pm\frac{1}{2}},$$

and therefore

$$\mathcal{F}_{i\pm\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(x_{i\pm\frac{1}{2}}), b_h(x_{i\pm\frac{1}{2}})\right),$$

so that

$$\widehat{\mathbf{F}}_{m\pm\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m\pm\frac{1}{2}}), b_h(\widetilde{x}_{m\pm\frac{1}{2}})\right).$$

As a consequence, we have

$$\widehat{\mathbf{F}}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}} = \int_{S_m} \partial_x \mathbf{F}(\mathbf{v}_h, b_h)\,\mathrm{d}x, \tag{3.28}$$

and injecting (3.28) into (3.27) gives

$$\begin{aligned}\overline{\mathbf{v}}_m^{n+1} &= \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|}\int_{S_m} \partial_x \mathbf{F}(\mathbf{v}_h, b_h)\mathrm{d}x + \frac{\Delta t^n}{|S_m|}\int_{S_m} \mathbf{B}(\mathbf{v}_h, \partial_x b_h)\mathrm{d}x \\ &= \overline{\mathbf{v}}_m^n.\end{aligned}$$

2. **Neighbor of a troubled subcell**: S_m, S_{m-1} are not marked and S_{m+1} is marked. The corresponding scheme, in this case, is the following:

$$\overline{\mathbf{v}}_m^{\star,n+1} = \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|}\left(\mathcal{F}^l_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_m,$$

with $\mathcal{F}^l_{m+\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ respectively defined in (3.24) and (3.14). At steady state, the reconstruction (3.23) yields $\overline{\eta}_m^+ = \overline{\eta}_{m+1}^- = \eta^c$ (see Fig. 3.4 and 3.5) and $\overline{q}_m^+ = \overline{q}_{m+1}^- = 0$. It leads to:

$$\mathcal{F}\left(\overline{\mathbf{v}}_m^+, \overline{\mathbf{v}}_{m+1}^-, \overline{b}_m^+\right) = \frac{1}{2}\left[\mathbf{F}(\overline{\mathbf{v}}_m^+, \overline{b}_m^+) + \mathbf{F}(\overline{\mathbf{v}}_{m+1}^-, \overline{b}_m^+)\right] = \frac{1}{2}\begin{pmatrix} 0 \\ g\left((\eta^e)^2 - 2\eta^e \overline{b}_m^+\right)\end{pmatrix},$$

and then to:

$$\mathcal{F}^l_{m+\frac{1}{2}} = \frac{1}{2}\begin{pmatrix} 0 \\ g\left((\eta^e)^2 - 2\eta^e b_{\widetilde{x}_{m+\frac{1}{2}}}\right)\end{pmatrix} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m+\frac{1}{2}}), b_h(\widetilde{x}_{m+\frac{1}{2}})\right). \tag{3.29}$$

Moreover, as in the previous case:

$$\widehat{\mathbf{F}}_{m-\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m-\frac{1}{2}}), b_h(\widetilde{x}_{m-\frac{1}{2}})\right). \tag{3.30}$$

Gathering (3.29) and (3.30), we then have

$$\mathcal{F}^l_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}} = \int_{S_m} \partial_x \mathbf{F}(\mathbf{v}_h, b_h)\,\mathrm{d}x,$$

so that

$$\overline{\mathbf{v}}_m^{\star,n+1} = \overline{\mathbf{v}}_m^n.$$

3. **Corrected subcell**: S_m is marked.
   In this case, the corresponding scheme reduces to (3.17). Following the lines of the previous cases, we have:

$$\mathcal{F}^l_{m+\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m+\frac{1}{2}}), b_h(\widetilde{x}_{m+\frac{1}{2}})\right), \qquad \mathcal{F}^r_{m-\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m-\frac{1}{2}}), b_h(\widetilde{x}_{m-\frac{1}{2}})\right),$$

   and therefore

$$\mathcal{F}^l_{m+\frac{1}{2}} - \mathcal{F}^r_{m-\frac{1}{2}} = \int_{S_m} \partial_x \mathbf{F}(\mathbf{v}_h, b_h)\,\mathrm{d}x,$$

   so that

$$\overline{\mathbf{v}}^{\star,\,n+1}_m = \overline{\mathbf{v}}^n_m.$$

□

We have just shown that schemes (3.12)-(3.17)-(3.18)-(3.19) do ensure the well-balanced property in wet subcells for all contexts, wet/wet and wet/dry. As for dry subcells, we can also simply show well-balancing property. Considering a dry zone at time level n, under the assumptions $\eta^n = b$ and $q^n = 0$, one can easily show that the dry zone stays a dry zone a the next time level $n+1$, *i.e.* $\eta^{n+1} = b$ and $q^{n+1} = 0$, by following a very similar procedure as in the previous proofs.

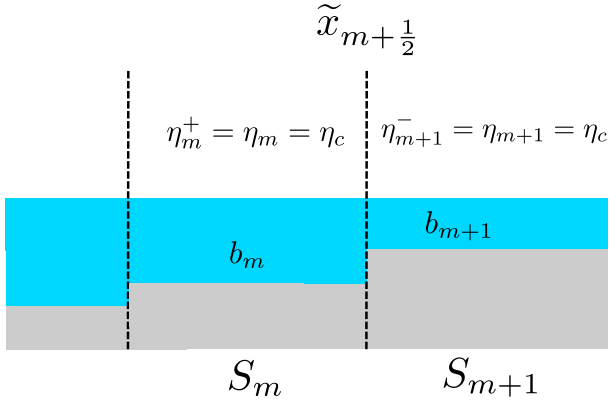

Figure 3.4: The sub-mesh reconstructed interface values for the water elevation: considering a wet/wet context

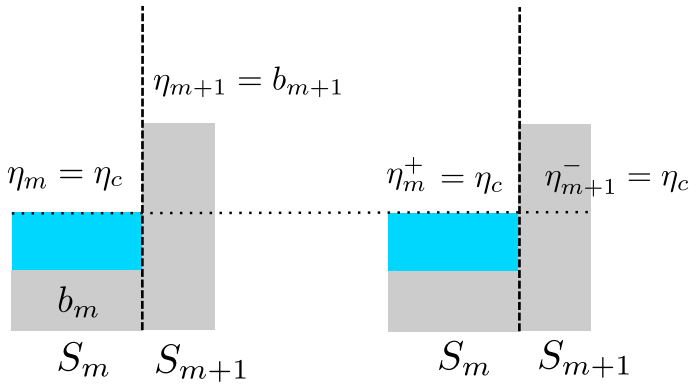

Figure 3.5: The sub-mesh reconstructed interface values for the water elevation: considering a wet/dry context

Remark 33. As we mentioned in Remark 24, the topography approximation b_h is a globally continuous function over the mesh. Indeed, the piecewise constant bathymetry values presented in Fig. 3.4 and 3.5 do not represent a discontinuous bathymetry, but a piecewise constant representation of b_h on the sub-grid.

Remark 34. We recall that the use of a non-smooth topography parameterization may be allowed, while still ensuring the well-balancing property, at the price of considering interface reconstructions also for the cells interfaces, for the DG scheme, in the spirit of [62, 116], see also §2.1.

3.9 Preservation of the water-height positivity

After computing the candidate solution $\mathbf{v}_h^{n+1}$ through the *uncorrected* DG scheme (3.3), if we detect a negative sub-mean-value on an arbitrary subcell, this subcell is then marked and a new (corrected) sub-mean-value is evaluated by means of the first-order FV-Subcell scheme (3.17). As a consequence, scheme (3.17) with reconstruction (3.23) should preserve positivity.

Proposition 2. Under the CFL condition (3.16), if $\forall \omega \in \mathscr{T}_h$, $\forall S_m \in \mathscr{T}_\omega$, $\overline{v}_m^{\omega,n} \in \Theta$, then $\forall \omega \in \mathscr{T}_h$, $\forall S_m \in \mathscr{T}_\omega$, $\overline{v}_m^{\omega,n+1} \in \Theta$.

Proof. As the positivity-preserving property of our *a posteriori* LSC of DG schemes relies on the positivity of the first-order FV scheme used as the correction method, let us prove that if $\overline{H}_m^n$ and $\overline{H}_{m\pm1}^n$ are non-negative, then scheme (3.17) does produce a water-height $\overline{H}_m^{n+1}$ also non-negative. Let us first recall the equation corresponding to the time evolution of the discrete surface elevation:

$$\overline{\eta}_m^{n+1} = \overline{\eta}_m^n - \frac{\Delta t^n}{|S_m|}\left(\mathcal{F}_1\left(\overline{\mathbf{v}}_m^{n,+}, \overline{\mathbf{v}}_{m+1}^{n,-}, \overline{b}_m^+\right) - \mathcal{F}_1\left(\overline{\mathbf{v}}_{m-1}^{n,+}, \overline{\mathbf{v}}_m^{n,-}, \overline{b}_m^-\right)\right), \tag{3.31}$$

where $\mathcal{F}_1$ represents the first component of the numerical flux $\mathcal{F}$ and $\overline{\mathbf{v}}_m^{n,\pm}$, $\overline{b}_m^{n,\pm}$ are defined in (3.23) and (3.21)-(3.22). For sake of simplicity, we drop in the following the superscript n. Equation (3.31)

rewrites explicitly as:

$$\begin{aligned}\overline{\eta}_m^{n+1} = \overline{\eta}_m \quad &- \quad \frac{\Delta t^n}{2\,|S_m|}\left(\overline{H}_m^+\frac{\overline{q}_m}{\overline{H}_m} + \overline{H}_{m+1}^-\frac{\overline{q}_{m+1}}{\overline{H}_{m+1}} - \sigma\left(\overline{\eta}_{m+1}^- - \overline{\eta}_m^+\right)\right)\\ &- \quad \frac{\Delta t^n}{2\,|S_m|}\left(\overline{H}_m^-\frac{\overline{q}_m}{\overline{H}_m} + \overline{H}_{m-1}^+\frac{\overline{q}_{m-1}}{\overline{H}_{m-1}} - \sigma\left(\overline{\eta}_m^- - \overline{\eta}_{m-1}^+\right)\right).\end{aligned} \tag{3.32}$$

Noticing that $\overline{\eta}_{m+1}^- - \overline{\eta}_m^+ = \overline{H}_{m+1}^- - \overline{H}_m^+$ as well as $\overline{\eta}_m^- - \overline{\eta}_{m-1}^+ = \overline{H}_m^- - \overline{H}_{m-1}^+$, and subtracting $\overline{b}_m$ on both sides of this last expression, equation (3.32) can be reformulated as:

$$\begin{aligned}\overline{H}_m^{n+1} \quad &= \quad \left[1 - \frac{1}{2}\lambda\left(\sigma - \overline{u}_m\right)\frac{\overline{H}_m^-}{\overline{H}_m} - \frac{1}{2}\lambda\left(\sigma + \overline{u}_m\right)\frac{\overline{H}_m^+}{\overline{H}_m}\right]\overline{H}_m\\ &+ \quad \left[\frac{1}{2}\lambda\left(\sigma + \overline{u}_{m-1}\right)\frac{\overline{H}_{m-1}^+}{\overline{H}_{m-1}}\right]\overline{H}_{m-1} + \left[\frac{1}{2}\lambda\left(\sigma - \overline{u}_{m+1}\right)\frac{\overline{H}_{m+1}^-}{\overline{H}_{m+1}}\right]\overline{H}_{m+1},\end{aligned} \tag{3.33}$$

with $\overline{u}_m = \dfrac{\overline{q}_m}{\overline{H}_m}$ and $\lambda = \dfrac{\Delta t^n}{|S_m|}$. Therefore, $\overline{H}_m^{n+1}$ reads as a convex combination of $\overline{H}_{m-1}, \overline{H}_m$ and $\overline{H}_{m+1}$. Furthermore, since by construction $0 \leq \overline{H}_p^{\pm} \leq \overline{H}_p$, $\forall p \in [\![1,\, k+1]\!]$ and by respect of the CFL condition (3.16), $\lambda\alpha \leq 1$, and then all the coefficients involved in the convex combination (3.33) are non-negative. It follows that $\overline{H}_m^{n+1} \geq 0$. □

Remark 35. The proposed positivity criteria are based on subcell values, and indeed, our goal is to show that our scheme preserves the positivity at the subcell level. Hence, the chosen strategy, as it is, does not ensure the pointwise positivity of H at some specific nodes: such a property is not needed. If one requires such pointwise positivity, for some specific reasons, we emphasize that an additional "positivity limiter", as the one provided in [179] for instance, can be combined with our approach, to ensure the positivity of the polynomial solution at any chosen points. We emphasize that enforcing the positivity of H at the subcell level, we are able to compute the approximated eigenvalues $u \pm \sqrt{gH}$ appearing in the CFL condition (3.16).

3.10 Numerical validations

In this numerical results section, we make use of several widely addressed and challenging test-cases to demonstrate the performance and robustness of DG schemes provided the *a posteriori* local subcell correction presented. In all following test-cases, if not stated differently, sub-mean-values are displayed. It allows us to fully illustrate the very precise subcell resolution of our scheme.

3.10.1 A smooth sinusoidal solution

This first test-case aims at numerically evaluating the rates of convergence of the present *a posteriori* LSC of DG schemes. To do so, following the methodology introduced in [167] in the context of compressible gas dynamics, we make use of a smooth solution of the NSW equations. Details on the design of such solution can be found in Appendix B. To do so, we initialize the problem with the following initial data:

$$\eta_0 = \frac{u_0^2}{4\,g} \qquad \text{and} \qquad q_0 = \frac{u_0^3}{4\,g},$$

with the initial constant velocity perturbed by a sinusoidal signal:

$$u_0(x) = 1 + 0.1\sin(2\pi x).$$

We run this test-case with the fourth-order scheme on 60 cells on the domain $[0, 1]$, up to the stopping time t = 0.3, with periodic boundary conditions. The result is plotted in Fig. 3.6. In Table 3.1, we gather the global L^2-errors as well as the rates of convergence for different order of approximation, computed on the surface elevation at $t = 0.3s$. As expected, the computed rates of convergence scale as $O(k + 1)$. A similar behavior can be observed for the horizontal discharge q.

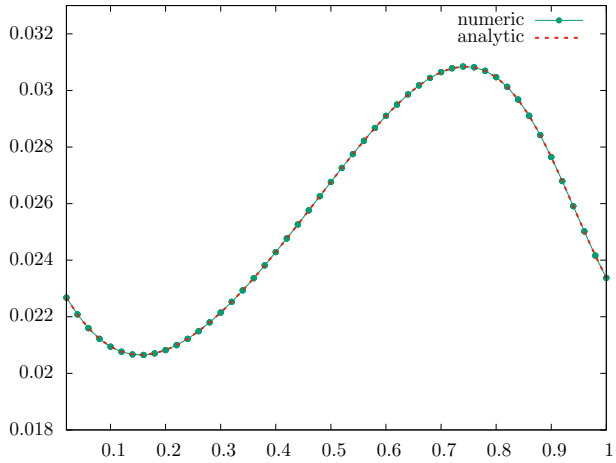

Figure 3.6: Test 4 - A smooth sinusoidal analytical solution for the NSW equations - Free surface elevation computed at $t = 0.3\,s$ with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 60$.

k	1		2		3	
h	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$
$\frac{1}{15}$	1.093E-5	2.05	1.91E-7	2.99	9.390E-7	4.32
$\frac{1}{30}$	2.62E-6	2.02	2.40E-8	3.004	4.70E-8	4.27
$\frac{1}{60}$	6.43E-7	2.01	2.99E-9	3.003	2.43E-9	3.89
$\frac{1}{120}$	1.59E-7	-	3.73E-10	-	1.64E-10	-

Table 3.1: Test 4 - A smooth sinusoidal analytical solution for the NSW equations: L^2-errors between numerical and analytical solutions and convergence rates for η at time $t = 0.3s$.

3.10.2 A new analytical solution for the NSW equations

An other test-case that also aims to numerically evaluate the rates of convergence of the present *a posteriori* LSC of DG schemes is presented. To do so, we follow always the methodology introduced in [167] in the context of compressible gas dynamics, see Appendix B. This solution has the very interesting features to achieve any arbitrary regularity, *i.e.* $\mathbf{v}(., t) \in \mathcal{C}^{N_s}(\Omega)$, $\forall\, t < t_c(N_s)$ and any $N_s \in \mathbb{N}^*$, allowing the study of convergence up to any order of accuracy, while involving almost

vanishing water depth, together with a loss of regularity and the occurrence of discontinuous profiles for $t \geq t_c$.

We consider here the computational domain $\Omega = [-0.5, 2.5]$, and the particular case of $N_s = 3$. It follows that the critical time reads $t_c \approx 0.44\,s$, see B for more details. We initialize the problem with the following initial data:

$$\eta_0 = \frac{u_0^2}{4\,g} \qquad \text{and} \qquad q_0 = \frac{u_0^3}{4\,g},$$

with the following $\mathcal{C}^{N_s}$ smooth initial velocity

$$u_0(x) = \begin{cases} 1 & \text{if } x \leq 0, \\ e^{-x^{N_s+1}} & \text{elsewhere.} \end{cases}$$

While the *uncorrected* DG scheme (3.3) allows to compute the solution without any robustness issue for small enough values of time, nonphysical oscillations may be generated for larger values of time, leading to the activation of the *a posteriori* LSC method . A comparison between our fourth-order numerical solution computed on a mesh made of 60 cells, and the analytical solution at $t = 0.1\,s$ is shown on Fig. 3.7. One can see in Fig. 3.7 that only the cell mean-values are displayed. We can also

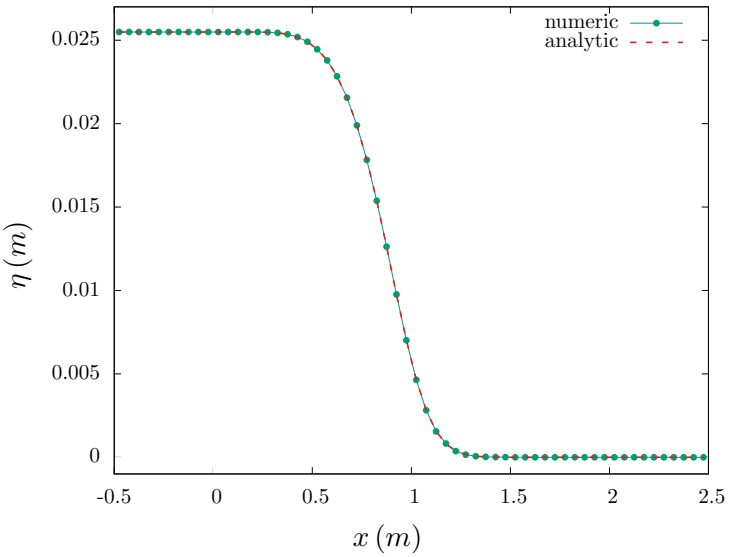

Figure 3.7: Test 5 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.1\,s$ with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 60$.

observe how the numerical scheme has very accurately captured to exact solution. In Table 3.2, we gather the global L^2-errors as well as the rates of convergence for different order of approximation, computed on the surface elevation at $t = 0.1s$. As expected, the computed rates of convergence scale as $O(k+1)$. A similar behavior can be observed for the horizontal discharge q.

k	1		2		3	
h	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$
$\frac{1}{15}$	5.91E-4	1.96	2.13E-5	3.19	3.20E-6	4.05
$\frac{1}{30}$	1.52E-4	2.02	2.33E-6	2.85	1.93E-7	4.18
$\frac{1}{60}$	3.73E-5	2.02	2.99E-7	2.95	1.06E-8	3.95
$\frac{1}{120}$	9.21E-6	-	4.18E-8	-	6.91E-10	-

Table 3.2: Test 5 - A new analytical solution for the NSW equations: L^2-errors between numerical and analytical solutions and convergence rates for η at time $t = 0.1s$

In a second time, we consider a larger final computational time $t > t_c$, so that a right-going discontinuity has developed from the initially regular profile, allowing to check the ability of the proposed *a posteriori* LSC method to stabilize the computation, namely to get rid of the spurious oscillations as well as enforcing the positivity of the water-height. We run the previous case until $t = 0.55$, with $k = 3$ and $n_{\text{el}} = 100$ mesh elements. Note that the standard DG method crashes in this case, since nonphysical undershoots would be rapidly amplified. In Fig. 3.8, a comparison between the *a posteriori* corrected DG solution and a reference solution obtained with a robust first-order FV method and $n_{\text{el}} = 10000$ mesh elements.

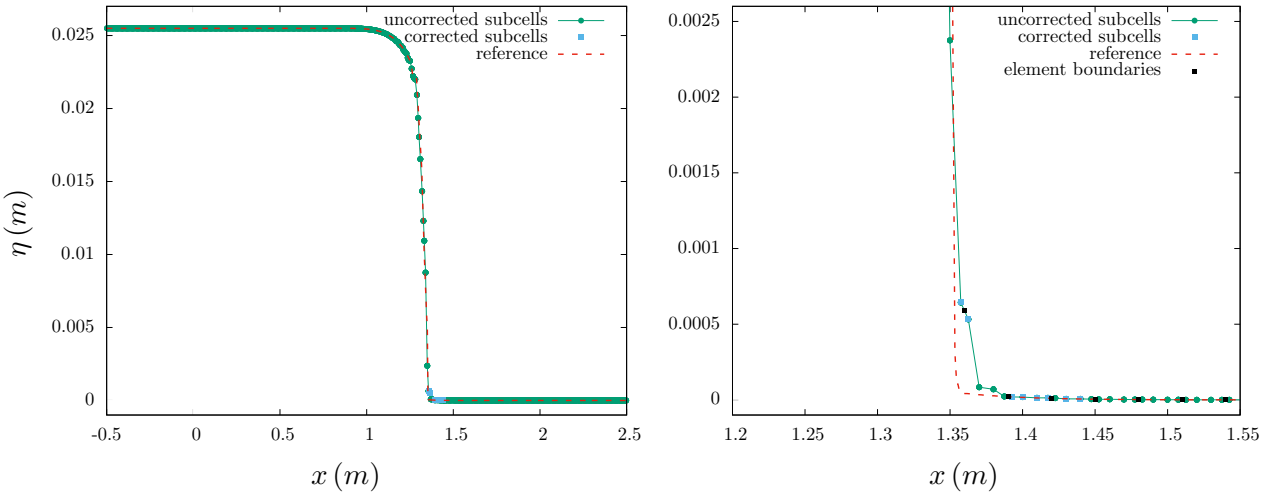

Figure 3.8: Test 5 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.55\,s$ with the *a posteriori* LSC method (left) for $k = 3$ and $n_{\text{el}} = 100$, with a zoom on the discontinuity and wet/dry interface (right).

This is a challenging computation for high-order methods since small values of water-height occur and thus small undershoots generally quickly lead to larger undershoots and possibly loss of positivity. In practice, the sensor starts to be activated when the strong gradient appears, slightly before the apparition of the discontinuity. A particular emphasize is put in Fig. 3.8 on the location of marked subcells, where uncorrected subcells are plotted by green dots while corrected subcells are plotted with blue squares. We observe that the particular combination of admissibility criteria introduced in §3.7 works quite well in practice, as the detection has been able to accurately track the moving front, and doing so removed the spurious oscillations without impacting smooth areas.

To conclude this test, we show on Fig. 3.9 the numerical results obtained with the *a posteriori* LSC method with a high-order polynomial approximation $k = 8$, along with a quite coarse mesh made of 20 elements, at time $t = 0.55\,s$. The use of such a coarse mesh permits to highlight the particularly interesting subcell resolution capabilities of our method, allowing to accurately locate the wet-dry interface inside a mesh element.

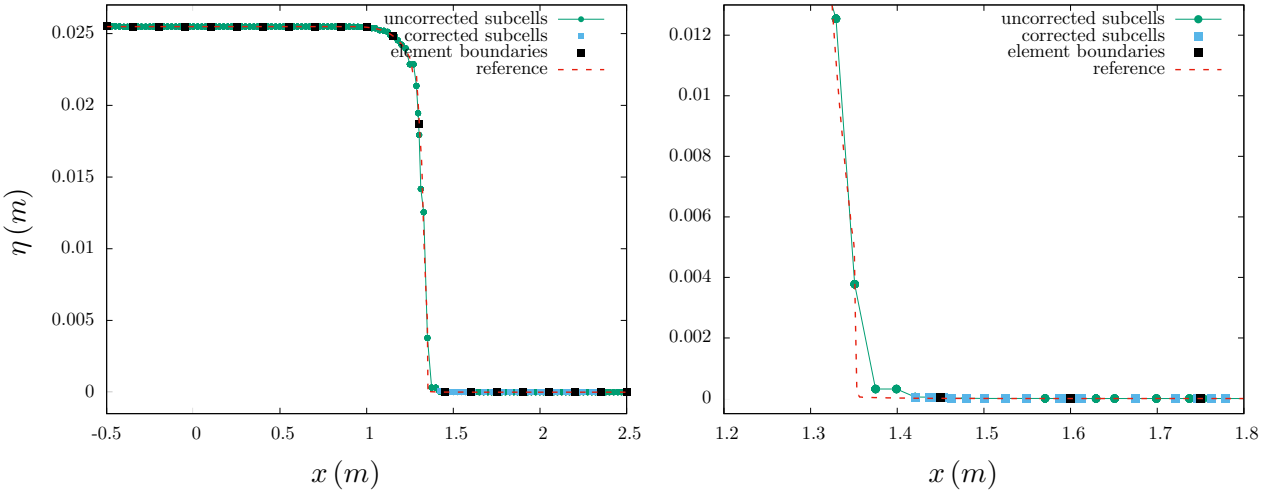

Figure 3.9: Test 5 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.55\,s$ with the *a posteriori* LSC method (left) for $k = 8$ and $n_{\text{el}} = 20$, with a zoom on the discontinuity and wet/dry interface (right).

3.10.3 Dam-break

In this second test-case, we focus on two dam-break problems over flat bottoms. The computational domain is set to $\Omega = [0, 1]$ and the first set of initial conditions is defined as follows:

$$\eta_0(x) = \begin{cases} 1 & \text{if} \quad x \leq 0.5, \\ 0.5 & \text{elsewhere,} \end{cases}, \qquad q_0 = 0, \qquad b = 0.$$

The final time is set to $t = 0.075\,s$. In Fig. 3.10, on a 50 cells mesh, fourth-order *uncorrected* DG solution is displayed on the left figure, while the *corrected* solution is plotted on the right one. This illustrates very clearly that even if the correction has been activated on in a very sharp area in the vicinity of the discontinuity, the solution has still been cleansed from its spurious oscillations. Now, we compare our *a posteriori* LSC method with the limitation process introduced in [179] (refereed to as PL/TVB method in what follows) and presented in (2.20)-(2.7), which combines the positivity-preserving limiter [184] with a standard TVB limiter [44]. Following [179], the constant M involved in the TVB limiter is set to $M = 0$. The results are plotted in Fig. 3.11 and Fig. 3.12.
In Fig. 3.11 and 3.12, one can observe that the present correction technique outperforms the positivity-preserving + TVB limiter, both in the rarefaction and shock resolution. Finally, to demonstrate how this *a posteriori* LSC method scales going to very high-orders of accuracy and very coarse meshes, we run the same test with $k = 9$ and a 10 mesh elements. The corresponding numerical result is shown on Fig. 3.13.

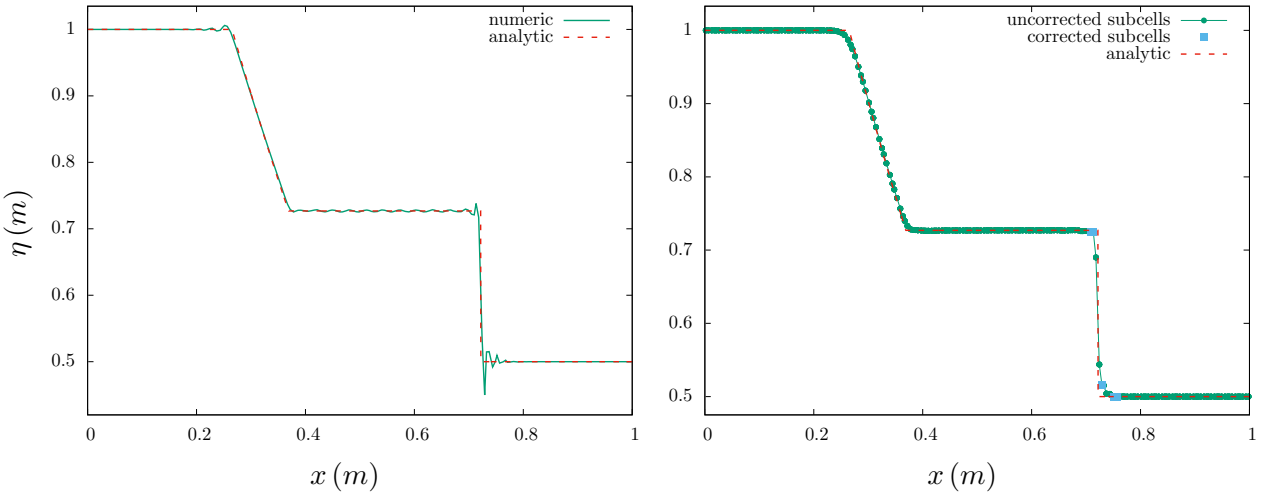

Figure 3.10: Test 6 - Dam break on a wet bottom- Free surface elevation computed at $t = 0.075\,s$ with the *uncorrected* DG method (left) and the *a posteriori* LSC method (right), with $k = 3$ and $n_{\text{el}} = 50$ mesh elements.

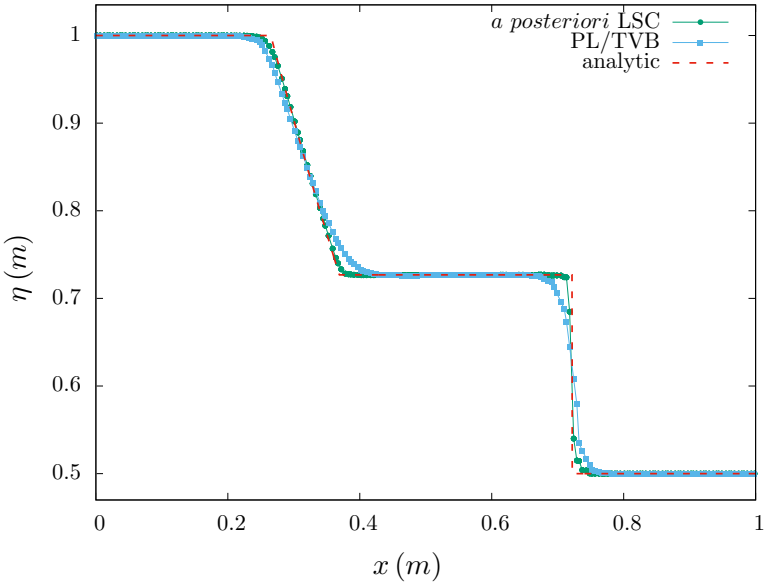

Figure 3.11: Test 6 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method and PL/TVB method for $k = 3$ and $n_{\text{el}} = 50$.

Fig. 3.13 illustrates the high capability of this *a posteriori* LSC method to retain the precise subcell resolution of DG schemes, allowing the use of very coarse meshes, along with being able to avoid the appearance of spurious oscillations or any unfortunate crash of the code. This figure also displays how the present correction affects the solution only at the subcell level, allowing the resolution of the shock in only one mesh element.

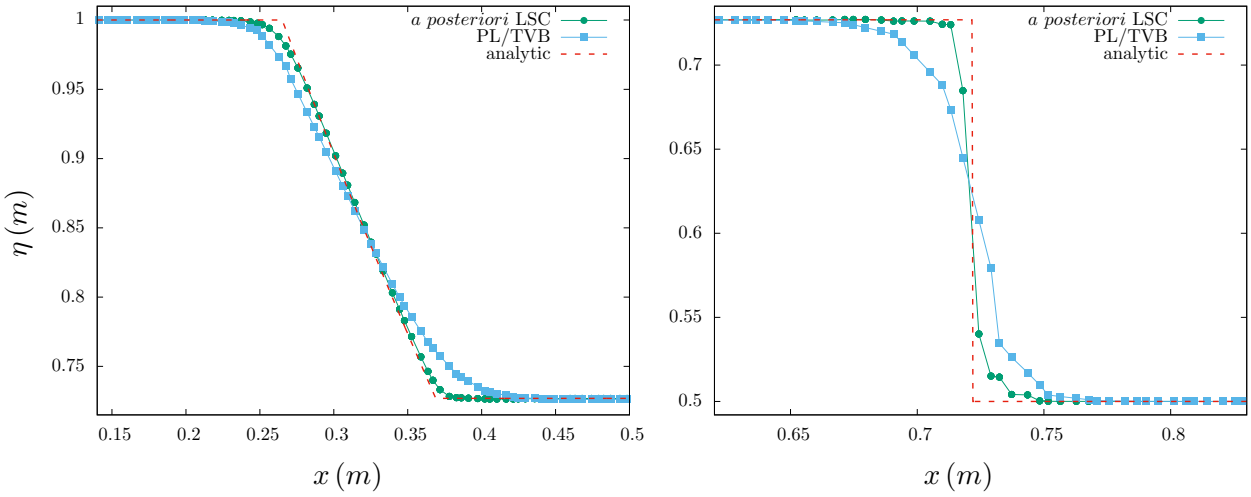

Figure 3.12: Test 6 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method and PL/TVB method for $k = 3$ and $n_{\text{el}} = 50$, with a zoom on the rarefaction wave (left) and the shock wave (right)

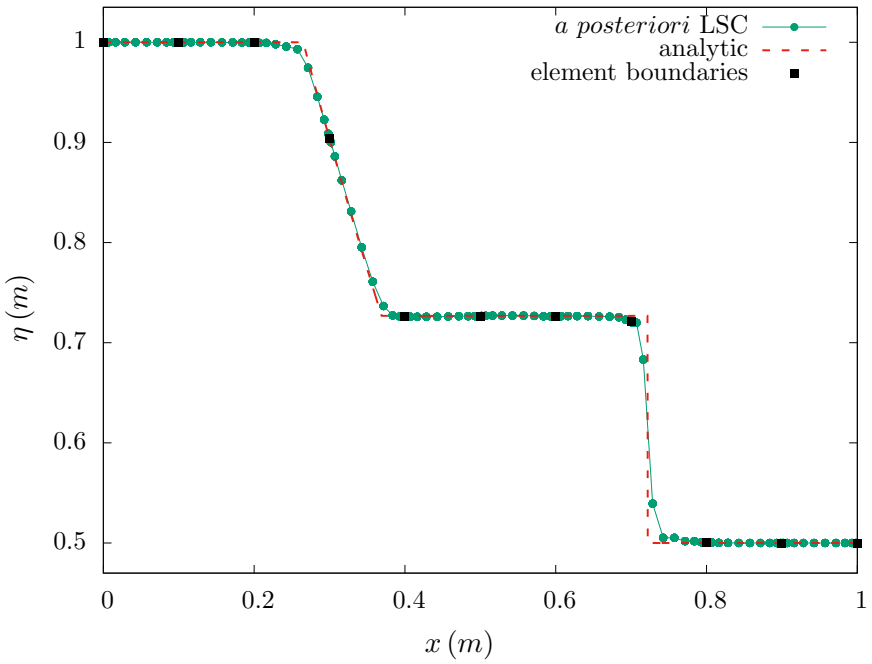

Figure 3.13: Test 6 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method (right) and PL/TVB method (left) for $k = 9$ and $n_{\text{el}} = 10$.

In a second time, we modify the initial conditions as follows:

$$\eta_0(x) = \begin{cases} 1 & \text{if} \quad x \leq x_0 \\ 0 & \text{elsewhere} \end{cases}, \quad q_0(x) = 0.$$

We compute the evolution up to $t = 0.05\,s$, with $k = 3$ and $n_{\text{el}} = 50$, in order to show the ability of the proposed method to compute the propagation of a wet/dry front. A comparison between the numerical results obtained with the *a posteriori* LSC method and the analytical solution is shown on Fig. 3.14 (left). Additionally, we compare these results with those obtained with the PL/TVB limitation process at the same times on Fig. 3.14 (right), together with zoomed profiles on Fig. 3.15 and Fig. 3.16.

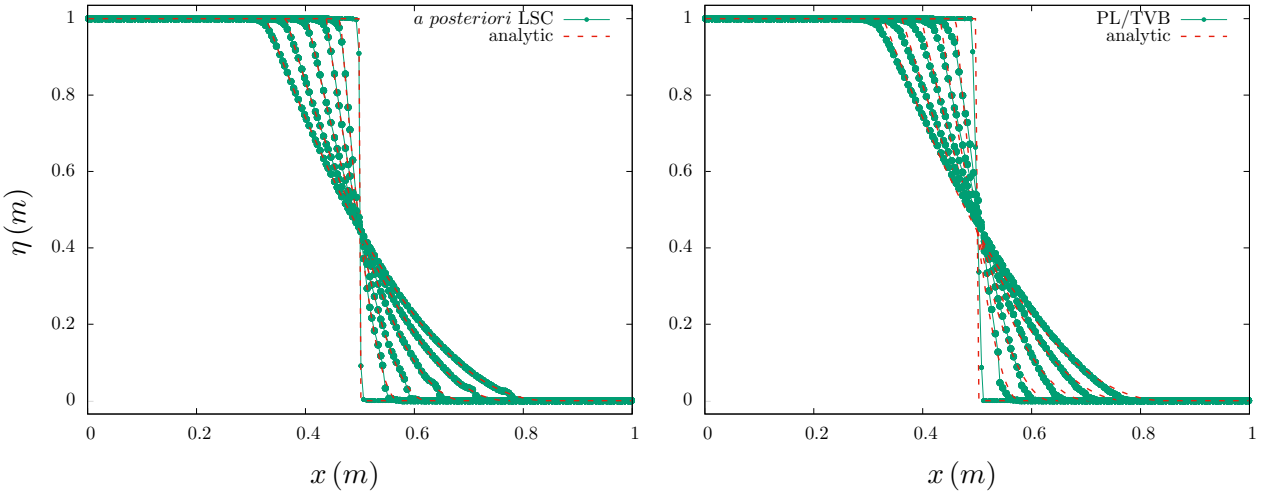

Figure 3.14: Test 7 - Dam break on a dry bottom - Free surface elevation computed at different times between $0.002s$ and $0.05\,s$ - Comparison between *a posteriori* LSC method (left) and PL/TVB method (right) for $k = 3$ and $n_{\text{el}} = 50$.

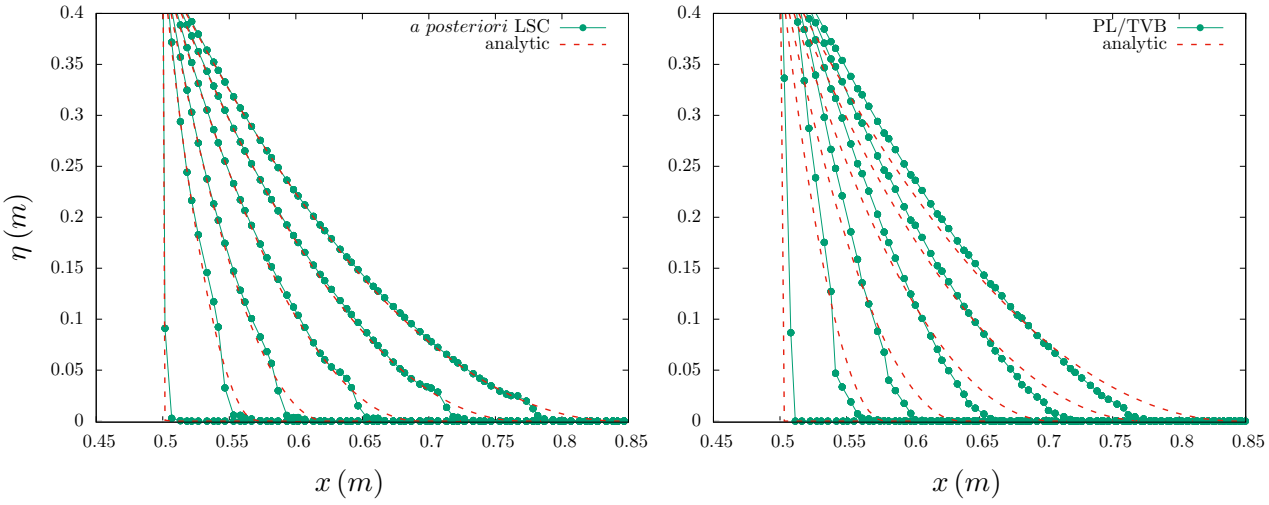

Figure 3.15: Test 7 - Dam break on a dry bottom - Free surface elevation computed at different times between $0.002s$ and $0.05\,s$ - Comparison between *a posteriori* LSC method (left) and PL/TVB method (right) for $k = 3$ and $n_{\text{el}} = 50$, with a zoom on the wet/dry interface.

Those results show how our subcell correction technique behaves in comparison to the PL/TVB limiter, in the context of the propagation of a wet/dry front. Finally, to exhibit once more the high

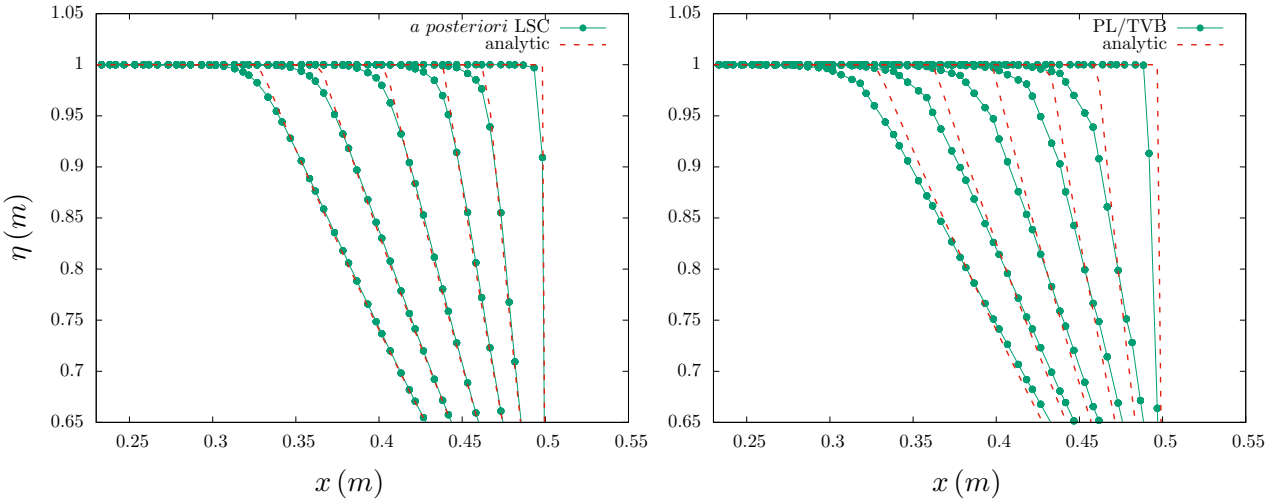

Figure 3.16: Test 7 - Dam break on a dry bottom - Free surface elevation computed at different times between $0.002s$ and $0.05\,s$ - Comparison between *a posteriori* LSC method (left) and PL/TVB method (right) for $k = 3$ and $n_{\text{el}} = 50$, with a zoom on the top.

scalability of the present *a posteriori* LSC method to very high-order of accuracy, we set $k = 8$ and $n_{\text{el}} = 10$. The corresponding numerical result is shown on Fig. 3.17.

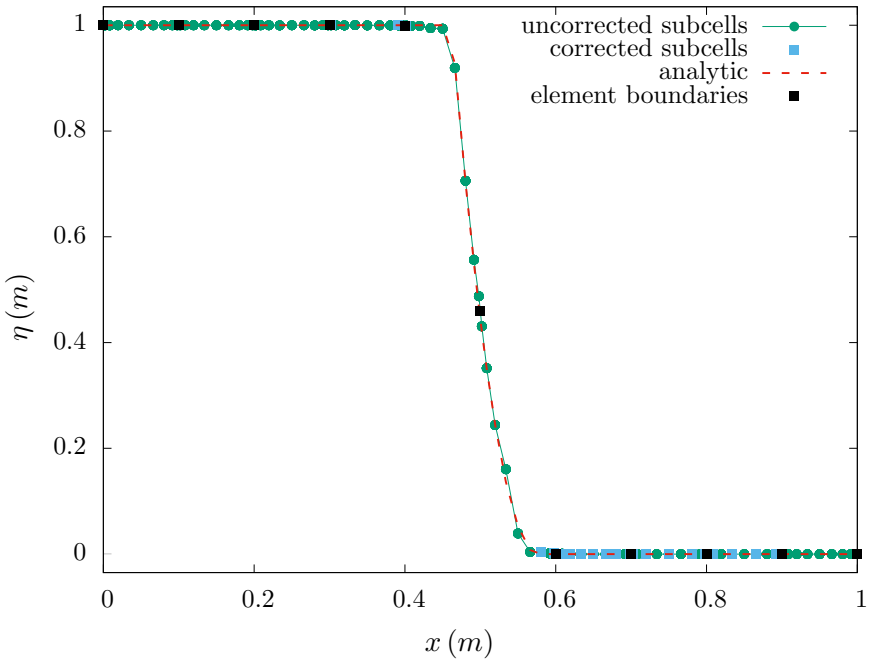

Figure 3.17: Test 7 - Dam break on a dry bottom - with the *a posteriori* LSC method for $k = 8$ and $n_{\text{el}} = 10$ at t = 0.01s.

3.10.4 Well-balancing property

In this third test, we focus on the preservation of the motionless steady states. The computational domain is $\Omega = [0, 1]$. The topography profile is defined as follows

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{(x - x_1)\cdot\pi}{75}\right)\right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere,} \end{cases} \tag{3.34}$$

where $A = 4.75$, $x_1 = 0.125$ and $x_2 = 0.875$. The initial data is defined as

$$\eta_0(x) = \max(3, b(x)) \qquad \text{and} \qquad q_0(x) = 0.$$

We evolve this initial configuration in time up to 100000 time iterations, with a fourth-order approximation and 120 mesh elements. The numerical results obtained with the *a posteriori* LSC method are shown on Fig. 3.18 and Fig. 3.19.

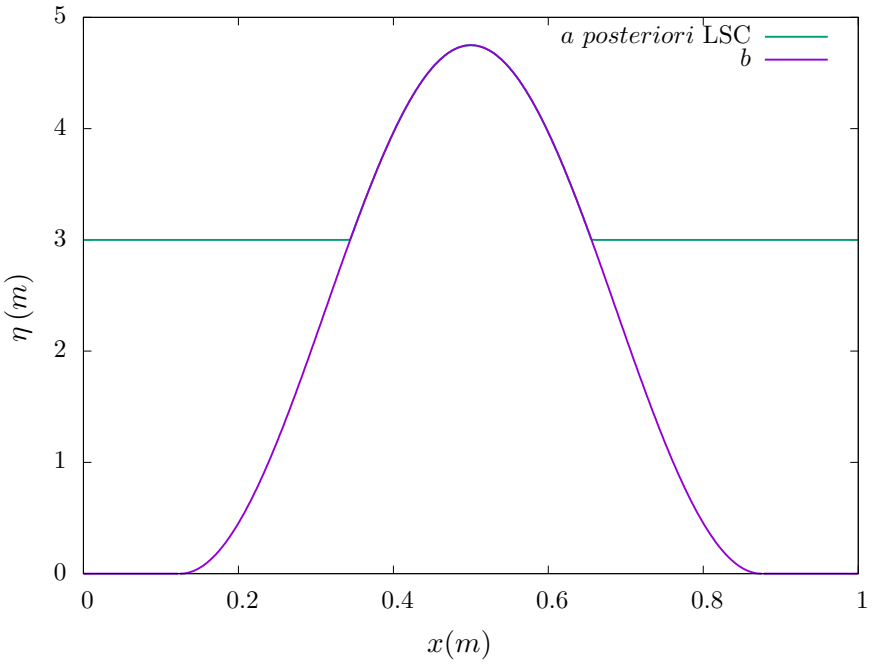

Figure 3.18: Test 8 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$.

We highlight in Fig. 3.19 the particular marked subcells, in which the correction has been performed. We emphasize that the steady state is effectively preserved up to the machine accuracy, validating numerically the compatibility of the *a posteriori* LSC method with the well-balancing property. A similar behavior is reported for other values of k and n_{el}.

Next, we slightly modify the initial condition for the water-height in order to have the bump totally submerged:

$$\eta_0(x) = 10 \qquad \text{and} \qquad q_0(x) = 0.$$

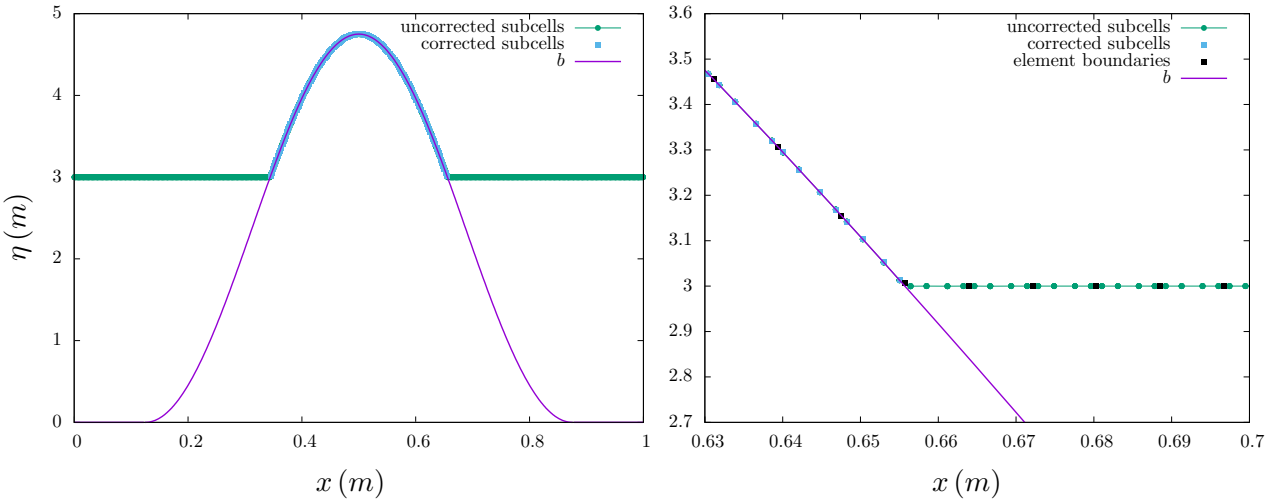

Figure 3.19: Test 8 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$ (left), with a zoom on the wet/dry interface (right).

We evolve this initial configuration in time up to $100,000$ time iterations, with a fourth-order approximation and 120 mesh elements. The numerical results obtained with the *a posteriori* LSC method are shown on Fig. 3.20.

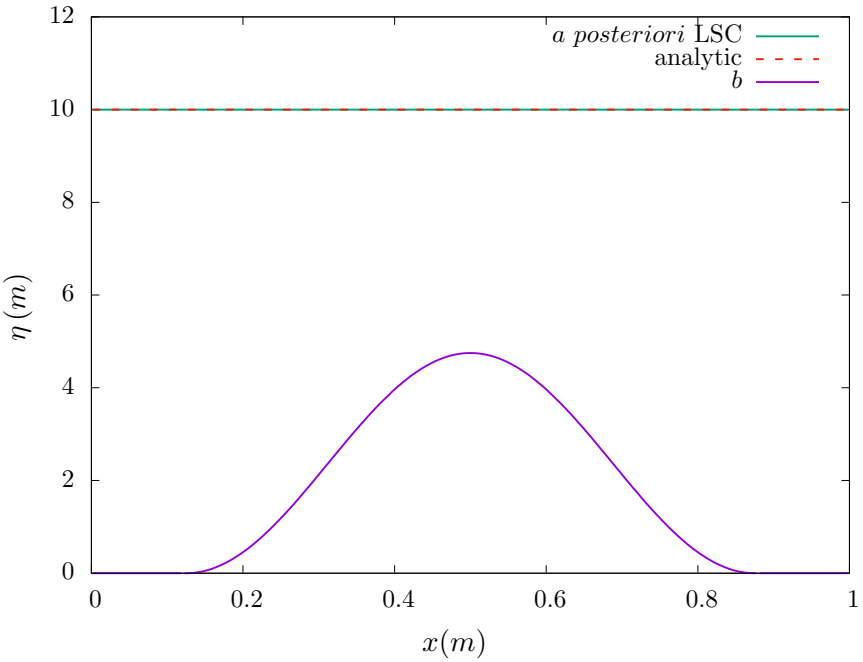

Figure 3.20: Test 9 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$.

In Table 3.3, we gather the global L^2-errors obtained for several orders of approximation, for the surface elevation at $t = 50s$. As expected the steady state is preserved up to double precision

accuracy.

k	1	2	3
h	$E^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$E^{\eta}_{L_2}$
$\frac{1}{15}$	1.35E-15	1.12E-15	6.32E-16
$\frac{1}{30}$	4.70E-16	2.61E-16	9.01E-17
$\frac{1}{60}$	1.52E-16	5.64E-17	1.03E-17
$\frac{1}{120}$	6.57E-17	1.27E-17	1.48E-18

Table 3.3: Test 9 - Preservation of a motionless steady state: L^2-errors between numerical and exact steady state solutions for η at time $t = 50s$.

3.10.5 Trans-critical flow over a bump: without shock

We focus in this test on a classical trans-critical flow without shock, see for instance [75] for a complete description. The computational domain is $\Omega = [0, 25]\,(m)$. The topography profile is defined as follows:

$$b(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8 < x < 12, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.35}$$

In this test, the incoming flow is enforced to be fluvial upstream and becomes torrential at the top of the bump. The initial data is defined as:

$$\eta_0(x) = 0.66 \;\; m \quad \text{and} \quad q_0(x) = 0 \;\; m^3.s^{-1},$$

and we prescribe the following boundary conditions:

$$\begin{cases} \text{upstream: } q = 1.53 \;\; m^3.s^{-1}, \\ \text{downstream: } h = 0.66 \;\; m \text{ while the flow is subcritical.} \end{cases}$$

We run this test-case with $k = 3$, $n_{\text{el}} = 100$ and $t_{max} = 200s$. We show on Fig. 3.21 the free surface elevation and the discharge obtained with the *a posteriori* LSC method , at several moments during the transient part of the flow ($3.55\,s$ and $20.3\,s$) and when the steady state is reached ($200\,s$), showing a very good agreement with the analytical solution.

3.10.6 Transcritical flow over a bump: with shock

Now we include some modifications in order to obtain a transcritical flow over a bump with shock. The computational domain is always $\Omega = [0, 25]\,(m)$ and the topography profile is defined as in (3.35). The initial data is defined as:

$$\eta_0(x) = 0.33 \;\; m \quad \text{and} \quad q_0(x) = 0 \;\; m^3.s^{-1},$$

and we prescribe the following boundary conditions:

$$\begin{cases} \text{upstream: } q = 0.18 \;\; m^3.s^{-1}, \\ \text{downstream: } h = 0.33 \;\; m \text{ while the flow is subcritical.} \end{cases}$$

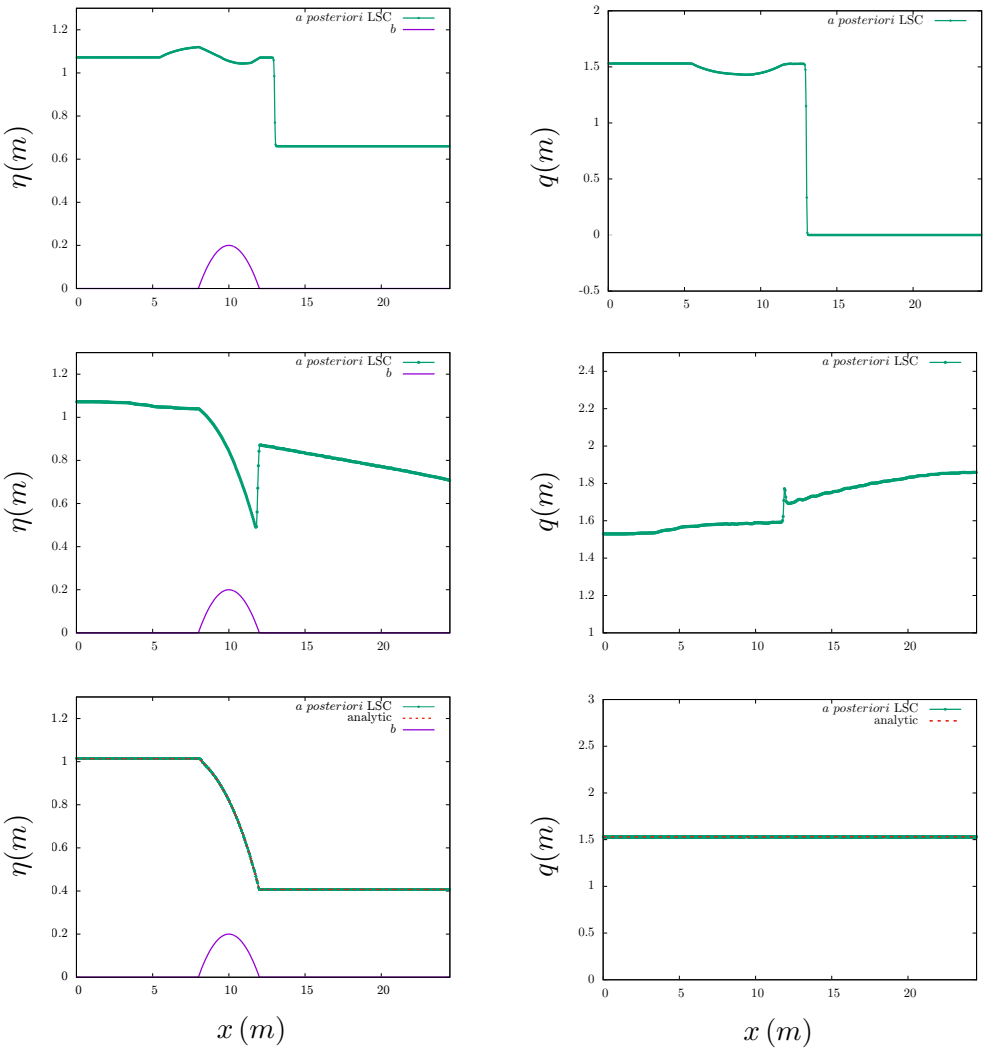

Figure 3.21: Test 10 - Transcritical flow over a bump without shock - Free surface elevation and discharge computed at several moments, $3.55s$, $20.3s$ and $200s$, with the *a posteriori* LSC method , for $k = 3$ and $n_{\text{el}} = 100$.

We run this test-case with $k = 3$, $n_{\text{el}} = 100$ and $t_{max} = 200s$. We show on Fig. 3.22 the free surface elevation obtained with the *a posteriori* LSC method , when the steady state is reached at $t = 200s$, showing a very good agreement with the analytical solution.

3.10.7 Transcritical flow over a bump and through a contraction

Here we conducted flow simulation as a response to both a contraction and a bump. The bump used in this simulation is always (3.35). The initial condition used are:

$$\eta_0(x) = 0.5 \;\; m \quad \text{and} \quad q_0(x) = 0 \;\; m^3.s^{-1}.$$

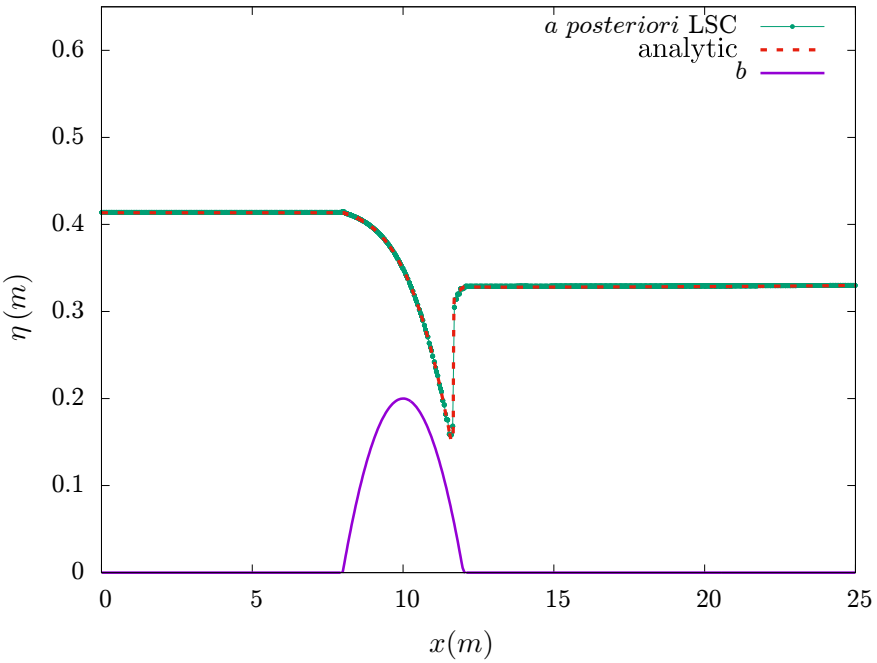

Figure 3.22: Test 11 - Transcritical flow over a bump with shock - Free surface elevation at $t = 200\,s$, with the *a posteriori* LSC method , for $k = 3$ and $n_{\text{el}} = 100$.

Boundary conditions used here are a hard wall boundary on the left, and an absorbing boundary on the right. The computed water elevation at subsequent times is plotted in Fig. 3.23, showing a very good behavior at all time stages and a very good agreement with the analytical solution at steady state. As time progresses, water on the downstream part is draining out, by the implementation of the absorbing right boundary, whereas on the upstream part, the water is being trapped by the hard wall and the bump.

3.10.8 Carrier and Greenspan's transient solution

This test-case, introduced in [33], describes the physical process in which the water level near the shoreline of a sloping beach is initially depressed, the fluid held motionless and then released at $t = 0$. A transient wave is generated which runs up the beach, before returning to equilibrium state in a slow convergence process, reproducing some interesting conditions for assessing the robustness of the *a posteriori* LSC method in computing long waves run-up. In [33], a hodograph transformation is used to solve the NSW equations and obtain an analytical solution. The transformation makes use of two dimensionless variables (in the following, starred variables denote dimensionless quantities) σ^* and λ^* which are, respectively, a space-like and a time-like coordinate given by

$$\sigma^* = 4c^*, \quad \lambda^* = 2\,(u^* + t^*)\,.$$

Let l be the typical length scale of this specific problem and α the beach slope. The scales used to obtain the nondimensionalized variables are:

$$x^* = x/l, \quad \eta^* = \eta/(\alpha l), \quad u^* = u/\sqrt{g\alpha l}, \quad t^* = t/\sqrt{l/\alpha g}, \tag{3.36}$$

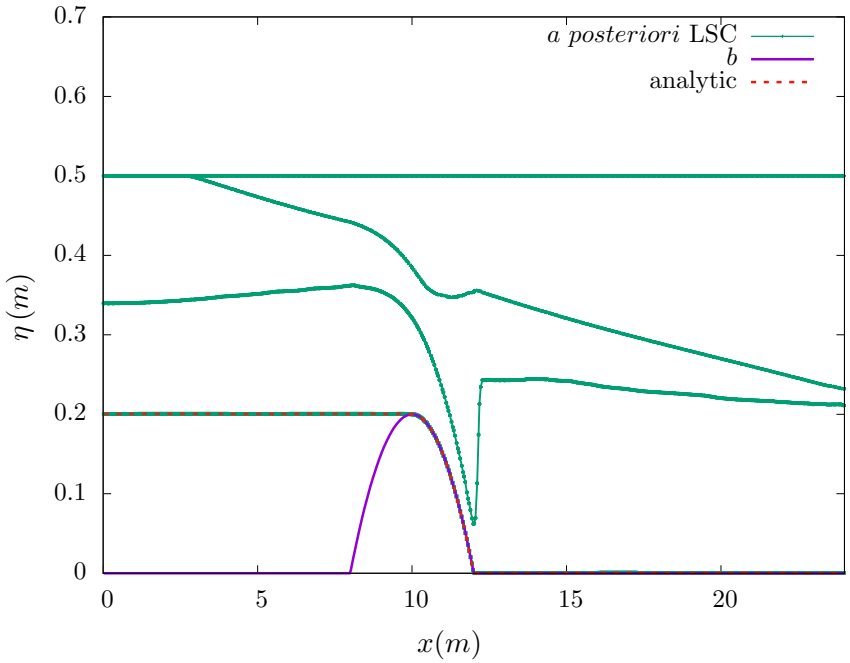

Figure 3.23: Test 12 -Transcritical flow over a bump and through a contraction - Free surface elevation at water elevation at subsequent times with the *a posteriori* LSC method , for $k = 3$ and $n_{\text{el}} = 100$.

and the non-dimensional phase speed is given by:

$$c^* = \sqrt{\eta^* - x^*}. \tag{3.37}$$

The initial solution is specified by the following initial conditions:

$$\eta_0^*(\sigma^*) = e\Big(1 - \frac{5}{2}\frac{a^3}{(a^2+\sigma^{*2})^{\frac{3}{2}}} + \frac{3}{2}\frac{a^5}{(a^2+\sigma^{*2})^{\frac{5}{2}}}\Big), \qquad q_0^*(\sigma^*) = 0 \qquad \text{and} \qquad x^* = -\frac{\sigma^{*2}}{16} + \eta_0^*, \tag{3.38}$$

where $a = \frac{3}{2}(1+0.9e)^{\frac{1}{2}}$ and e is a small parameter which characterizes the surface elevation profile. The analytical solution is then given by

$$\begin{cases} \eta^*(\sigma,\lambda) = -\dfrac{u^{*2}}{2} + eR_e\left[1 - 2\dfrac{5/4 - i\lambda}{[(1-i\lambda)^2+\sigma^2]^{\frac{3}{2}}} + \dfrac{3}{2}\dfrac{(1-i\lambda)^2}{[(1-i\lambda)^2+\sigma^2]^{\frac{5}{2}}}\right], \\ u^*(\sigma,\lambda) = \dfrac{8e}{a} I_m\left[\dfrac{1}{[(1-i\lambda)^2+\sigma^2]^{\frac{3}{2}}} - \dfrac{3}{4}\dfrac{1-i\lambda}{[(1-i\lambda)^2+\sigma^2]^{\frac{5}{2}}}\right], \\ t^* = \dfrac{1}{2}a\lambda - u^* \qquad \text{and} \qquad x^* = \eta^* - \dfrac{a^2\sigma^2}{16}, \end{cases}$$

where we have set $\sigma^* = a\sigma, \lambda^* = a\lambda$. This set of equations may be solved by some iterative process. In what follows, we set $e = 0.1$, $\alpha = 1/50$, the initial surface profile (3.38) is provided in

the dimensional case with the length scale $l = 20$ m and we define $\beta := \alpha\, e\, l$. We run this test-case with $k = 3$ and 50 mesh elements, for different values of discrete time t in the range $[0.5\,s, 23\,s]$, see Fig. 3.24 (left) and at $t = 200\,s$ on Fig. 3.24 (right).

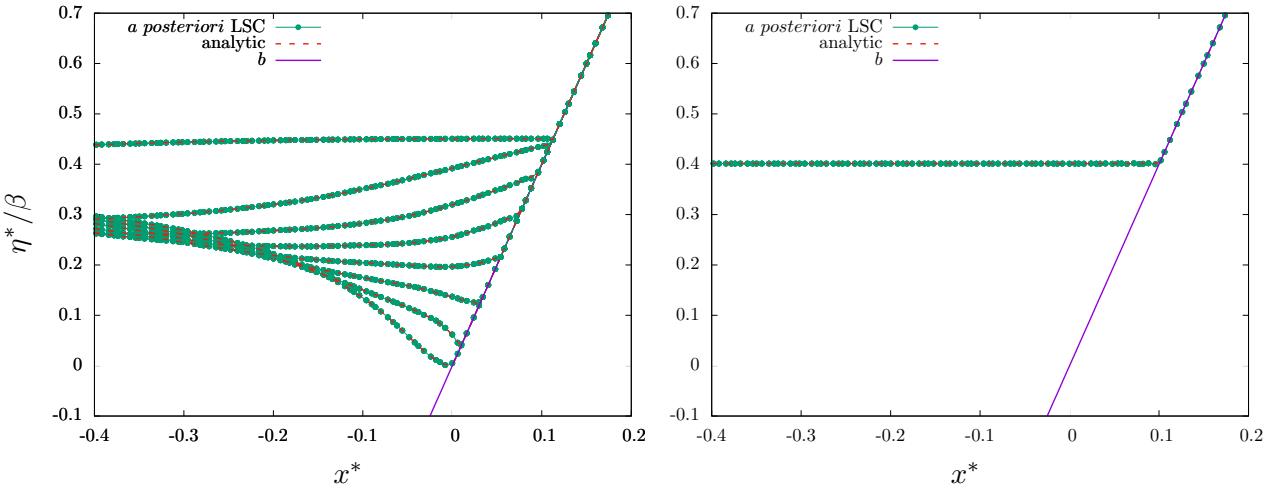

Figure 3.24: Test 13 - Carrier and Greenspan's transient solution - Free surface elevation η^*/β plotted versus the onshore coordinate x^* - Free surface elevation for different values of time in the range $[0.5\,s,\ 23\,s]$ (left) and at $t = 200\,s$ (right) for $k = 3$ and $n_{\text{el}} = 50$.

In view of result displayed in Fig.3.24, one can see how accurate DG scheme along with our *a posteriori* LSC method is, as the numerical solution is extremely close to the exact solution and is able to simulate the return to the equilibrium state. This is due to the ability of our correction method to surgically modified the numerical solution only in the very few concerned subcells, as illustrated on Fig. 3.25. Additionally, we compare these results with those obtained with the PL/TVB method on Fig. 3.27 with $M = 0$ (left) and with $M = 32$ (right). Let us note that in [179], the authors make use of $M = 0$ in every situations, except for the convergence rate analysis where $M = 32$ is used. As this test-case is for the most part smooth (except at the wet/dry transition point), a non-zero value of M can be used in order to improve the quality of the results, as depicted by Fig. 3.27. However, even for higher value of M ($M = 32$), the PL/TVB limiter is outperformed by the present *a posteriori* LSC method .

We finally assess the use of a high-order polynomial approximation ($k = 8$) on a very coarse mesh ($n_{\text{el}} = 10$) to emphasize the very accurate and interesting subcell resolution ability of the proposed approach. The results obtained at $t = 7\,s$ are plotted on Fig. 3.28.

In Table 3.4, we gather the global L^2-errors associated with the computation of η, for different polynomial orders, and increasing mesh refinements, computed at $t^* = 1\,s$, for Carrier and Greenspan's transient test-case. We emphasize that in such situations, in which the *a posteriori* LSC is activated, the resulting scheme is not a "pure" DG scheme, but a combination of a pure DG scheme with a first order FV scheme. Also the solution is only $H^1(\Omega)$ and the regularity is not sufficient to obtain the "optimal" order. As a consequence, it is no surprise that, the observed rates of convergence are not the optimal rates generally associated for a "sufficiently regular" solution with pure DG schemes. Measuring order of convergence in such test-cases is not a criterion for accuracy of our DG scheme. The order of convergence of our DG scheme has already been calculated in § 3.10.1 and § 3.10.2, where a very good results for errors and order of convergence are shown.

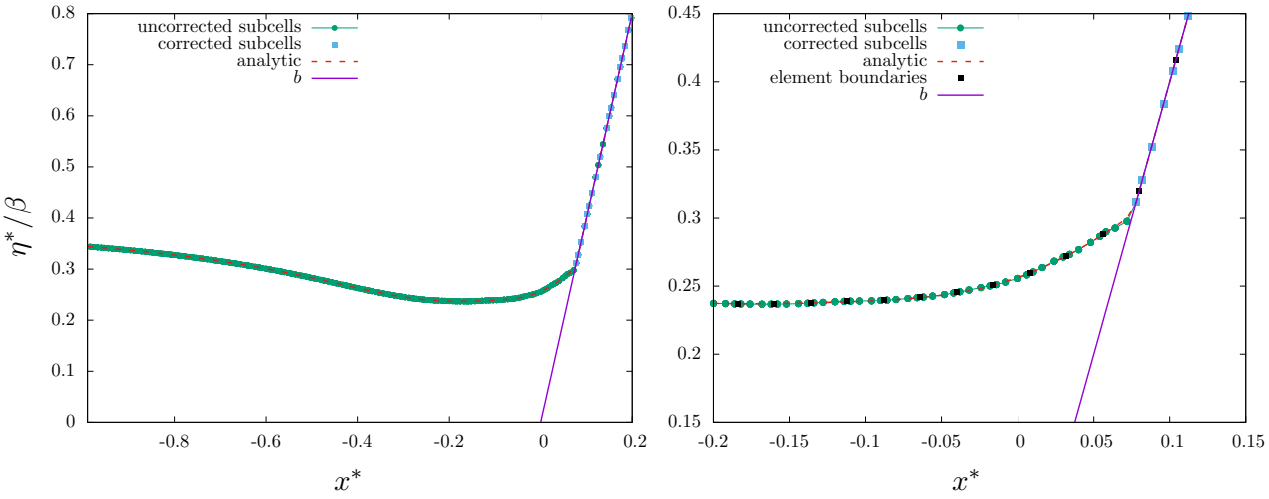

Figure 3.25: Test 13 - Carrier and Greenspan's transient solution - Free surface elevation computed at $t = 7\,s$ with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 50$ (left): corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the shoreline (right)

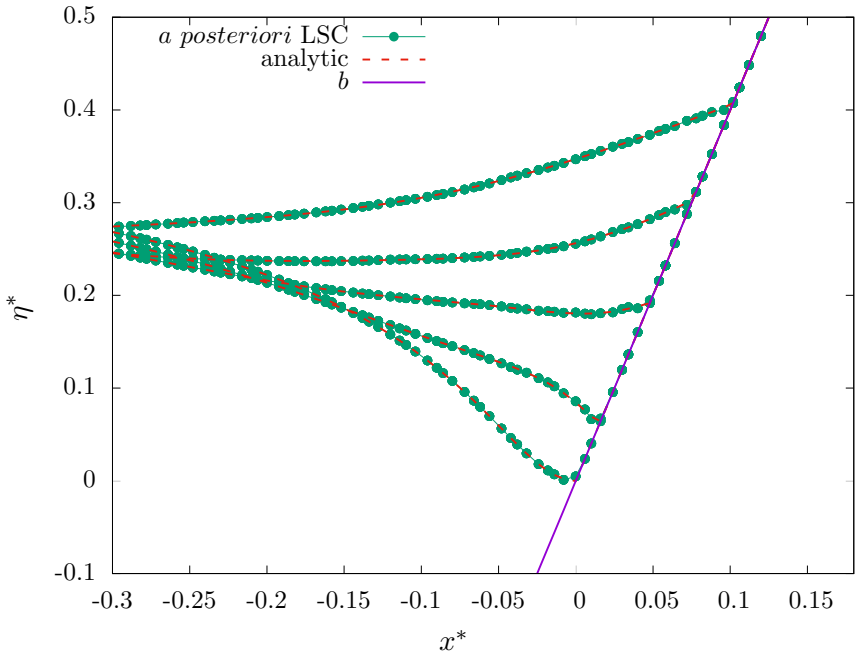

Figure 3.26: Test 13 - Carrier and Greenspan's transient solution - Free surface elevation computed for different values of time in the range $[0.5\,s, 15\,s]$ with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 50$.

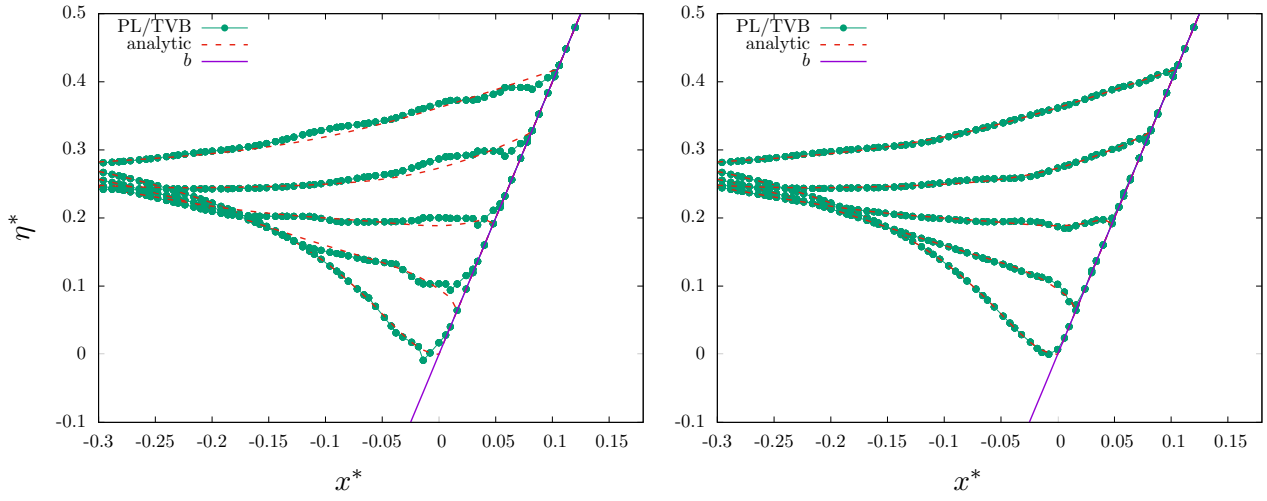

Figure 3.27: Test 13 - Carrier and Greenspan's transient solution - Free surface elevation computed for different values of time in the range $[0.5\,s, 15\,s]$ with the PL/TVB method for $k = 3$ and $n_{\text{el}} = 50$, with $M = 0$ (left) and $M = 32$ (right).

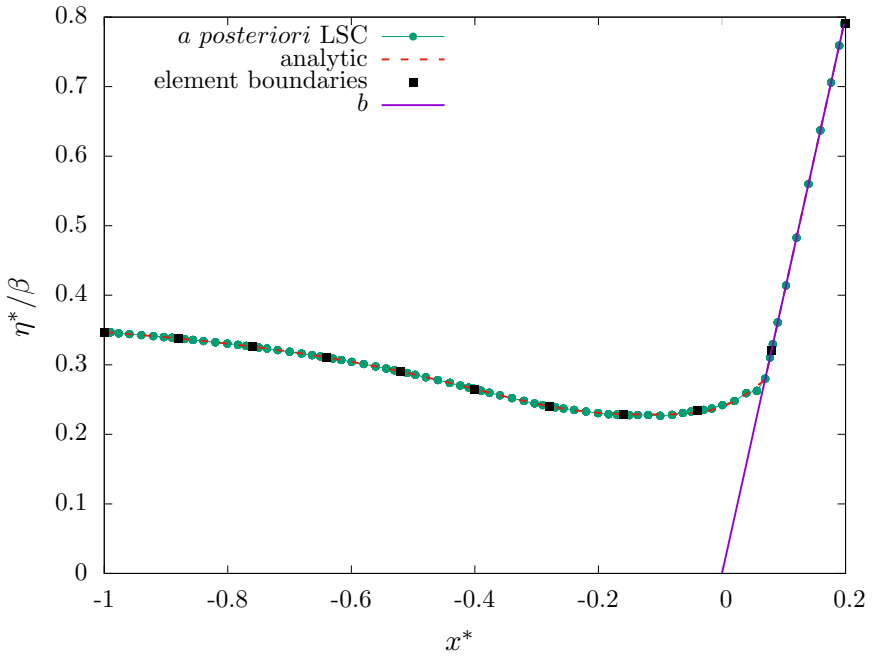

Figure 3.28: Test 13 - Carrier and Greenspan's transient solution - Free surface elevation computed at $t = 7\,s$ with the *a posteriori* LSC method for $k = 8$ and $n_{\text{el}} = 10$.

3.10.9 Carrier and Greenspan's periodic solution

In this test-case, a monochromatic wave is let run-up and run-down on a plane beach. This solution represents the motion of a periodic wave of dimensionless amplitude A^* and frequency ω^* traveling shoreward and being reflected out to sea generating a standing wave on a plane beach. Recalling

k	1		2		3	
h	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$	$E^{\eta}_{L_2}$	$q^{\eta}_{L_2}$
$\frac{1}{15}$	7.02E-3	3.49	1.27E-3	4.23	8.72E-4	4.95
$\frac{1}{30}$	6.23E-4	1.94	6.77E-5	2.43	2.80E-5	2.05
$\frac{1}{60}$	1.61E-4	1.98	1.25E-5	2.10	6.75E-6	1.95
$\frac{1}{120}$	4.07E-5	-	2.92E-6	-	1.74E-6	-

Table 3.4: Test 13 - Carrier and Greenspan's transient solution: L^2-errors between numerical and analytical solutions for η at time $t^* = 1s$

the dimensionless quantities (3.36) and (3.37), the analytical solution is formulated as follows:

$$\begin{cases} u^* = -\dfrac{A^* J_1(\sigma^*)\sin(\lambda^*)}{\sigma^*}, \\ \eta^* = \dfrac{A^*}{4} J_0(\sigma^*)\cos(\lambda^*) - \dfrac{u^{*2}}{4}, \\ t^* = \dfrac{1}{2}\lambda^* - u^* \quad \text{and} \quad x^* = \eta^* - \dfrac{\sigma^{*2}}{16}, \end{cases}$$

where J_0 and J_1 stand for the Bessel functions of zero and first order. We consider the solution obtained for $A^* = 0.6$ and $\omega^* = 1$ (non-breaking wave), together with the length scale $l = 20\,m$ and a bottom slope $\alpha = 1/30$. The value of this solution at $t = 0$ is supplied as initial condition, and similarly to the previous transient case, the analytical variations of the surface elevation at the left boundary is used as an offshore inlet boundary condition, generating the motion. We refer the reader to [33] for a complete description. We set $k = 3$ and $n_{\text{el}} = 50$ and we compute the time evolution up to $t = 1.5\,T$, where T is the time period of the periodic forcing. We show on Fig. 3.29 some snapshots of the free surface elevation plotted at several discrete time in the range $[1.25T, 1.5T]$ with the *a posteriori* LSC method , showing a very good agreement between the numerical solution and the analytical one. Additionally, we compare these results with those obtained with the PL/TVB method on Fig. 3.30 with $M = 0$ (left) and with $M = 32$ (right). However, even for higher value of M ($M = 32$), the PL/TVB limiter is outperformed by the present *a posteriori* LSC method .
In order to emphasize the accuracy of the proposed approach for long time integration, we set $t = 15T$ and show on Fig. 3.31 the free surface elevation obtained at times $t = 14.5T$ (left) and $t = 15T$ (right), for $k = 3$ and $n_{\text{el}} = 50$. We observe that such a long time-integration has a negligible impact on the accuracy of the predictions of the shoreline location. Such a result can be reproduced with a high-order approximation $k = 8$ and a very coarse mesh $n_{\text{el}} = 10$, showing again the ability of our approach to provide a high-order accurate subcell description of the motion, see Fig. 3.33. In Fig. 3.32, we show time-series of the shoreline elevation "η_s" in the range $[0, 6T]$ (left) and $[0, 15T]$ (right). We can see that the minimum and maximum water elevations are accurately computed, even after a large number of periods.

3.10.10 Run-up of a solitary wave on a plane beach

The last test-case is devoted to the computation of the run-up of a solitary wave on a constant slope. Such run-up phenomena are investigated experimentally and numerically in [152]. In this test, a solitary wave traveling from the shoreward is let run-up and run-down on a plane beach, before being fully reflected and evacuated from the computational domain. The topography is made of a constant

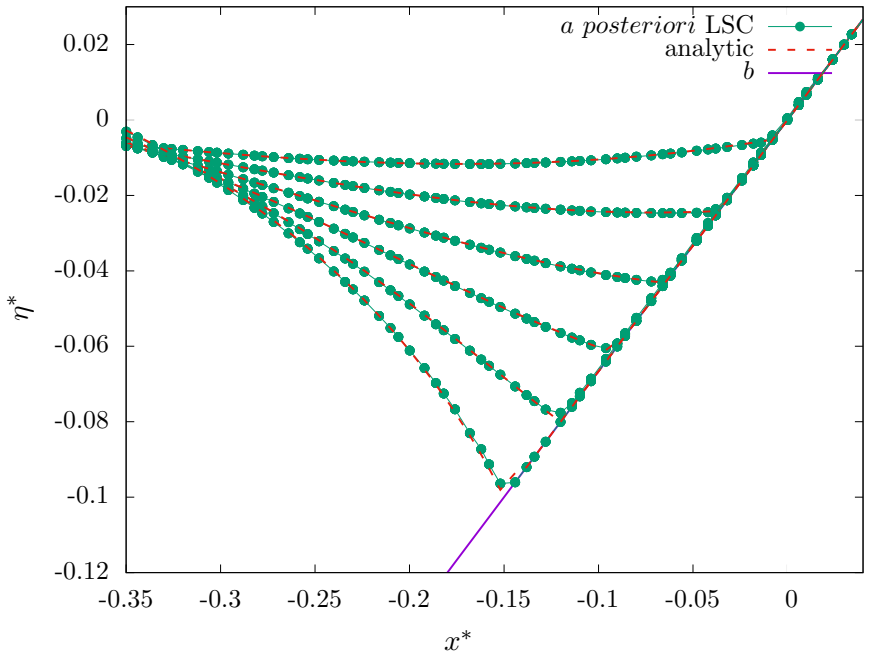

Figure 3.29: Test 14 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[1.25T, 1.5T]$ with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 50$

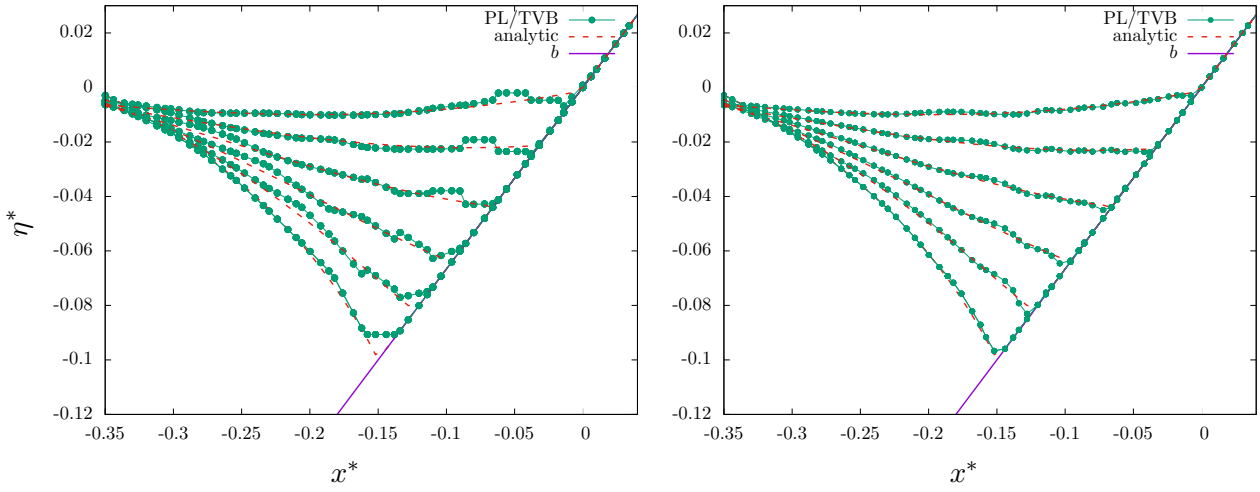

Figure 3.30: Test 14 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[1.25T, 1.5T]$ with the PL/TVB method for $k = 3$ and $n_{\text{el}} = 50$, with $M = 0$ (left) and $M = 32$ (right).

depth area juxtaposed with a plane sloping beach of constant slope α such that $\cot(\alpha) = 19.85$. The right boundary condition is transmissive. The initial condition is defined as follows:

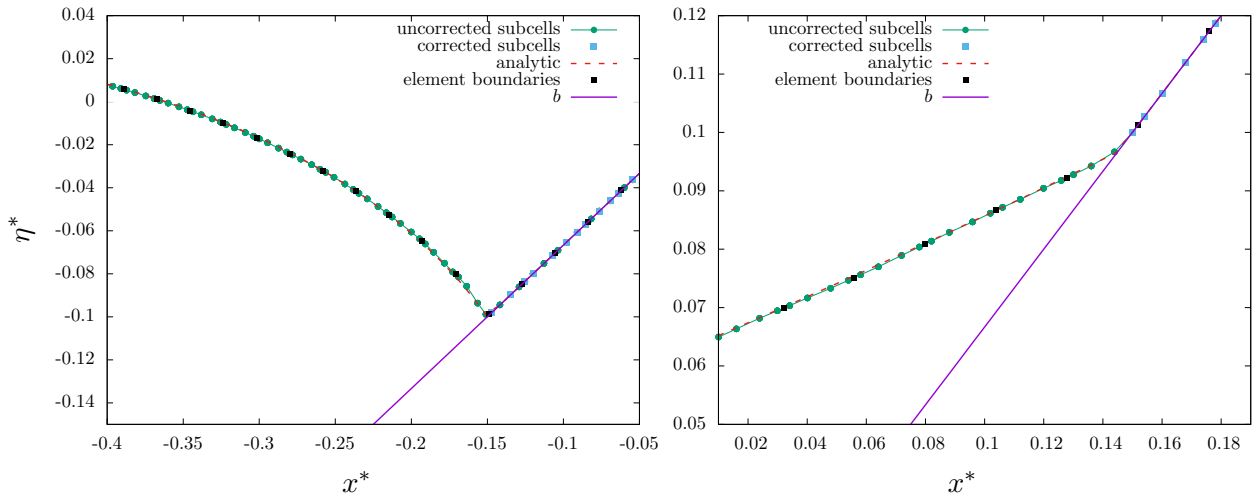

Figure 3.31: Test 14 - Carrier and Greenspan's periodic solution - Free surface elevation computed at $t = 14.5T$ (left) and $t = 15T$ (right) with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 50$.

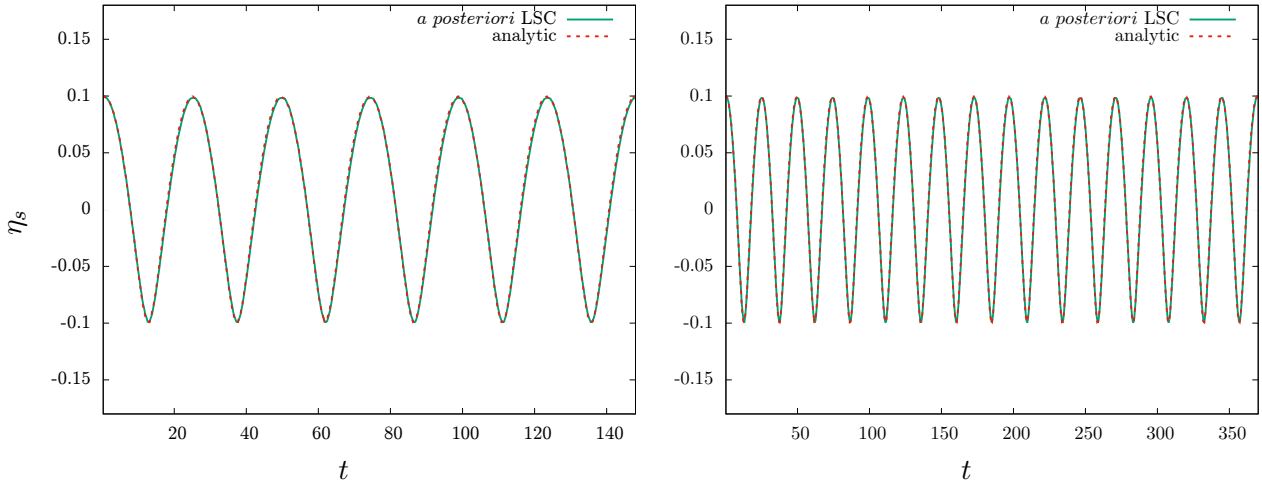

Figure 3.32: Test 14 - Carrier and Greenspan's periodic solution - Time-series of the shoreline elevation in the range $[0, 6T]$ (left) and $[0, 15T]$ (right), with the *a posteriori* LSC method for $k = 3$ and $n_{\text{el}} = 60$.

$$\eta_0(x) = H_0 + \frac{A}{H_0}\,\text{sech}^2\left(\gamma\,(x - x_1)\right) \qquad \text{and} \qquad u_0(x) = \sqrt{\frac{g}{H_0}}\,(\eta_0(x) - H_0)\,,$$

where $\gamma = \sqrt{\dfrac{3A}{4H_0}}$ and $x_1 = \sqrt{\frac{4H_0}{3A}}\,\text{arcosh}\left(\sqrt{\frac{1}{0.05}}\right)$ is nothing but the initial position of the center of the solitary wave. This test is run with $A = 0.019\,m$, $H_0 = 1.0\,m$, $k = 8$, $n_{\text{el}} = 20$ and $t = 40\,s$. We show on Fig. 3.34 the free surface obtained with the *a posteriori* LSC method at several times in the range $[1\,s,\ 40\,s]$, showing once more a very good agreement with the reference solution obtained with a robust FV method on a very fine mesh $n_{\text{el}} = 10000$.

In this work, we have introduced a new well-balanced high-order DG discrete formulation with a

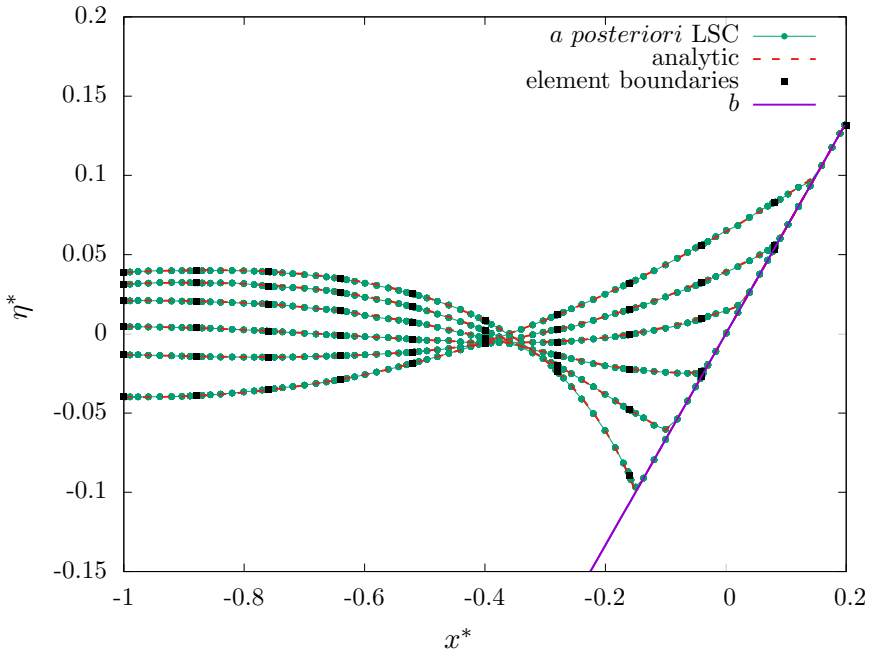

Figure 3.33: Test 14 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[14.5T, 15T]$ for $k = 8$ and $n_{\text{el}} = 10$.

FV-Subcell correction patch designed for the NSW equations. This formulation, based on [164], combines the very high accuracy of DG schemes along with a robust correction procedure ensuring the water-height positivity as well as addressing the issue of spurious oscillations in the vicinity of discontinuities. This robustness is enforced by means of an *a posteriori* LSC of the conservative variables. This procedure relies on an advantageous reformulation of DG schemes as a FV-like method on a sub-grid, which makes the correction strategy surgical and flexible, as well as conservative at the subcell level. Indeed, only the non-admissible subcells are marked and subject to correction, retaining as much as possible the very accurate subcell resolution of high-order DG formulations. The proposed strategy is investigated through an extensive set of benchmarks, including a brand new smooth solution for the computation of convergence rates, stabilization of flows with discontinuities, the preservation of motionless steady states, or moving shorelines over varying bottoms. We observe in particular that this approach provides a very accurate description of wet/dry interfaces even with the use of very high-order schemes on coarse meshes.
Regarding potential advantages of this *a posteriori* limiting strategy compared to *a priori* limiters, because the troubled zone detection is performed *a posteriori*, the correction can be done only where it is absolutely necessary. Furthermore, positivity preservation of the water-height is included without any additional effort, while it is generally not the case of *a priori* limitations of high-order schemes. Let us further emphasize that this *a posteriori* LSC method scalability to any order of accuracy is also perfectly natural. Finally, it is important to note that this new correction procedure is totally parameter free.

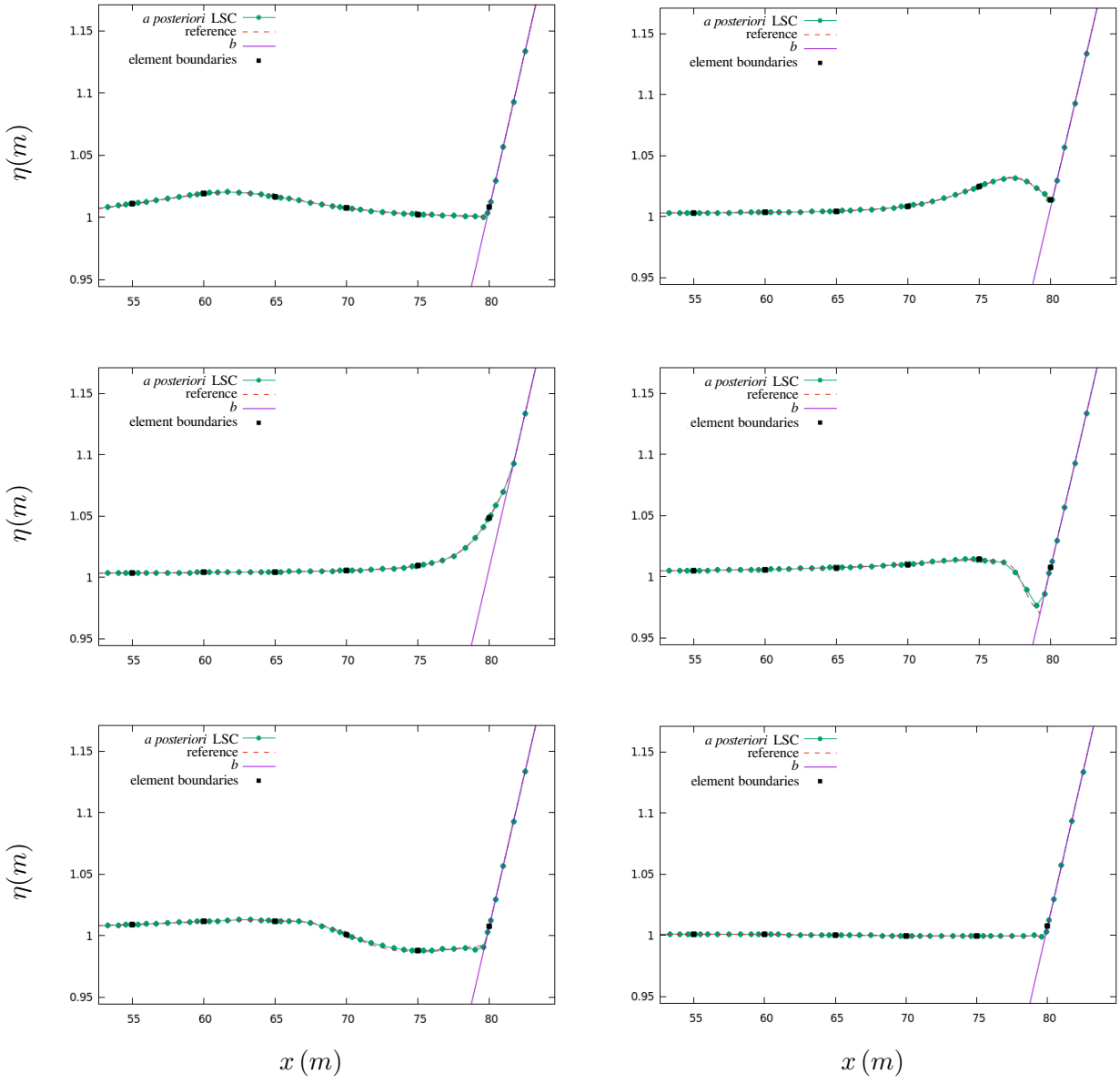

Figure 3.34: Test 15 - Run-up of a solitary wave on a plane beach - Free surface elevation computed for different values of time in the range $[1\,s, t = 40\,s]$ with the *a posteriori* LSC method obtained for $k = 8$ and $n_{\text{el}} = 20$.

Chapter 4

An *a posteriori* LSC method for the NSW equations: the 2d case

The *a posteriori* LSC procedure introduced in chapter § 3 for the one-dimensional NSW model can be extended to the 2d NSW system, following the steps presented in [165]. Many articles have been devoted to the study of correction strategies for high-order DG scheme using FV-Subcell method for multi-dimensional hyperbolic systems. The reader may refer for instance to [151, 61, 59, 88]. In all these aforementioned papers, the idea of the proposed correction procedure is the following: If the numerical solution in a cell is detected as bad, the cell is then subdivided into subcells and a first-order FV, or alternatively other robust scheme is applied on each subcell, *i.e.*, the entire cell is corrected. Contrary to our *a posteriori* LSC method, where the correction is strictly local. In fact, only the non-admissible subcells are corrected via a first-order FV scheme on those marked subcells without impacting the high-order DG solution elsewhere in the cell. There are very few works in the literature which are devoted to the study of such a local correction procedure, and none up to our knowledge for multi-dimensional hyperbolic systems on unstructured grids. In this chapter, an arbitrary-order DG discretization is proposed for the 2d NSW system with topography source term. Then, similarly to what has been done in [165], our *a posteriori* LSC method is extended to this 2d system. Let us mention that this 2d extension is still an ongoing project, and will be the topic of a near future article.

Denoting by $\mathbf{q} = (q_x, q_y)^t$ the water discharge as a vector variable and $\mathbf{u} = (u_x, u_y)^t$ the depth-averaged water velocity vector, the 2d NSW equations are commonly written as follows :

$$\frac{\partial \mathbf{V}}{\partial t} + \nabla \cdot \mathbf{G}(\mathbf{V}) = \mathbf{B}(\mathbf{V}, \nabla b) \tag{4.1}$$

with

$$\mathbf{V} = \begin{pmatrix} \eta \\ q_x \\ q_y \end{pmatrix}, \mathbf{G}(\mathbf{V}) = \big(\mathbf{G}^1(\mathbf{V}), \mathbf{G}^2(\mathbf{V})\big) = \begin{pmatrix} q_x & q_y \\ u_x q_x + \frac{1}{2}\mathrm{g}(\eta^2 - 2\eta b) & u_y q_x \\ u_x q_y & u_y q_y + \frac{1}{2}\mathrm{g}(\eta^2 - 2\eta b) \end{pmatrix} \tag{4.2}$$

and the bathymetry source term defined by :

$$\mathbf{B}(\mathbf{V},\nabla b) = \begin{pmatrix} 0 \\ -g\eta\partial_x b \\ -g\eta\partial_y b \end{pmatrix}.$$

4.1 Discrete formulation

In this section, we use smilar notations than for the 1d case, and introduce some new ones when it is necessary. Let Ω be the computational domain and we consider a triangulation $\mathcal{T} = \{T_1, \ldots, T_{n_{\text{el}}}\}$ of Ω in open disjoint triangles T of boundary ∂T such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}} \overline{T}$. The partition is characterized by the mesh size $h := \max_{T\in\mathcal{T}} |T|$, where $|T|$ is the volume of element T. For a given mesh element $T_i \in \mathcal{T}$, we note by c_i its barycenter.

We aim at computing an approximate vector solution on this triangulation. Given an integer polynomial degree $k \geq 1$, we define:

$$\mathbb{P}^k(\mathcal{T}) := \left\{ v \in L^2(\Omega), \quad v_{|T} \in \mathbb{P}^k(T), \quad \forall T \in \mathcal{T} \right\},$$

where $\mathbb{P}^k(T)$ denotes the space of 2-variables polynomials in T of degree at most k.

A weak formulation of the problem is obtained by multiplying (4.1) by a test function $\phi \in \mathbb{P}^k(\mathcal{T})$. We integrate locally on a mesh element T_i and the flux term is integrated by part to obtain :

$$\begin{array}{ll} \int_{T_i} \frac{\partial}{\partial t}\mathbf{V}(\mathbf{x},t)\phi(\mathbf{x})d\mathbf{x} - \int_{T_i} & \mathbf{G}(\mathbf{V},b) \cdot \nabla\phi(\mathbf{x})d\mathbf{x} + \\ & \int_{\partial T_i} \mathbf{G}(\mathbf{V},b) \cdot \vec{\mathbf{n}}_{\partial T_i}\phi(s)ds = \int_{T_i} \mathbf{B}(\mathbf{V},\nabla b)\phi(\mathbf{x})d\mathbf{x}, \end{array} \tag{4.3}$$

where $\vec{\mathbf{n}}_{\partial T_i}$ is the unit outward normal of ∂T_i. The approximated vector solution $\mathbf{V}_h \in \mathbb{P}^k(\mathcal{T})^3$ is expressed as a polynomial of order k on each element T :

$$\mathbf{V}_h(\mathbf{x},t) = \sum_{j=1}^{n_d} \mathbf{V}_j^T(t)\varphi_j^T(\mathbf{x}), \quad \forall \mathbf{x} \in T, \forall t \in [0, t_{\max}]$$

where $\left\{\varphi_j^T\right\}_{j=1}^{n_d}$ are the polynomial basis functions of $\mathbb{P}^k(T)$, and $\left\{\mathbf{V}_j^T(t)\right\}_{j=1}^{n_d}$ are the local degrees of freedom vectors associated to $\mathbf{V}_h^T := \mathbf{V}_{h|T}$ with $\mathbf{V}_j^T(t) = \left(\eta_j^T(t), (q_x)_j^T(t), (q_y)_j^T(t)\right)^t$, with $n_d := \frac{(k+1)(k+2)}{2} = \dim \mathbb{P}^k(T)$ is the number of the degrees of freedom. Let also consider a polynomial expansion of the bathymetry parameterization b :

$$b_h(\mathbf{x}) = \sum_{j=1}^{n_d} b_j^T \varphi_j^T(\mathbf{x}), \quad \forall \mathbf{x} \in T. \tag{4.4}$$

We replace the exact solution $\mathbf{V}(x,t)$ by the approximation $\mathbf{V}_h(x,t)$ in order to obtain the discrete

formulation of (4.3) and the test function ϕ by the basis function $\varphi \in \mathbb{P}^k(T)$:

$$\begin{aligned} &\int_T \left(\sum_{j=1}^{n_d} \frac{d}{dt} \mathbf{V}_j^T(t) \varphi_j^T(\mathbf{x}) \right) \varphi_l^T(\mathbf{x}) d\mathbf{x} - \int_T \mathbf{G}\left(\mathbf{V}_h^T, b_h^T\right) \cdot \nabla \varphi_l^T(\mathbf{x}) d\mathbf{x} + \\ &\int_{\partial T} \mathbf{G}\left(\mathbf{V}_h, b_h\right) \cdot \vec{\mathbf{n}}_{\partial T} \ \varphi_l^T(s) ds = \int_T \mathbf{B}\left(\mathbf{V}_h^T, \nabla b_h^T\right) \varphi_l^T(\mathbf{x}) d\mathbf{x}, \quad 1 \leq l \leq n_d. \end{aligned} \tag{4.5}$$

and the semi-discrete DG formulation of (4.3) writes:

$$\begin{aligned} &\sum_{j=1}^{n_d} \frac{d}{dt} \mathbf{V}_j^T(t) \int_T \varphi_j^T(\mathbf{x}) \varphi_l^T(\mathbf{x}) d\mathbf{x} - \int_T \mathbf{G}\left(\mathbf{V}_h^T, b_h^T\right) \cdot \nabla \varphi_l^T(\mathbf{x}) d\mathbf{x} + \\ &\sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \mathbf{G}^\star_{ij(k)} \varphi_l^T(s) ds = \int_T \mathbf{B}\left(\mathbf{V}_h^T, \nabla b_h^T\right) \varphi_l^T(\mathbf{x}) d\mathbf{x}, \quad 1 \leq l \leq n_d. \end{aligned}$$

Noting that we have :

$$\mathbf{G}^\star_{ij(k)} = \mathbf{G}\left(\mathbf{V}_h, b_h\right)|_{\Gamma_{ij(k)}} \cdot \vec{\mathbf{n}}_{ij(k)}.$$

To approximate $\mathbf{G}^\star_{ij(k)}$ on the k-th interface $\Gamma_{ij(k)}$ of the triangle (element) T_i, we may use any consistent numerical flux, like the global LF numerical flux for instance, see Fig. 4.1. We introduce in the following a simple choice for the interfaces numerical fluxes $\mathbf{G}^\star_{ij(k)}$, inspired from the FV well-balanced discretization detailed in § 3.

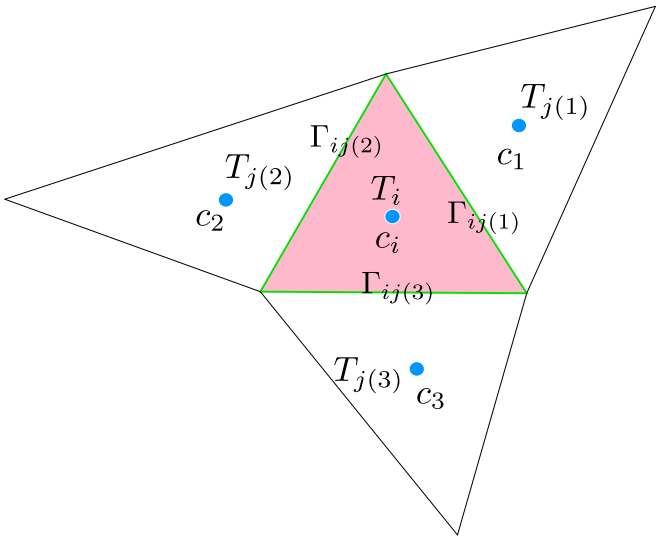

Figure 4.1: Test 2 - T_i element and its first neighbors

4.2 DG well-balancing

Let us define, for a given interface $\Gamma_{ij(k)}$, $\mathbf{V}_k^-$ and $\mathbf{V}_k^+$ respectively the restrictions of $\mathbf{V}_{h|T_i}$ and $\mathbf{V}_{h|T_{j(k)}}$ to $\Gamma_{ij(k)}$ (the interior and exterior traces, with respect to the element T_i). Similarly, b_k^- and b_k^+ stand

for the interior and exterior values of b_h on $\Gamma_{ij(k)}$. For each interface $\Gamma_{ij(k)_{k=1,\dots,3}}$, we follow exactly the same procedure as in the FV frame § 3.5:

$$\tilde{b}_k = \max\left(b_k^-, b_k^+\right), \quad \check{b}_k = \tilde{b}_k - \max\left(0, \tilde{b}_k - \eta_k^-\right),$$

and

$$\begin{aligned} \check{H}_k^- &= \max\left(0, \eta_k^- - \tilde{b}_k\right), \quad \check{H}_k^+ = \max\left(0, \eta_k^+ - \tilde{b}_k\right), \\ \check{\eta}_k^- &= \check{H}_k^- + \check{b}_k, \qquad\qquad \check{\eta}_k^+ = \check{H}_k^+ + \check{b}_k, \end{aligned} \tag{4.6}$$

leading to the new interior and exterior values :

$$\check{\mathbf{V}}_k^- = \left(\check{\eta}_k^-, \frac{\check{H}_k^-}{H_k^-}\mathbf{q}_\mathbf{k}^-\right)^t, \quad \check{\mathbf{V}}_k^+ = \left(\check{\eta}_k^+, \frac{\check{H}_k^+}{H_k^+}\mathbf{q}_\mathbf{k}^+\right)^t.$$

Now we set:

$$\mathbf{G}_{ij(k)}^\star = \mathbf{G}^\star\left(\check{\mathbf{V}}_k^-, \check{\mathbf{V}}_k^+, \check{b}_k, \check{b}_k, \vec{\mathbf{n}}_{ij(k)}\right) + \check{\mathbf{G}}_{ij(k)}, \tag{4.7}$$

as the numerical flux function through the interface between T_i and $T_{j(k)}$, with $\mathbf{G}^\star\left(\mathbf{V}^-, \mathbf{V}^+, b^-, b^+, \vec{\mathbf{n}}\right)$ is the global LF numerical flux:

$$\mathbf{G}^\star\left(\mathbf{V}^-, \mathbf{V}^+, b^-, b^+, \vec{\mathbf{n}}\right) = \frac{1}{2}\left(\mathbf{G}\left(\mathbf{V}^-, b^-\right) \cdot \vec{\mathbf{n}} + \mathbf{G}\left(\mathbf{V}^+, b^+\right) \cdot \vec{\mathbf{n}}\right) - \frac{\sigma}{2}\left(\mathbf{V}^+ - \mathbf{V}^-\right),$$

where

$$\sigma = \max_{i\in\mathbb{Z}} \lambda_i, \tag{4.8}$$

with

$$\lambda_i = \max_{\partial T_i}\left(|\mathbf{u}_i \cdot \vec{\mathbf{n}}_{ij}| + \sqrt{gH_i}\right),$$

$\mathbf{V}_i$ to refer to the restriction of $\mathbf{V}_h$ on the element T_i . This also stands stand for b and each scalar component of $\mathbf{V}_h$. $\check{\mathbf{G}}_{ij(k)}$ is a correction term needed to ensure flux balancing at motionless steady states, defined as follows:

$$\check{\mathbf{G}}_{ij(k)} = \begin{pmatrix} 0 & 0 \\ g\check{\eta}_k^-\left(\check{b}_k - b_k^-\right) & 0 \\ 0 & g\check{\eta}_k^-\left(\check{b}_k - b_k^-\right) \end{pmatrix} \cdot \vec{\mathbf{n}}_{ij(k)}.$$

Remark 36. For the 1d case in § 3, to have the well-balance property, this modified flux strategy (4.7) was applied only for FV scheme and we didn't need to apply it to the DG scheme. Actually, the well-balanced property can be satisfied for DG scheme simply by ensuring the continuity of η_h and b_h globally at initial time (under steady state hypothesis), using a corresponding interpolation method, see Remark 24. Here, for the 2d context we can use a similar strategy. Among the interpolation

points of a mesh element T, one have to choose k interpolation points on each element boundary Γ_{ij} so the interpolated $\mathbb{P}^k$ polynomial be continuous on all elements boundaries, and thus, globally continuous. If one would like to use discontinuous and complex topographies, then he can refer to the modified flux strategy (4.7) for DG scheme. We choose not complicate things and we initialize b_h with a globally continuous polynomial. To ensure that the scheme is indeed well-balanced, and particularly in wet/dry context, we initialize the surface elevation η_h in dry areas by setting $\eta_h = b_h$. Then, water-height positivity is also ensured in dry areas at initial time since $h_h = \eta_h - b_h = 0$ by construction, and in this case we simply use the following classic DG numerical flux without any additional modification strategies .

$$\mathbf{G}^{\star}_{ij(k)} = \mathbf{G}^{\star}\left(\mathbf{V}_k^-, \mathbf{V}_k^+, b_k^-, b_k^-, \vec{\mathbf{n}}_{ij(k)}\right). \tag{4.9}$$

4.3 Sub-partition

For any mesh element (triangle) $T_i \in \mathcal{T}$, we introduce a sub-partition $\mathscr{T}_{T_i}$ into $n_d = \frac{(k+1)(k+2)}{2}$ open disjoint subcells:

$$\overline{T_i} = \bigcup_{m=1}^{n_d} \overline{S}_m^{T_i},$$

where the subcell $S_m^{T_i}$ has a polygonal shape of volume $\left|S_m^{T_i}\right|$, see Fig. 4.2, and Fig. 4.3 for two subdivision examples for a triangular cell T.

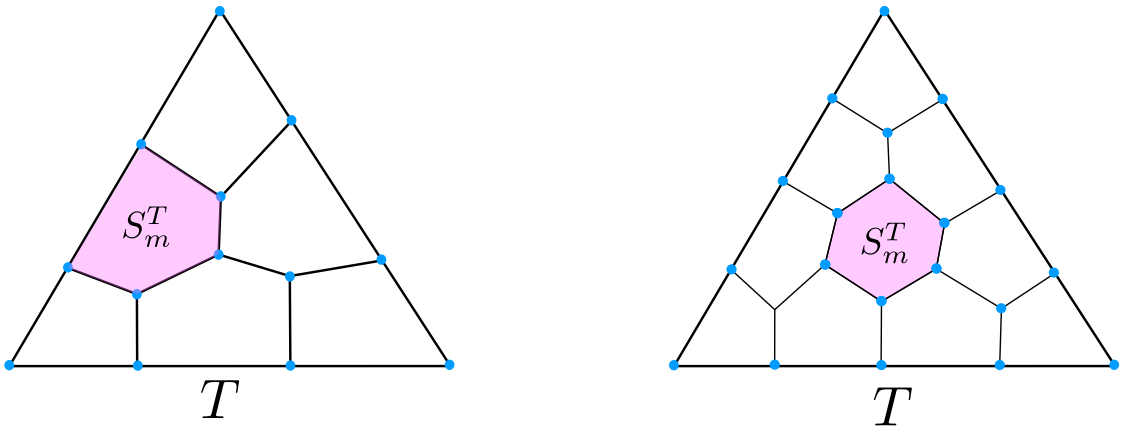

Figure 4.2: Example 1: Partition of a mesh element T in n_d subcells for $\mathbb{P}^2$ (left) and $\mathbb{P}^3$ (right) cases.

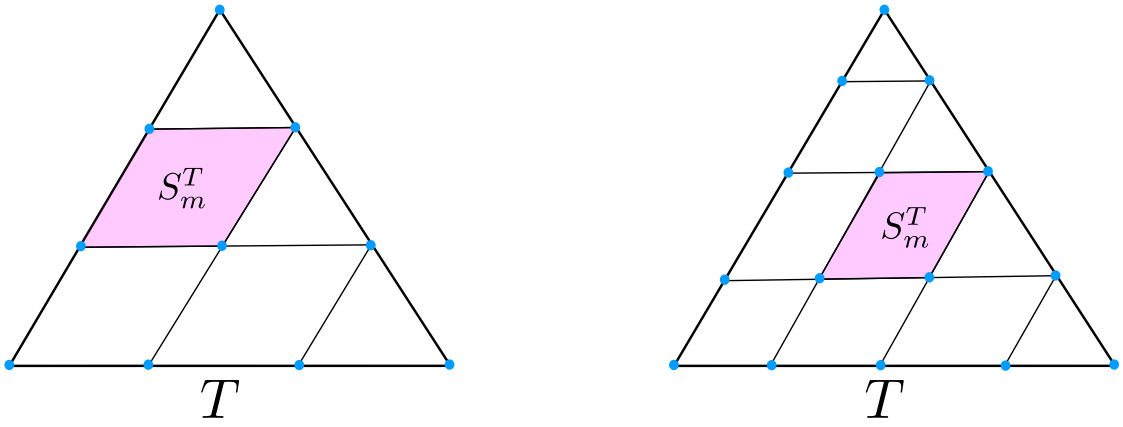

Figure 4.3: Example 2: Partition of a mesh element T in n_d subcells for $\mathbb{P}^2$ (left) and $\mathbb{P}^3$ (right) cases.

For further details concerning the cell subdivision, we refer to [165]. To define the *sub-resolution* basis functions, required in § 3.2, we introduce for a given mesh element $T \in \mathcal{T}$ the following set of *subcell indicator* functions $\left\{\mathbb{1}_m^T,\ \ m \in [\![1,\, n_d]\!]\right\}$, with:

$$\mathbb{1}_m^T(x) = \begin{cases} 1 & \text{if } x \in S_m^T, \\ 0 & \text{if } x \notin S_m^T, \end{cases} \qquad \forall m \in [\![1,\, n_d]\!].$$

Recalling that p_T^k is L^2-orthogonal projector onto $\mathbb{P}^k(T)$, the set of *sub-resolution* basis functions $\left\{\phi_m^T \in \mathbb{P}^k(T),\ \ m \in [\![1,\, n_d]\!]\right\}$ are defined as follows:

$$\phi_m^T = p_T^k(\mathbb{1}_m^T), \quad \forall m \in [\![1,\, n_d]\!], \tag{4.10}$$

$$\int_T \phi_m^T \varphi \, d\mathbf{x} = \int_T \mathbb{1}_m^T \varphi \, d\mathbf{x} = \int_{S_m^T} \varphi \, d\mathbf{x}, \quad \forall m \in [\![1,\, n_d]\!], \quad \forall\ \varphi \in \mathbb{P}^k(T). \tag{4.11}$$

Now, similarly to what we have done in the 1d case, we now seek to reformulation DG scheme as a FV-like scheme on a subgrid. To this end, we follow the step presented in [165].

4.4 DG formulation as a FV-like scheme on a sub-grid

Let us introduce the global L^2-projector $p_{\mathcal{T}}^k$ onto $\mathbb{P}^k(\mathcal{T})$ that gather all the local L^2-projectors p_T^k on each element T. Now, let $\mathbf{G}_h$ and $\mathbf{B}_h$ be the L^2-projections of the flux function and source term onto $\mathbb{P}^k(\mathcal{T})$

$$\mathbf{G}_h = p_{\mathcal{T}}^k(\mathbf{G}(\mathbf{V}_h, b_h)) \ \text{ and } \ \mathbf{B}_h = p_{\mathcal{T}}^k(\mathbf{B}(\mathbf{V}_h, \nabla b_h)).$$

By replacing the flux function and the source term by their L^2-projections in (4.5) we get:

$$\int_T \frac{\partial}{\partial t} \mathbf{V}_h^T \varphi d\mathbf{x} - \int_T \mathbf{G}_h^T \,.\, \nabla\varphi d\mathbf{x} + \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \mathbf{G}_{ij(k)}^{\star} \varphi(s) ds = \int_T \mathbf{B}_h^T \varphi d\mathbf{x}, \quad \forall \varphi \in \mathbb{P}^k(\mathcal{T}). \tag{4.12}$$

Equivalently, using an integration by parts leads to:

$$\int_T \frac{\partial}{\partial t}\mathbf{V}_h^T \varphi d\mathbf{x} + \int_T \nabla \cdot \mathbf{G}_h^T \varphi d\mathbf{x} - \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij(k)} - \mathbf{G}^{\star}_{ij(k)}\right) \varphi(s) ds = \int_T \mathbf{B}_h^T \varphi d\mathbf{x}, \quad \forall \varphi \in \mathbb{P}^k(\mathcal{T}),$$

(4.13)

which is generally referred to as the strong form of DG scheme. Now, since $\partial_t \mathbf{V}_h^T$, $\nabla \cdot \mathbf{G}_h^T$ and $\mathbf{B}_h^T$ belong to $(\mathbb{P}^k(T))^3$, by substituting ϕ_m^T into (4.13), one gets:

$$\int_{S_m^T} \frac{\partial}{\partial t}\mathbf{V}_h^T d\mathbf{x} = -\int_{S_m^T} \nabla \cdot \mathbf{G}_h^T d\mathbf{x} + \int_{S_m^T} \mathbf{B}_h^T d\mathbf{x} + \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij(k)} - \mathbf{G}^{\star}_{ij(k)}\right) \phi_m^T(s) ds, \qquad \forall m \in [\![1,\, n_d]\!].$$

(4.14)

It immediately follows that:

$$\frac{\partial}{\partial t}\overline{\mathbf{V}}_m^T = -\frac{1}{|S_m^T|}\left(\sum_{k=1}^{n_f^m} \int_{\Gamma^T_{mp(k)}} \mathbf{G}_h^T \cdot \vec{\mathbf{n}} ds - \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij(k)} - \mathbf{G}^{\star}_{ij(k)}\right) \phi_m^T(s) ds\right) + \overline{\mathbf{B}}_m^T, \qquad \forall m \in [\![1,\, n_d]\!],$$

(4.15)

where Γ^T_{mp} denotes the face between subcell S_m^T and its neighbor S_p while n_f^m is the number of subcell S_m^T faces. Let us mention that S_p can either be inside cell T (S_p=S_p^T) or in one of its neighboring cell V (S_p=S_p^V). In (4.15), $\overline{\mathbf{V}}_m^T$ and $\overline{\mathbf{B}}_m^T$ stand respectively for the mean-values of $\mathbf{V}_h$ and $\mathbf{B}_h$ on subcell S_m^T, defined as:

$$\overline{\mathbf{V}}_m^T = \frac{1}{|S_m^T|}\int_{S_m^T} \mathbf{V}_h^T d\mathbf{x} \quad \text{and} \quad \overline{\mathbf{B}}_m^T = \frac{1}{|S_m^T|}\int_{S_m^T} \mathbf{B}_h^T d\mathbf{x}. \tag{4.16}$$

We now introduce the DG reconstructed flux $\widehat{\mathbf{G}}_n$ such that

$$\frac{\partial}{\partial t}\overline{\mathbf{V}}_m^T = -\frac{1}{|S_m^T|}\left(\sum_{k=1}^{n_f^m} \int_{\Gamma^T_{mp(k)}} \widehat{\mathbf{G}}_n \, ds\right) + \overline{\mathbf{B}}_m^T, \qquad \forall m \in [\![1,\, n_d]\!]. \tag{4.17}$$

with

$$\sum_{k=1}^{n_f^m} \int_{\Gamma^T_{mp(k)}} \widehat{\mathbf{G}}_n \, ds = \sum_{k=1}^{n_f^m} \int_{\Gamma^T_{mp(k)}} \mathbf{G}_h^T \cdot \vec{\mathbf{n}} ds - \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij(k)} - \mathbf{G}^{\star}_{ij(k)}\right) \phi_m^T(s) ds \tag{4.18}$$

or equivalently

$$\int_{\partial S_m^T} \widehat{\mathbf{G}}_n \, ds = \int_{\partial S_m^T} \mathbf{G}_h^T \cdot \vec{\mathbf{n}} ds - \int_{\partial T} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij} - \mathbf{G}^{\star}_{ij}\right) \phi_m^T(s) ds. \tag{4.19}$$

To further develop relation (4.19), we impose that on the boundary of cell T the reconstructed flux coincide with the DG numerical flux (same as what we did in the 1d case):

$$\widehat{\mathbf{G}}_{n|\partial T} = \mathbf{G}^{\star}. \tag{4.20}$$

Expression (4.19) then rewrites

$$\int_{\partial S_m^T \backslash \partial T} \widehat{\mathbf{G}}_n \, ds = \int_{\partial S_m^T \backslash \partial T} \mathbf{G}_h^T \cdot \vec{\mathbf{n}} ds - \int_{\partial T} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij} - \mathbf{G}_{ij}^{\star} \right) \widetilde{\phi}_m^T(s) ds, \tag{4.21}$$

where $\widetilde{\phi}_m^T$ reads as follows

$$\widetilde{\phi}_m^T = \begin{cases} \phi_m^T & \text{if } x \in \partial T \backslash \partial S_m^T \\ \phi_m^T - 1 & \text{if } x \in \partial T \cap \partial S_m^T. \end{cases}$$

Similarly to [165], we now make use here of a face integrated version of the high-order DG reconstructed flux. Indeed, for a face Γ_{mp}^T, let $\widehat{\mathbf{G}_{mp}}$ be defined as follows

$$\int_{\Gamma_{mp}^T} \widehat{\mathbf{G}}_n \, ds = \varepsilon_{mp}^T \widehat{\mathbf{G}_{mp}}.$$

Similarly, let $\mathbf{G}_{mp}$ be the face integrated integrated value of the polynomial interior flux

$$\int_{\Gamma_{mp}^T} \mathbf{G}_h^T \cdot \vec{\mathbf{n}} ds = \varepsilon_{mp}^T \mathbf{G}_{mp}.$$

In those definitions, the sign function ε_{mp}^T imposes an orientation for each face Γ_{mp}^T:

$$\varepsilon_{mp}^T = \begin{cases} 1 & \text{if face } \Gamma_{mp}^T \text{ is direct or if } \Gamma_{mp}^T \subset \partial T, \\ -1 & \text{if face } \Gamma_{mp}^T \text{ is indirect,} \\ 0 & \text{if } \ S_p \notin \mathcal{V}_m^T, \end{cases}$$

where $\mathcal{V}_m^T$ denotes the set of the face neighboring subcells of S_m^T, and $\breve{\mathcal{V}}_m^T$ stands for the set containing only the face neighboring subcells of S_m^T inside T. Actually, $\forall \ S_p^T \in \breve{\mathcal{V}}_m^T$, we have $\varepsilon_{pm}^T = -\varepsilon_{mp}^T$. Now, let $\mathbf{G}_T \in \mathbb{R}^{n_f^T}$ be the vector containing all the interior faces fluxes, while $\widehat{\mathbf{G}_T} \in \mathbb{R}^{n_f^T}$ would be the vector containing all the interior faces reconstructed fluxes. By denoting by n_f^T the number of subcells faces inside T, meaning not belonging to ∂T, one finally gets

$$A_T \widehat{\mathbf{G}_T} = A_T \mathbf{G}_T - R_T$$

where $A_T \in \mathcal{M}_{n_d \times n_f^T}$, defined as $(A_T)_{mp} = \varepsilon_{mp}^T$, stands for the adjacency matrix, and R_T contains the boundary contribution as

$$(R_T)_m = \int_{\partial T} \left(\mathbf{G}_h^T \cdot \vec{\mathbf{n}}_{ij} - \mathbf{G}_{ij}^{\star} \right) \widetilde{\phi}_m^T(s) ds.$$

Finally, by means of the graph Laplacian technique employed in [1, 165], we are able to solve such system and express explicitly the reconstructed flux $\widehat{\mathbf{G}_T}$ through the interior flux and a boundary correction term. We note by L_T the Laplacian matrix of the interior subgrid graph $L_T = A_T {A_T}^t$,

and by $\mathcal{L}_T^{-1}$ the inverse of L_T on the orthogonal of its kernel . For any $\lambda \neq 0$, this generalized inverse writes:

$$\mathcal{L}_T^{-1} = (L_T + \lambda\Pi)^{-1} - \frac{1}{\lambda}\Pi$$

with $\Pi = \frac{1}{n_d}(1 \otimes 1) \in \mathcal{M}_{n_d}$. We are now able to exhibit the following definition of the reconstructed flux

$$\widehat{\mathbf{G}_T} = \mathbf{G}_T - A_T^{\mathrm{t}} \mathcal{L}_T^{-1} R_T. \tag{4.22}$$

Theorem 37. DG scheme, expressed in the following equation in cell T through volume and boundary flux contribution

$$\int_T \frac{\partial}{\partial t}\mathbf{V}_h^T \varphi d\mathbf{x} - \int_T \mathbf{G}(\mathbf{V}_h^T, b_h^T) \cdot \nabla\varphi d\mathbf{x} + \sum_{k=1}^{3} \int_{\Gamma_{ij(k)}} \mathbf{G}^{\star}_{ij(k)} \varphi(s) ds = \int_T \mathbf{B}(V_h^T, \nabla b_h^T)\varphi d\mathbf{x}, \quad \forall \varphi \in \mathbb{P}^k(T), \tag{4.23}$$

can be recast into n_d FV-like subcell schemes as

$$\frac{\partial}{\partial t}\overline{\mathbf{V}}_m^T = -\frac{1}{|S_m^T|} \sum_{S_p \in \mathcal{V}_m^T} \varepsilon_{mp}^T \widehat{\mathbf{G}_{mp}} + \overline{\mathbf{B}}_m^T, \qquad \forall m \in [\![1,\, n_d]\!] \tag{4.24}$$

where the FV-like fluxes $\widehat{\mathbf{G}_{mp}}$, referred to as reconstructed fluxes, are defined in equation (4.22) if $S_p \subset T$ and in (4.20) otherwise.

4.5 Corrected scheme

In this section, we show that the reconstructed fluxes may be locally corrected to enforce some required properties. As investigated in § 3 for the 1d NSW equations, lowest-order FV fluxes may be introduced in order to prevent high-order approximations from spurious oscillations in the vicinity of discontinuities or sharp gradients, as well as to ensure the preservation of water-height positivity. Additionally, we introduce the same states reconstructions as the one presented in §4.2 in order to ensure a well-balancing property. For a neighbor subcell $S_p^{\mathcal{V}}$ sharing with S_m^T the face Γ_{mp}^T, we denote by $\overline{V}_m^T$ and $\overline{V}_p^{\mathcal{V}}$ respectively the interior (S_m^T sub-mean-value) and exterior ($S_p^{\mathcal{V}}$ sub-mean-value) mean-values with respect to the Γ_{mp}^T. We proceed exactly as in (4.6) by defining $\breve{\overline{\mathbf{V}}}_m^T$ and $\breve{\overline{\mathbf{V}}}_p^{\mathcal{V}}$ as follows:

$$\tilde{\bar{b}}_{mp} = \max\left(\bar{b}_m^T, \bar{b}_p^{\mathcal{V}}\right), \quad \breve{\bar{b}}_{mp} = \tilde{\bar{b}}_{mp} - \max\left(0, \tilde{\bar{b}}_{mp} - \bar{\eta}_m^T\right),$$

and

$$\breve{\overline{H}}_m^T = \max\left(0, \bar{\eta}_m^T - \tilde{\bar{b}}_{mp}\right), \quad \breve{\overline{H}}_p^{\mathcal{V}} = \max\left(0, \bar{\eta}_p^{\mathcal{V}} - \tilde{\bar{b}}_{mp}\right), \tag{4.25}$$

where $\breve{\bar{\eta}}_m^T$ and $\breve{\bar{\eta}}_p^{\mathcal{V}}$ write:

$$\breve{\bar{\eta}}_m^T = \breve{\overline{H}}_m^T + \breve{\bar{b}}_{mp}, \quad \breve{\bar{\eta}}_p^{\mathcal{V}} = \breve{\overline{H}}_p^{\mathcal{V}} + \breve{\bar{b}}_{mp}. \tag{4.26}$$

Those definitions lead to new interior and exterior values:

$$\breve{\overline{\mathbf{V}}}_m^T = \begin{pmatrix} \breve{\overline{\eta}}_m^T, \dfrac{\breve{\overline{H}}_m^T}{\overline{H}_m^T}\overline{\mathbf{q}}_m^T \end{pmatrix}^t, \quad \breve{\overline{\mathbf{V}}}_p^{\mathcal{V}} = \begin{pmatrix} \breve{\overline{\eta}}_p^{\mathcal{V}}, \dfrac{\breve{\overline{H}}_p^{\mathcal{V}}}{\overline{H}_p^{\mathcal{V}}}\overline{\mathbf{q}}_p^{\mathcal{V}} \end{pmatrix}^t.$$

Finally, we set the FV corrected numerical flux to:

$$\mathbf{G}_{mp} = l_{mp}^T \left(\mathbf{G}^* \left(\breve{\overline{\mathbf{V}}}_m^T, \breve{\overline{\mathbf{V}}}_p^{\mathcal{V}}, \breve{\overline{b}}_{mp}, \breve{\overline{b}}_{mp}, \vec{\mathbf{n}}_{mp} \right) + \check{\mathbf{G}}_{mp} \right), \tag{4.27}$$

where $\check{\mathbf{G}}_{mp}$, defined in the expression right bellow, is a correction term required to ensure flux balancing at motionless steady states:

$$\check{\mathbf{G}}_{mp} = \begin{pmatrix} 0 & 0 \\ \mathrm{g}\breve{\overline{\eta}}_m^T \left(\breve{\overline{b}}_{mp} - b_{mp} \right) & 0 \\ 0 & \mathrm{g}\breve{\overline{\eta}}_m^T \left(\breve{\overline{b}}_{mp} - b_{mp} \right) \end{pmatrix} \cdot \vec{\mathbf{n}}_{mp}. \tag{4.28}$$

In definition (4.28), b_{mp} is the mean-value of the bathymetry in T at the face Γ_{mp}^T. Noting by l_{mp}^T the length of Γ_{mp}^T, the bathymetry term b_{mp} writes:

$$b_{mp} = \frac{1}{l_{mp}^T} \int_{\Gamma_{mp}^T} b_h^T \, ds.$$

By means of such FV corrected numerical flux (4.27), it is possible to modify the reconstructed fluxes $\widehat{\mathbf{G}_{mp}}$ in a robust way, in some particular subcells, where the *uncorrected* DG scheme (4.23) has failed to produce an admissible solution. As we did for the 1D case, we compute the candidate solution with the fully DG high-order scheme (4.23) or (4.24). If all the sub-mean-values $\overline{\mathbf{V}}_m^{T,n+1}$ are admissible, we go further in time. Otherwise, we modify the corresponding reconstructed flux value on the troubled subcells faces through the first-order numerical flux as following

$$\begin{cases} \widetilde{\mathbf{G}_{mp}} = \varepsilon_{mp}^T \mathbf{G}_{mp} & \text{if } S_m^T \text{ or } S_p^{\mathcal{V}} \in \mathcal{V}_m^T \text{ is either marked,} \\ \widetilde{\mathbf{G}_{mp}} = \widehat{\mathbf{G}_{mp}} & \text{otherwise.} \end{cases}$$

Through the corrected reconstructed flux, we recompute the sub-mean-values for tagged subcells and their first neighboring subcells, as depicted in Figure 4.4, through a FV-like scheme as:

$$\partial_t \overline{\mathbf{V}}_m^T = -\frac{1}{|S_m^T|} \sum_{S_p^{\mathcal{V}} \in \mathcal{V}_m^T} \varepsilon_{mp}^T \widetilde{\mathbf{G}_{mp}} + \overline{\mathbf{B}}_m^T. \tag{4.29}$$

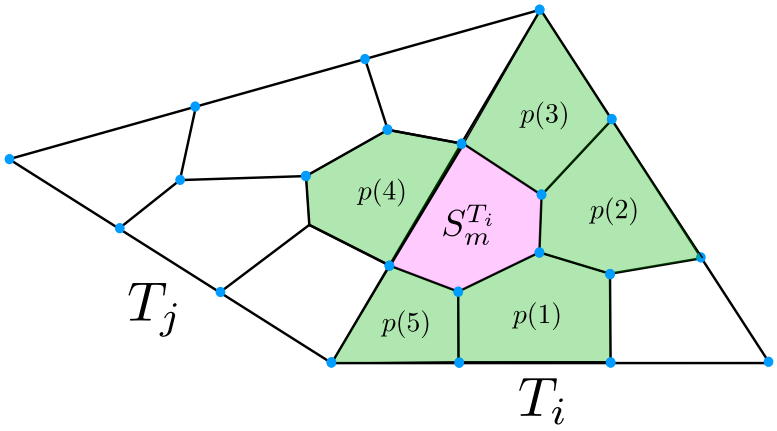

Figure 4.4: $S_m^{T_i}$ subcell and its first neighbors

4.6 Positivity-preserving and well-balancing property

4.6.1 Well-balancing property

Proposition 38. The semi-discrete DG formulation (4.23), combined with the local low-order correction (4.29), together with high-order SSP-RK time marching algorithms, preserves the motionless steady states, provided that the integrals of (4.23) are exactly computed at motionless steady states. Specifically, for any $n \in \mathbb{N}_+^*$:

$$\left(\eta_h^n = \eta^c \text{ and } \mathbf{q}_h^n = 0\right) \implies \left(\eta_h^{n+1} = \eta^c \text{ and } \mathbf{q}_h^{n+1} = 0\right).$$

4.6.2 Positivity-preserving

Proposition 39. Considering the first-order FV sheme on subcell S_m^T, together with a first-order Euler time-marching algorithm:

$$\overline{\mathbf{V}}_m^{T,n+1} = \overline{\mathbf{V}}_m^{T,n} - \frac{\Delta t}{|S_m^T|} \sum_{S_p^{\mathcal{V}} \in \mathcal{V}_m^T} \mathbf{G}_{mp} + \overline{\mathbf{B}}_m^T, \tag{4.30}$$

where $\mathbf{G}_{mp}$ is defined in (4.27). Under the CFL condition:

$$\Delta t^n = \frac{\min\limits_{T \in \mathcal{T},\ S_m^T \in \mathscr{T}_T} |S_m^T|}{\sigma \max\limits_{T \in \mathcal{T}} (\mathbf{p}^T)}, \tag{4.31}$$

with

$$\sigma = \max_{T \in \mathcal{T}} \lambda_T, \quad \lambda_T = \max_{m \in [\![1, n_d]\!],\ \partial S_m^T} \left(\left| \overline{\mathbf{u}}_m^T \cdot \vec{\mathbf{n}} \right| + \sqrt{\mathrm{g} \overline{H}_m^T} \right), \tag{4.32}$$

and $\mathbf{p}^T$ stands for the perimeter of cell T. If $\forall\, T \in \mathcal{T}$, $\forall\, S_m^T \in \mathscr{T}_T$, $\overline{H}_m^{T,n} \geq 0$, then $\forall\, T \in \mathcal{T}$, $\forall\, S_m^T \in \mathscr{T}_T$, $\overline{H}_m^{T,n+1} \geq 0$.

Remark 40. The two previous propositions have been proved for the 1d case in the previous chapter. As the 2d extension is still an ongoing work, complete description and proofs will be presented in details in a future paper. Here, we only present a brief description and some numerical results for the *a posteriori* LSC in the 2d case.

4.7 Numerical validations

4.7.1 Well-balancing property

The initial condition of this test is a flow at rest, with a varying topography, and a dry area. We consider a rectangular domain $[0,2] \times [0,1]$, with the following topography :

$$b(x,y) = \begin{cases} 0.5e^{-(100((x-1.2)^2+150(y-0.7)^2)} & \text{if } x > 0.68, \\ -0.5e^{-\left(100(x-0.45)^2+150(y-0.4)^2\right)} & \text{elsewhere.} \end{cases}$$

The domain is meshed with $n_{\text{el}} = 5000$ elements and we impose at $t = 0$ that:

$$\eta_0 = \max(b, 0.2) \quad \text{and} \quad \mathbf{q}_0 = 0.$$

Results are reported in Fig. 4.5 for the 3rd-order corrected scheme, up to time $t_{max} = 50s$. We emphasize that the steady state is effectively preserved up to the machine accuracy, validating numerically the compatibility of the *a posteriori LSC* method with the well-balancing property. A similar behavior is reported for higher orders of approximations and different grids.

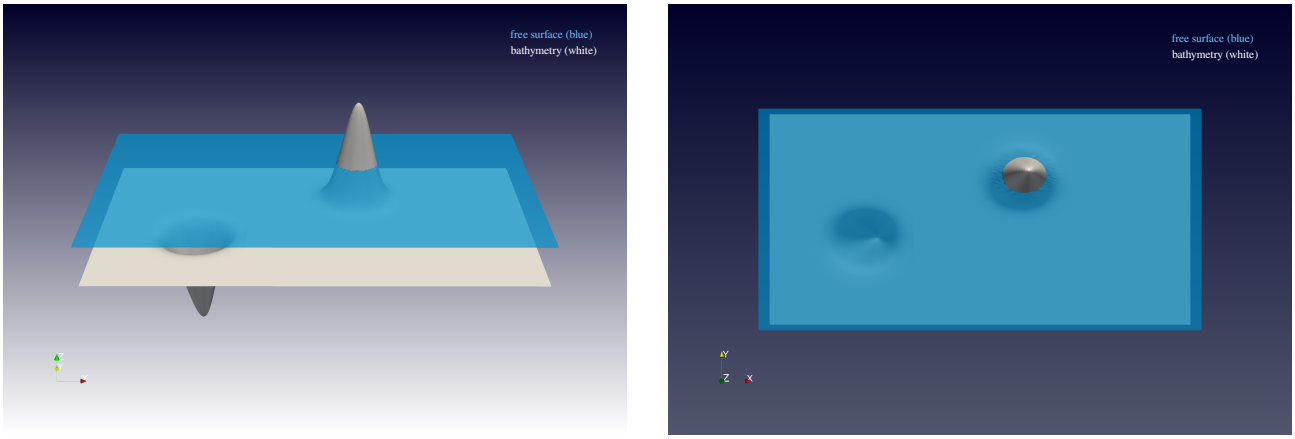

Figure 4.5: Test 16 - Preservation of a motionless steady state - Free surface elevation at $t_{max} = 50s$, with $k = 2$ and $n_{\text{el}} = 5000$.

Next, we slightly modify the initial condition for the water-height in order to have the bump above the water level:

$$\eta_0 = 0.8 \qquad \text{and} \qquad \mathbf{q}_0 = 0.$$

Under the same conditions (for k, t_{max} and n_{el}), numerical results are shown on Fig. 4.6.

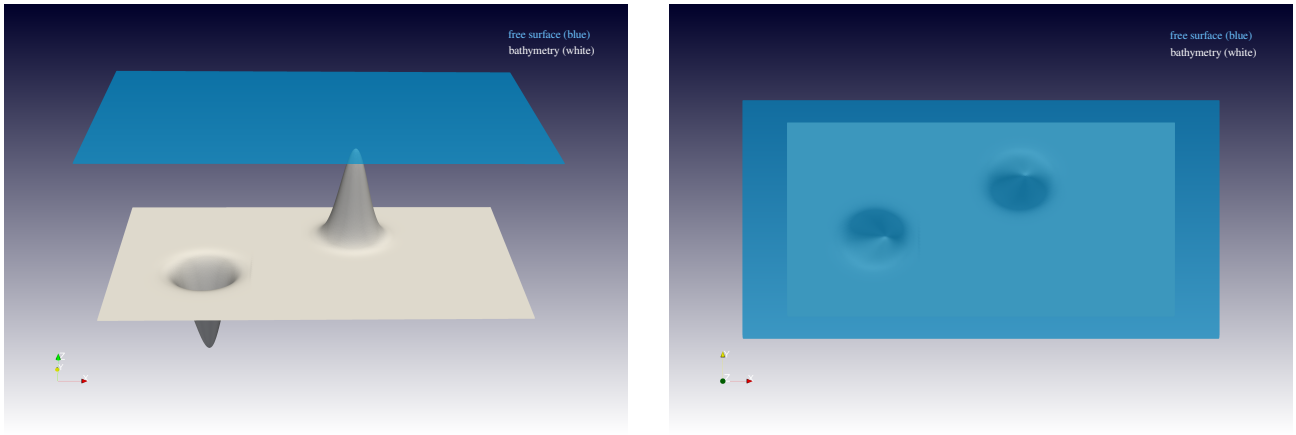

Figure 4.6: Test 17 - Preservation of a motionless steady state - Free surface elevation at $t_{max} = 50s$, with $k = 2$ and $n_{\text{el}} = 5000$.

4.7.2 Dam-break

Dam-break pseudo-1d

In this second test-case, we focus on two pseudo-1d dam-break problems over flat bottoms. We consider a rectangular domain $[0, 1] \times [0, 0.5]$ and the first set of initial conditions is defined as follows:

$$\eta_0(x) = \begin{cases} 1.5 & \text{if} \quad x \leq 0.5, \\ 0.5 & \text{elsewhere}, \end{cases}, \qquad \mathbf{q}_0 = 0, \qquad b = 0.$$

The final time is set to $t = 0.055\,s$. In Fig. 4.7, on a 3600 cells mesh, third-order solution is displayed. The solution has been cleansed from its spurious oscillations, which illustrate very clearly the efficiency of the *a posteriori* LSC method.

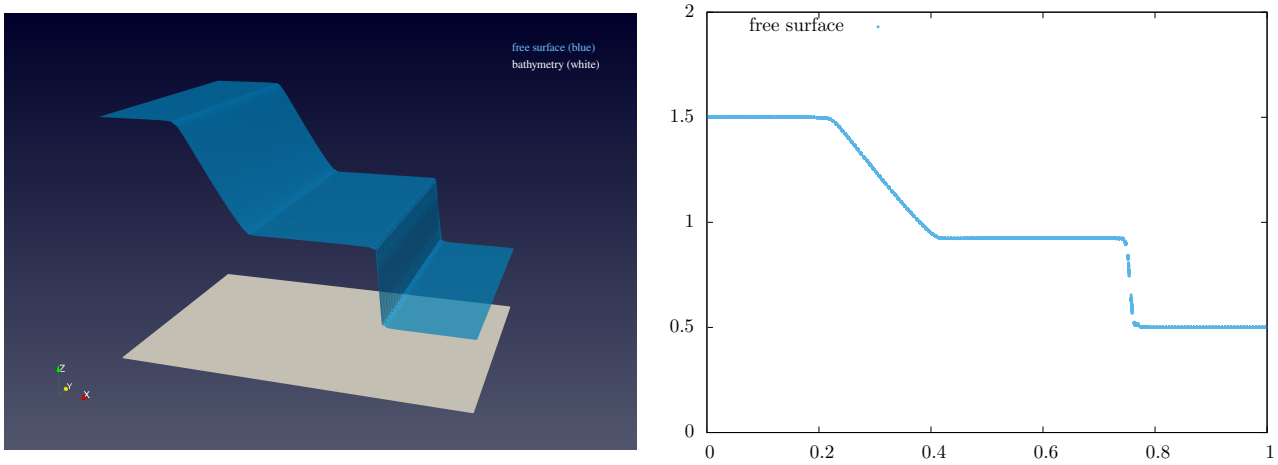

Figure 4.7: Test 18 - Dam break 1D on a wet bottom - Free surface elevation computed at $t = 0.055\,s$ for $k = 2$ and $n_{\text{el}} = 3600$, with the *a posteriori* LSC method.

In a second time, we modify the initial conditions as follows:

$$\eta_0(x) = \begin{cases} 1 & \text{if} \quad x \leq 0.5 \\ 0 & \text{elsewhere} \end{cases}, \quad \mathbf{q}_0(x) = 0.$$

We compute the evolution up to $t = 0.05\,s$, with $k = 2$ and $n_{\text{el}} = 3600$, in order to show the ability of the *a posteriori* LSC method to compute the propagation of a wet/dry front, see Fig. 4.8.

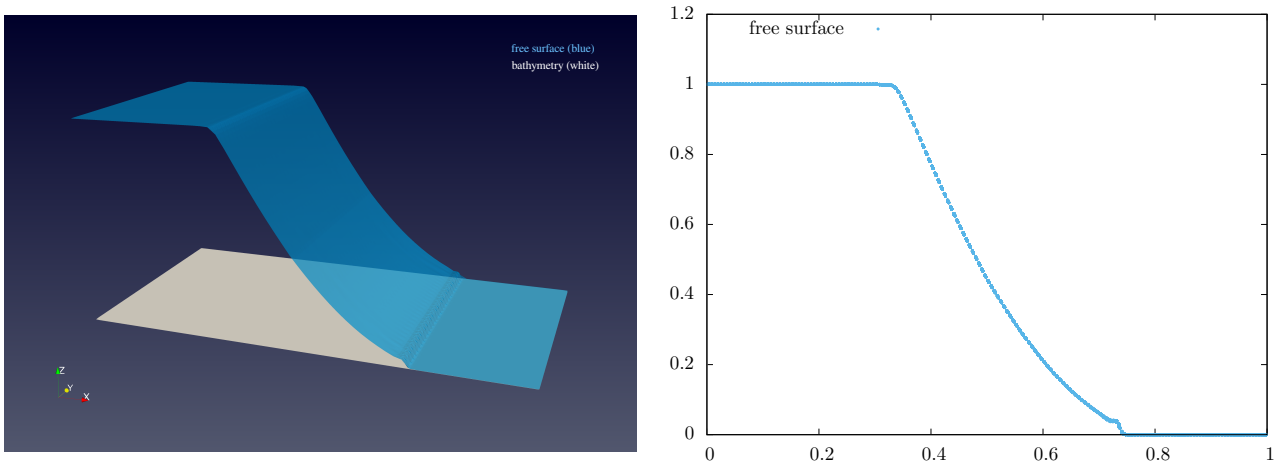

Figure 4.8: Test 19 - Dam break 1D on a dry bottom - Free surface elevation computed at $t = 0.05\,s$ for $k = 2$ and $n_{\text{el}} = 3600$, with the *a posteriori* LSC method.

Dam-break 2d

Now, we focus on two 2d polar dam-break problems over flat bottoms. We consider a polar domain $(R, \theta) \in [0, 1] \times [0, \frac{\pi}{4}]$ and we set the initial conditions as follows:

$$\eta_0(x) = \begin{cases} 1.5 & \text{if} \quad R \leq 0.6, \\ 0.5 & \text{elsewhere}, \end{cases}, \qquad \mathbf{q}_0 = 0, \qquad b = 0.$$

The final time is set to $t = 0.05\,s$. We make use of an unstructured totally anisotropic polar grid made of 2676 triangular cells, as displayed in Figure 4.9. This test-case consists in an expansion wave and a cylindrical diverging shock. If no correction is applied, the DG scheme would produce an oscillatory solution, which may even lead to negative water height for very high-order of accuracy. In Figure 4.10, the 3rd-order numerical solution obtained through our *a posteriori* LSC method is depicted. One can clearly see how the solution exhibits the correct radial wave structure, even in this anisotropic grid case, while still ensuring a non-oscillatory behavior.

The high accuracy of DG scheme is preserved, while ensuring a robust solution, since the *a posteriori* correction is done locally, at the subcell level. Indeed, as illustrated by Figure 4.11 where the troubled subcells are colored red and their first neighbors are colored green, only the subcells where the uncorrected DG method has failed as well as their face neighbors will be recomputed in our *a posteriori* LSC method . The remaining subcells, colored blue in Figure 4.11, do not require any additional treatment, which means that their corresponding mean value is nothing but the one

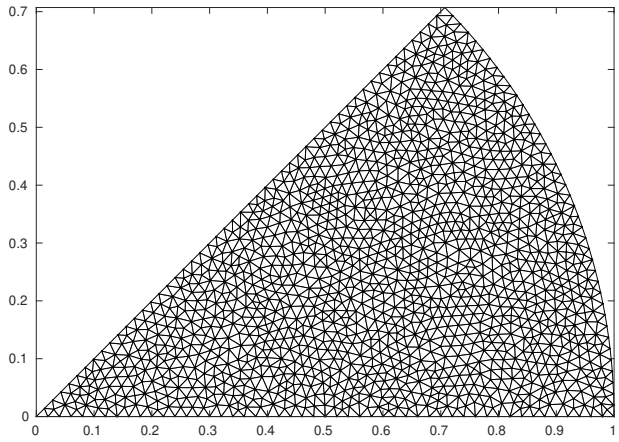

Figure 4.9: Test 20-21 - polar unstructured-grid made of 2676 triangular cells

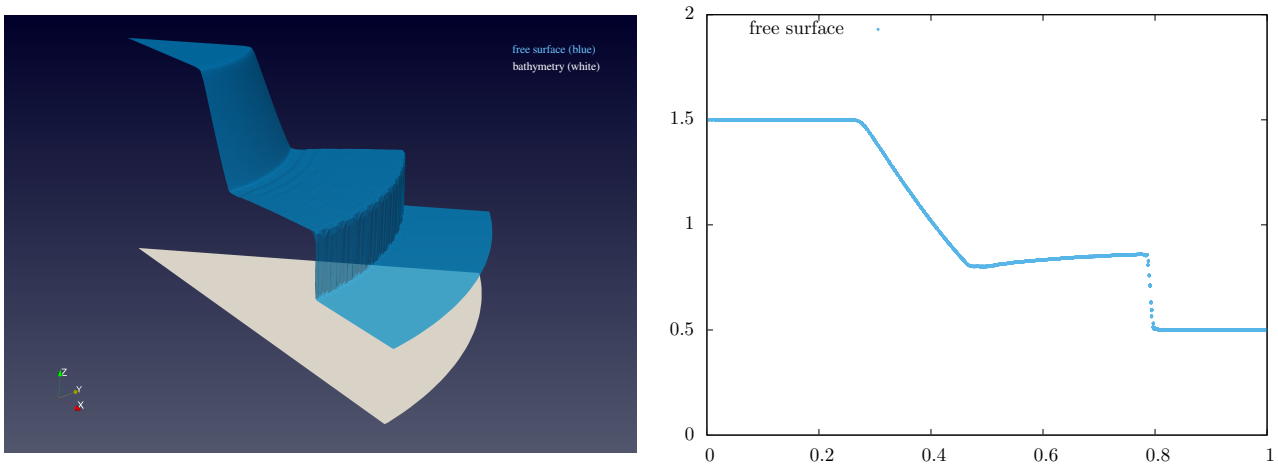

Figure 4.10: Test 20 - Dam break 2D on a wet bottom - Free surface elevation computed at $t = 0.05\,s$ for $k = 2$ and $n_{\text{el}} = 2676$, with the *a posteriori* LSC method.

obtained through the uncorrected DG scheme.

In a second time, we modify the initial conditions as follows to assess the capibility of the *a posteriori* local subcell corrected DG scheme in the case of a 2D cylindrical dam break on a dry bottom:

$$\eta_0(x) = \begin{cases} 1 & \text{if} \quad R \leq 0.6 \\ 0 & \text{elsewhere} \end{cases}, \quad \mathbf{q}_0(x) = 0.$$

We compute the evolution up to $t = 0.045\,s$, by means of the 3rd-order method and the same polar grid as before, see Fig. 4.12. This result demonstrates once more the ability of the proposed method to accurately and robustely compute the propagation of a wet/dry front, as it produced a very high accurate solution while ensuring the positivity of the water-height and avoiding the apparition of spurious oscillations.

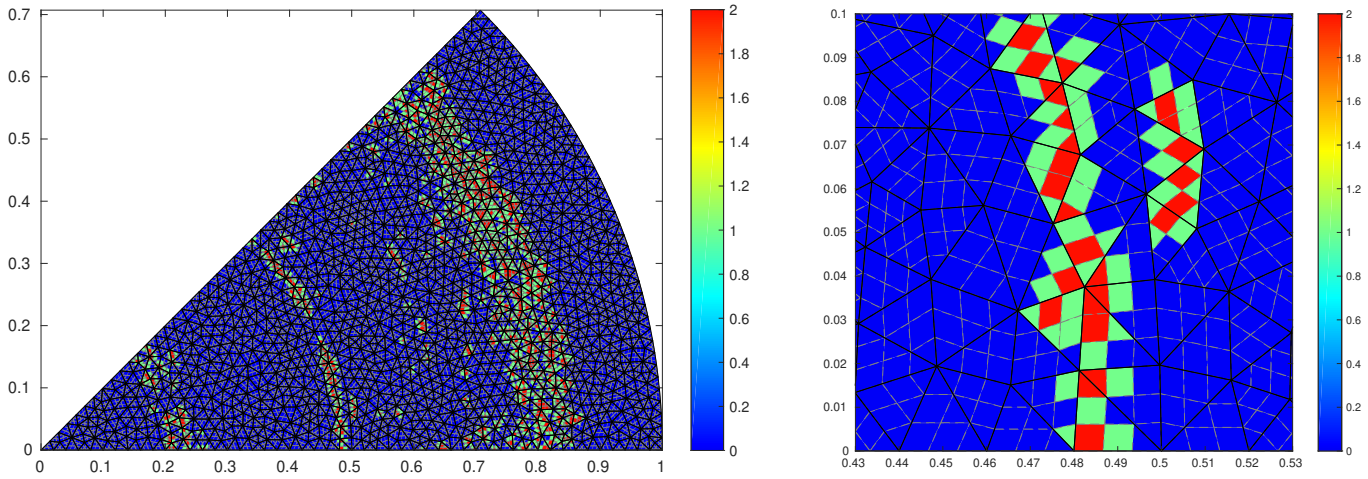

Figure 4.11: Test 20 - Dam break 2D on a wet bottom - computed at $t = 0.05\,s$ for $k = 2$ and $n_{\text{el}} = 2676$, with the *a posteriori* LSC method: flagged-subcells (red), neighboors of flagged-subcells (green), uncorrected-subcells (blue).

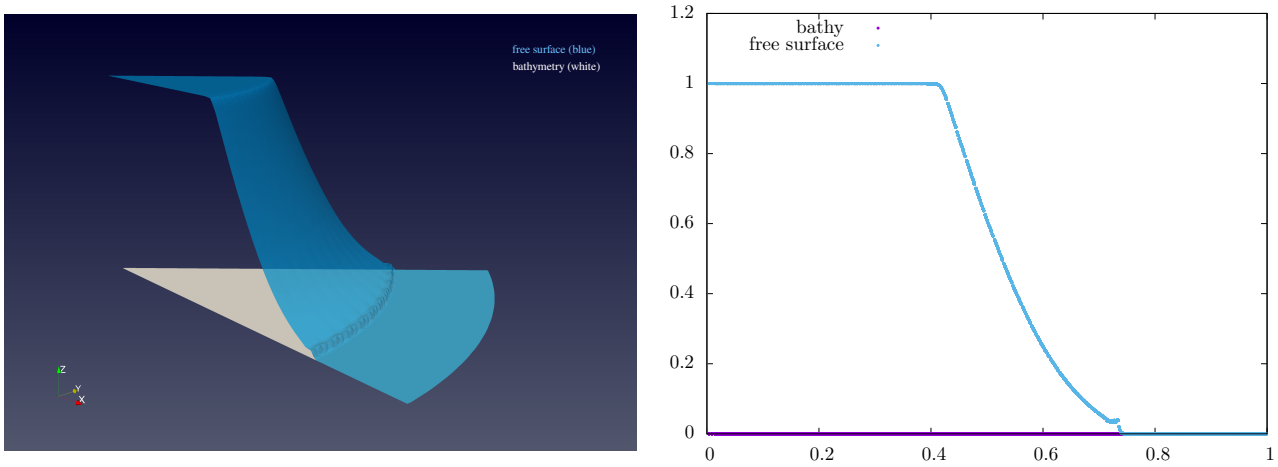

Figure 4.12: Test 21 - Dam break 2D on a dry bottom - Free surface elevation computed at $t = 0.045\,s$ for $k = 2$ and $n_{\text{el}} = 2676$, with the *a posteriori* LSC method.

4.7.3 Rock-wave interaction

We now consider a challenging 2d test-case consisting of the propagation of a solitary wave over a solid rock. We assume a rectangular domain $[5, 25] \times [0, 30]$, with the following topography (rock):

$$b(x, y) = 5e^{-(0.4((x-15)^2+0.2(y-15)^2)}.$$

The domain is meshed with $n_{\text{el}} = 7000$ triangular elements. At the initial time, the solitary wave is defined as follows:

$$\eta_0(x, y) = H_0 + A\,\text{sech}\left(\gamma\left(20y - y_1\right)\right) \qquad \text{and} \qquad \mathbf{q}_0(x, y) = \begin{pmatrix} 0 \\ 0.3\sqrt{g}\left(\eta_0(x, y) - H_0\right) \end{pmatrix},$$

where $\gamma = \sqrt{\dfrac{A}{600H_0}}$ and $y_1 = 180$. The computation is run with $A = 2\,m$ and $H_0 = 2\,m$. In Fig. 4.13, the free surface obtained with the third-order *a posteriori* LSC method is displayed at several times in the range $[0\,s,\ 3.5\,s]$.

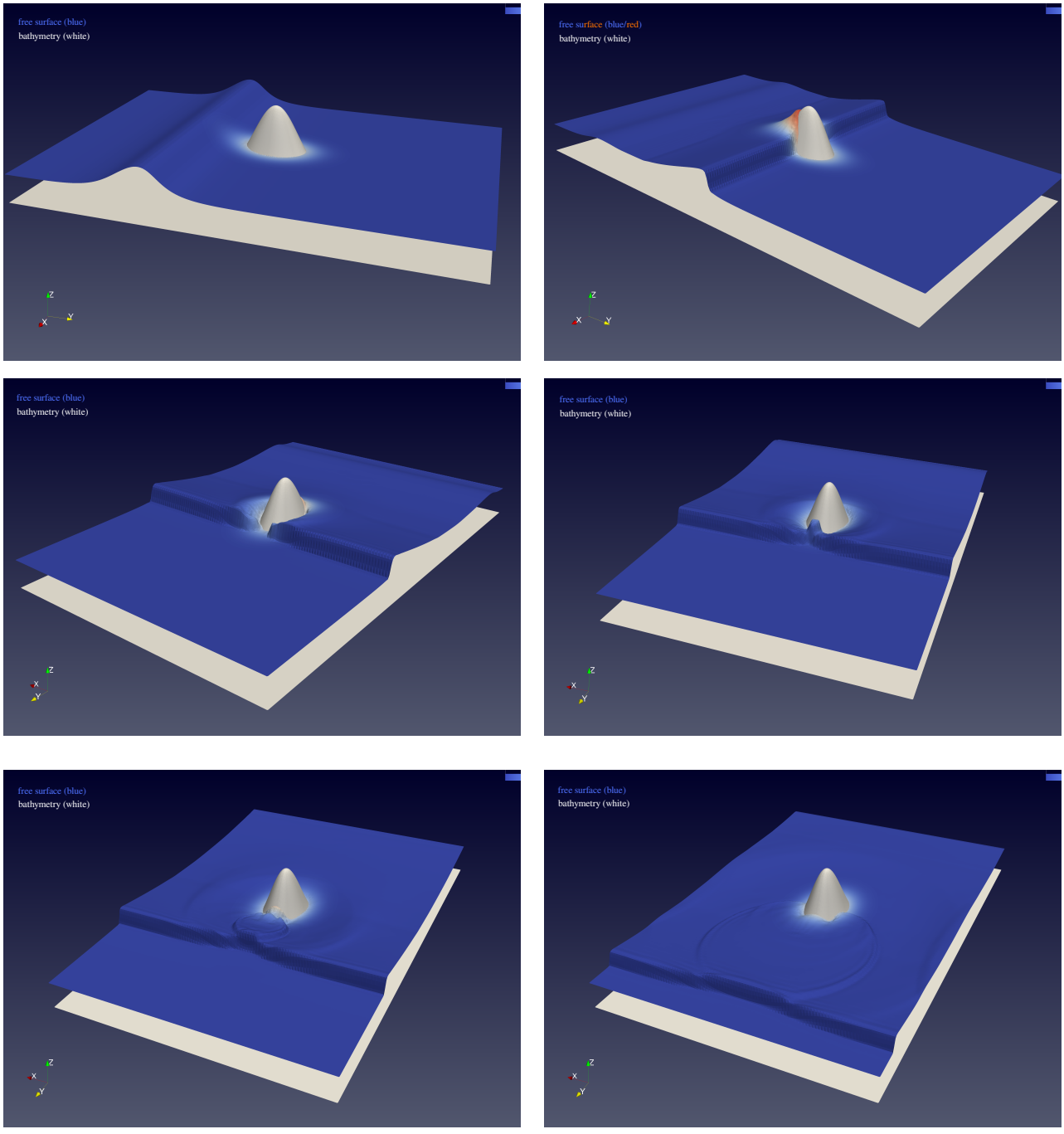

Figure 4.13: Test 22 - Propagation of a solitary wave over a solid rock - Free surface obtained at several times in the range $[0\,s,\ 3.5\,s]$, with $k = 2$ and $n_{\text{el}} = 7000$.

We can see in Figure 4.13 how the solitary wave collides with the rock while its front getting steeper, how the collision induces a cylindrical diverging shock shave around the rock and how those different waves interact with each other. Those results exhibit once more the robustness of the corrected, even in this quite challenging test-case.

4.7.4 Run-up of a solitary wave on a plane beach

The last test-case is devoted to the computation of the run-up of a solitary wave on a constant slope [152]. In this test, a solitary wave traveling from the shoreward is let run-up and run-down on a plane beach, before being fully reflected and evacuated from the computational domain. The topography is made of a constant depth area juxtaposed with a plane sloping beach of constant slope $\alpha = \frac{1}{11}$. The initial condition is defined as follows:

$$\eta_0(x,y) = H_0 + A\,\mathrm{sech}^2\left(\gamma\,(x - x_1)\right) \qquad \text{and} \qquad \mathbf{q}_0(x,y) = \begin{pmatrix} \sqrt{g}\,(\eta_0(x,y) - H_0) \\ 0 \end{pmatrix}.$$

where $\gamma = \sqrt{\dfrac{3A}{4H_0}}$ and $x_1 = 13$. This test is run with $A = 0.1\,m$, $H_0 = 0.3\,m$, $k = 2$, $n_{\text{el}} = 4000$ triangular cells and $t_{max} = 27\,s$. We show on Fig. 4.14 the free surface obtained with the *a posteriori* LSC method at several times in the range $[0\,s,\ 27\,s]$.

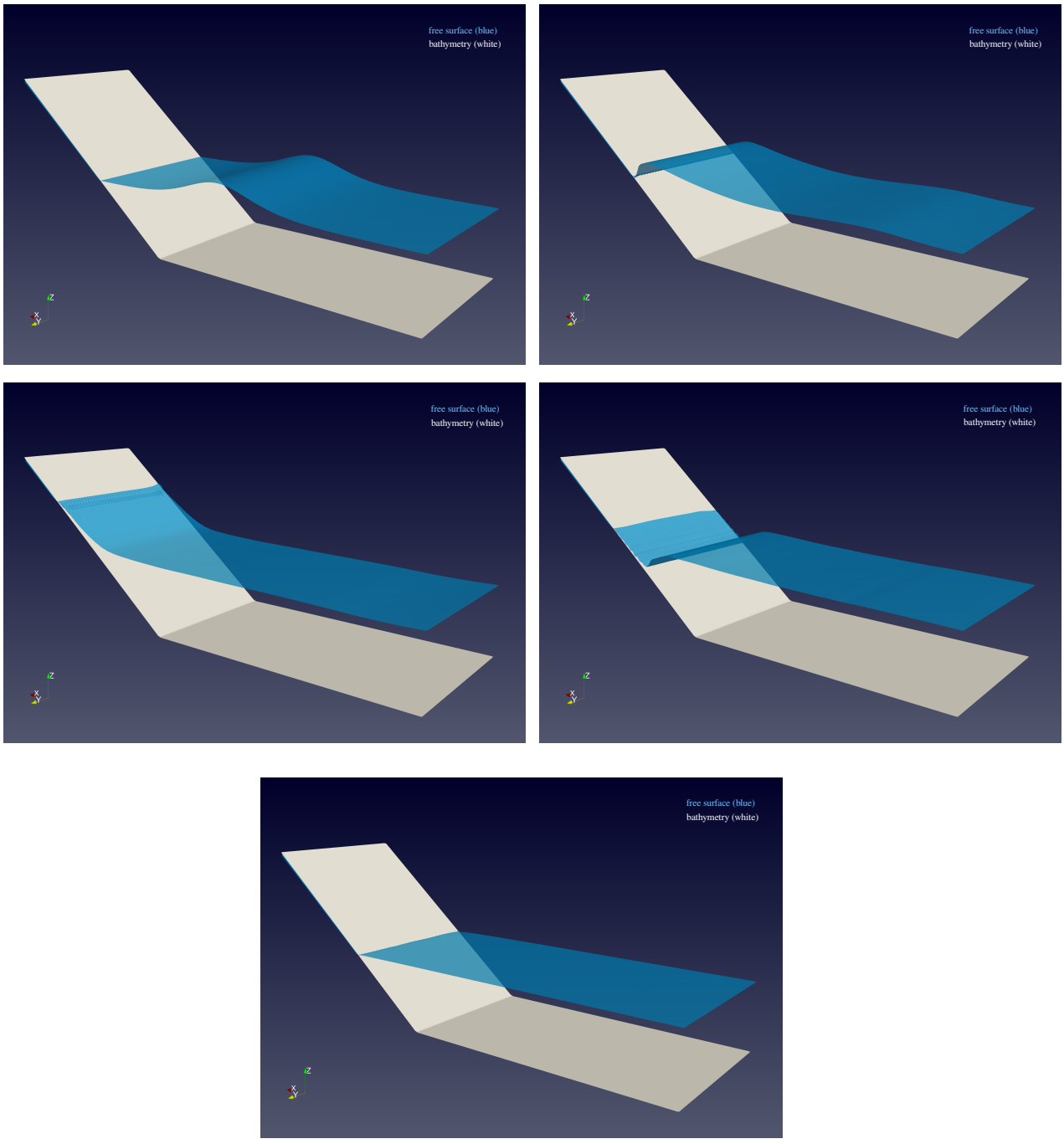

Figure 4.14: Test 23 - Run-up of a solitary wave on a plane beach - Free surface elevation computed for time different values in $[0\,s, t = 27\,s]$ with the 3rd-order *a posteriori* LSC method with 4000 cells.

Part II

Wave interactions with a floating structure in shallow-water

In this second part, we intend to model the interaction between water waves in NSW and a rigid floating structure. We follow the approach proposed in [107], where, under the body, the surface of the fluid coincides with the bottom of the body. As shown in [107], this approach can be used for the NSW approximation. In this work a NSW system with topography source term for the *pre-balance* formulation is considered (5.1)-(5.3). The coupling of our NSW system (5.1)-(5.3) with the floating structure is based on the model proposed in [87], and is achieved using ALE mesh displacement method.

In some situations, the boundary of the domain on which the equations are cast depends on time, where the function χ (boundary node position) is either a known function (prescribed boundary motion) or an unknown function determined by an equation involving the solution $\mathbf{v}$ of the hyperbolic system. In the proposed water-structure interaction model, we come across the second case, typically, χ satisfies $\chi'(t) = \varphi(\mathbf{v})$ for some smooth function φ and the regularity of χ' is the same as the regularity of the solution at the boundary. The free boundary problem that motivates this work is the evolution of the contact line between a floating structure and the water, in the situation where the motion of the waves is assumed to be governed by the hyperbolic NSW equations, and in horizontal dimension $d = 1$. This is the framework that we shall consider here, addressing three cases: the floating body is fixed; the motion of the body is prescribed (and is therefore not influenced by the surface waves); and the body floats freely (and is therefore submitted to the flow motion), according to Newton's laws under the action of the gravitational force and the pressure exerted by the water on the structure. The floating body is allowed to move with heave, surge and pitch motions.

A critical element when we consider sway, surge and pitch motions of a floating structure is to keep track of the contact points $\chi_{\pm}(t)$ ($\chi_{-}(t) < \chi_{+}(t)$) between water and structure as it defines the boundary between the free water surface, the body lateral surface and the air, see Fig. 5.1. The position of the contact points can be accounted using some tracking techniques as the one developed in [72] for congested shallow-water flow and adapt the roof model to a moving body. Here in our work we are inspired by the method developed by [87] which allow us to describe the position of the contact points $\chi_{\pm}(t)$ for translating and rotating structure which lateral walls are not necessarily vertical at the contact points. The approach consists in having water-structure interfaces that are time dependent and move accordingly to the water and body motion, which may cause a displacement of the mesh nodes. This can be treated by setting up the NSW system in a ALE framework in the *exterior* region (flow region), thus leading to what we call a DG-ALE formulation for the NSW system. As for the *interior* region (beneath the floating body) the computation of the water elevation is reduced to a nonlinear algebraic equation to solve, essentially ruled by the object's position and underside's shape. Where the discharge is solution of a nonlinear ordinary differential equation.

We extend the *a posteriori* LSC method to the proposed DG-ALE framework and enforce some nonlinear stability and monotonicity, that are minimal requirements for the high-order approximations of nonlinear flows with floating objects, which ultimately results in what we call a DG-ALE-LSC formulation. We show also that, besides this stabilization procedure we are able to ensure the well-balanced property and both GCL and DGCL properties for the DG-ALE-LSC formulation. These assets are numerically illustrated through an extensive set of manufactured benchmarks validating the water-structure interaction model.

Remark 41. For the sake of clarity and readability, we may recall in this second part some essential notations previously defined in previous chapters.

Chapter 5

Modeling and analysis

In this Chapter, we provide some general floating structure models based on the NSW equations, in the particular case $d = 1$, considering the three different cases: the floating body is fixed; the motion of the body is prescribed; and the body floats freely. The corresponding Initial Boundary Value Problems (IBVP) are also stated. This Chapter, mostly adapted from [87], does not introduce any new result, but the 1d hyperbolic theory for free-boundary problem introduced in [87] is very recent, and for the sake of completeness and consistency in the notations, we choose to completely recall the mathematical results that serve as a theoretical ground to our numerical work.

5.1 Free surface flow in shallow-water

Given a smooth parametrization of the topography $b : \mathbb{R} \to \mathbb{R}$, denoting by H the water-height, $\eta = H + b$ the water elevation, u the horizontal (depth-averaged) velocity and $q = Hu$ the horizontal discharge (see Fig.3.1), the NSW equations may be written as follows:

$$\partial_t \mathbf{v} + \partial_x \mathbf{F}(\mathbf{v}, b) = \mathbf{B}(\mathbf{v}, b'), \tag{5.1}$$

where $\mathbf{v} : \mathbb{R} \times \mathbb{R}_+ \to \Theta$ gathers the flow's conservative variables and is assumed to take values in the convex and open set Θ defined as

$$\Theta = \{(\eta, q) \in \mathbb{R}^2,\ H = \eta - b \geq 0\}, \tag{5.2}$$

$\mathbf{F} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the (nonlinear) flux function and $\mathbf{B} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the topography source term, defined as follows:

$$\mathbf{v} = \begin{pmatrix} \eta \\ q \end{pmatrix},\ \mathbf{F}(\mathbf{v}, b) = \begin{pmatrix} q \\ uq + \frac{1}{2}\mathrm{g}\eta(\eta - 2\,b) \end{pmatrix},\ \ \mathbf{B}(\mathbf{v}, b') = \begin{pmatrix} 0 \\ -\mathrm{g}\,\eta\,b' \end{pmatrix}. \tag{5.3}$$

The benefits of using this *pre-balanced* formulation instead of the classical form are highlighted in [116, 62] and also in chapter 3.

5.2 Shallow-water flow with a floating object

We consider a floating non-deformable object, denoted by $\mathcal{O}_{\mathrm{bj}}$, of mass m_o, inertia coefficient i_o and center of mass $\mathcal{M}_G$, which is partly immersed in an inviscid, incompressible and irrotational

shallow-water flow, under the assumption that there are only two *contact-points* where the water, the air, and the object meet, see Fig. 5.1, and that no wave overhanging occurs. For any given time value $t \geq 0$, the horizontal spatial coordinate of these contact-points are denoted by $\varkappa_{\text{-}}(t)$ and $\varkappa_{\text{+}}(t)$, with $\varkappa_{\text{-}}(t) < \varkappa_{\text{+}}(t)$. Let split the horizontal line into two time-dependent sub-domains, namely the *interior* sub-domain, denoted by $\mathcal{I}(t)$, and the *exterior* sub-domain $\mathcal{E}(t)$, $\mathcal{E}(t)$ and $\mathcal{I}(t)$ being the projections on the horizontal line of the areas where the water surface get in touch with the floating structure and the air:

$$\mathcal{I}(t) := \left]\varkappa_{\text{-}}(t),\, \varkappa_{\text{+}}(t)\right[, \quad \mathcal{E}(t) := \mathcal{E}^{-}(t) \cup \mathcal{E}^{+}(t), \quad \mathcal{E}^{-}(t) := \left]-\infty,\, \varkappa_{\text{-}}(t)\right[, \quad \mathcal{E}^{+}(t) := \left]\varkappa_{\text{+}}(t),\, +\infty\right[, \tag{5.4}$$

and we conveniently gather the contact-points into the set $\partial\mathcal{I}(t) := \{\varkappa_{\text{-}}(t), \varkappa_{\text{+}}(t)\}$. The topography variations are parameterized by a regular function denoted by $b : \mathbb{R} \to \mathbb{R}$, H^i and u^i respectively denote the water-height and the water averaged horizontal velocity in $\mathcal{I}(t)$, H^e and u^e the water-height and the velocity in $\mathcal{E}(t)$ and we set $\eta^i := H^i + b$, $\eta^e := H^e + b$, $q^e := H^e u^e$, $q^i := H^i u^i$ the free-surface elevations and the vertically averaged horizontal discharge respectively in $\mathcal{E}(t)$ and $\mathcal{I}(t)$. The vectors of conservative variables respectively in $\mathcal{E}(t)$ and $\mathcal{I}(t)$ are denoted by $\mathbf{v}^{\mathrm{e}}(x,t)$ and $\mathbf{v}^{\mathrm{i}}(x,t)$ with

$$\mathbf{v}^{\mathrm{e}} = \begin{pmatrix} \eta^e \\ q^e \end{pmatrix}, \qquad \mathbf{v}^{\mathrm{i}} = \begin{pmatrix} \eta^i \\ q^i \end{pmatrix}, \tag{5.5}$$

We also assume the pressure field to be hydrostatic, so that the pressure may be described as follows:

$$\mathrm{p}(x,z,t) := \begin{cases} \mathrm{p}_{\mathrm{atm}} - \rho \mathrm{g}(z - \eta^e(x,t)) & \text{in} \quad \mathcal{E}(t), \\ \underline{\mathrm{p}}^i(x,t) - \rho \mathrm{g}(z - \eta^i(x,t)) & \text{in} \quad \mathcal{I}(t), \end{cases} \tag{5.6}$$

where ρ is the density of the water, $\mathrm{p}_{\mathrm{atm}}$ the atmospheric pressure (at the fluid free-surface) and $\underline{\mathrm{p}}^i(x,t)$ is the inner pressure that applies on the underside of the floating object. Hence, we consider the following flow model:

$$\begin{cases} \mathcal{E}(t) = \left]-\infty,\, \varkappa_{\text{-}}(t)\right[\cup \left]\varkappa_{\text{+}}(t),\, +\infty\right[\quad \text{and} \quad \mathcal{I}(t) = \left]\varkappa_{\text{-}}(t), \varkappa_{\text{+}}(t)\right[, & \text{(5.7a)} \\ \left.\begin{aligned} &\partial_t \eta^e + \partial_x q^e = 0, \\ &\partial_t q^e + \partial_x\Big(u^e q^e + \frac{1}{2}\mathrm{g}\eta^e(\eta^e - 2b)\Big) = -\mathrm{g}\eta^e b', \end{aligned}\right\} & \text{in} \quad \mathcal{E}(t) \quad \text{(5.7b)} \\ \left.\begin{aligned} &\partial_t \eta^i + \partial_x q^i = 0, \\ &\partial_t q^i + \partial_x\Big(u^i q^i + \frac{1}{2}\mathrm{g}(H^i)^2\Big) = -\mathrm{g}H^i b' - \frac{1}{\rho}H^i \partial_x \underline{\mathrm{p}}^i, \end{aligned}\right\} & \text{in} \quad \mathcal{I}(t) \quad \text{(5.7c)} \\ \eta^e = \eta^i, \quad q^e = q^i \quad \text{and} \quad \underline{\mathrm{p}}^i = \mathrm{p}_{\mathrm{atm}} & \text{at} \quad \varkappa_{\pm}(t) \quad \text{(5.7d)} \end{cases}$$

or equivalently

$$\begin{cases} \partial_t \mathbf{v}^{\mathrm{e}} + \partial_x \mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) = \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \quad \text{in} \quad \mathcal{E}(t) = \left]-\infty, \chi_{\text{-}}(t)\right) \cup \left(\chi_{\text{+}}(t), +\infty\right[, & (5.8a) \\ \partial_t \mathbf{v}^{\mathrm{i}} + \partial_x \mathbf{F}(\mathbf{v}^{\mathrm{i}}, b) = \mathbf{B}(\mathbf{v}^{\mathrm{i}}, b') + \mathbf{P}(\mathbf{v}^{\mathrm{i}}, \partial_x \underline{\mathrm{p}}^i) \quad \text{in} \quad \mathcal{I}(t) = \left]\chi_{\text{-}}(t), \chi_{\text{+}}(t)\right[, & (5.8b) \\ \mathbf{v}^{\mathrm{e}} = \mathbf{v}^{\mathrm{i}} \quad \text{and} \quad \underline{\mathrm{p}}^i = \mathrm{p}_{\mathrm{atm}} \quad \text{on} \quad \partial\mathcal{I}(t), & (5.8c) \end{cases}$$

where

$$\mathbf{v}^{\mathrm{e}} \,:\, \mathcal{E}(t) \times [0, T_{\max}] \ni (x,t) \mapsto \mathbf{v}^{\mathrm{e}}(x,t) \in \Theta := \{(\eta, q) \in \mathbb{R}^2,\ H = \eta - b \geq 0\}, \tag{5.9}$$

$$\mathbf{v}^{\mathrm{i}} \,:\, \mathcal{I}(t) \times [0, T_{\max}] \ni (x,t) \mapsto \mathbf{v}^{\mathrm{i}}(x,t) \in \Theta\,, \tag{5.10}$$

respectively gather the flow's main variables in $\mathcal{E}(t)$ and $\mathcal{I}(t)$, $\mathbf{F} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is the (nonlinear) flux function, $\mathbf{B} : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is a topography source term, see (5.3), and $\mathbf{P}(\mathbf{v}^{\mathrm{i}}, \underline{\mathrm{p}}^i) : \Theta \times \mathbb{R} \to \mathbb{R}^2$ is a pressure source term defined as follows:

$$\mathbf{P}(\mathbf{v}^{\mathrm{i}}, \partial_x \underline{\mathrm{p}}^i) := \begin{pmatrix} 0 \\ -\dfrac{1}{\rho} H^i \partial_x \underline{\mathrm{p}}^i \end{pmatrix}. \tag{5.11}$$

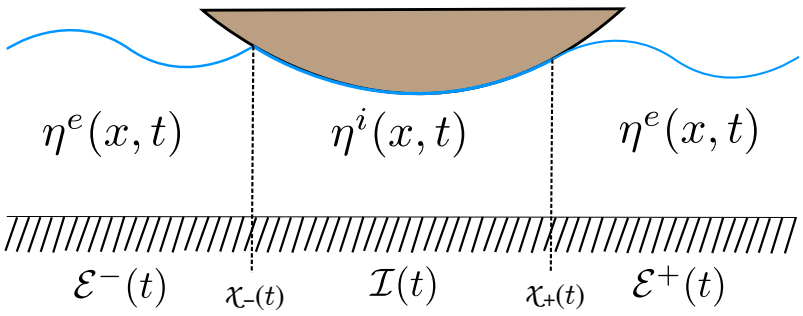

Figure 5.1: shallow-water interacting with a floating object.

For further use, let introduce the unit vectors $\mathbf{e}_{\mathrm{x}} := (1,0)^T$ and $\mathbf{e}_{\mathrm{z}} := (0,1)^T$ in the plane (Oxz), together with the following operators, respectively extracting an average and an oscillating part of any regular enough scalar function $v(\cdot, t)$ defined on $\mathcal{I}(t)$, as follows:

$$\langle\!\langle v \rangle\!\rangle_{\mathcal{I}(t)} := \left(\int_{\mathcal{I}(t)} \frac{1}{H^i}\, dx \right)^{-1} \int_{\mathcal{I}(t)} \frac{v}{H^i}\, dx, \qquad v^\star_{\mathcal{I}(t)} := v - \langle\!\langle v \rangle\!\rangle_{\mathcal{I}(t)}\,, \tag{5.12}$$

and the subscript $\mathcal{I}(t)$ may be forgotten when no confusion is possible.

5.3 A stationary partly immersed object

As we mentioned, we shall consider in this chapter three different cases of water-structure interactions: the floating body is fixed; the motion of the body is prescribed; and the body floats freely.

5.3.1 The model

We start by considering the first case, assuming that the surface-piercing structure is stationary, the profile η^i on the underside of the structure is prescribed by the parametrization of the structure's profile and does not explicitly depend on time (though it implicitly depends on time as $\mathcal{I}(t)$ does):

$$\eta^i(x,t) := \eta_{\text{lid}}(x) \quad \text{on} \quad \mathcal{I}(t) \subset \mathcal{I}_{\text{lid}}, \tag{5.13}$$

where η_{lid} is a given function defined on $\mathcal{I}_{\text{lid}}$, which is the open interval where the parametrization of the partially immersed object's underside is defined, see Fig. 5.2. The continuity equation in $\mathcal{I}(t)$ in (5.8) yields $\partial_x q^i = 0$ and therefore $q^i(x,t) = \underline{q^i}(t)$. Injecting this information into the momentum equation for the interior sub-domain in (5.8), we obtain:

$$\frac{1}{H_i}\frac{\mathrm{d}\,\underline{q^i}}{\mathrm{d}t} + \partial_x\Big(\frac{1}{2}\Big(\frac{\underline{q^i}}{H^i}\Big)^2 + \mathrm{g}H^i\Big) = -\mathrm{g}b' - \frac{1}{\rho}\partial_x \underline{\mathrm{p}^i},$$

so that $\underline{\mathrm{p}^i}$ satisfies the following Boundary Value Problem (BVP):

$$\begin{cases} \partial_x \underline{\mathrm{p}^i} = -\rho\left(\dfrac{1}{H_i}\dfrac{\mathrm{d}\,\underline{q^i}}{\mathrm{d}t} + \partial_x\big(\dfrac{1}{2}\Big(\dfrac{\underline{q^i}}{H^i}\Big)^2 + \mathrm{g}\eta^i\big)\right) & \text{in} \quad \mathcal{I}(t), \qquad (5.14a) \\ \underline{\mathrm{p}^i} = \mathrm{p}_{\text{atm}} & \text{on} \quad \mathcal{E}(t)\cap\mathcal{I}(t). \qquad (5.14b) \end{cases}$$

Integrating (5.14a) on $\mathcal{I}(t)$, we get:

$$\frac{\mathrm{d}\,\underline{q^i}}{\mathrm{d}t} = -\Big(\int_{\mathcal{I}(t)}\frac{1}{H^i}\,\mathrm{d}x\Big)^{-1}\Big[\!\Big[\frac{1}{2}\Big(\frac{\underline{q^i}}{H^i}\Big)^2 + \mathrm{g}\eta^i\Big]\!\Big]_{\mathcal{I}(t)}. \tag{5.15}$$

As a consequence, in the particular case of free surface shallow-water flows with a stationary surface-piercing partially immersed object, model (5.8) may be simplified as follows:

$$\begin{cases} \mathcal{E}(t) = \,]-\infty,\, \varkappa_-(t)[\,\cup\,]\varkappa_+(t),\,+\infty[\quad \text{and} \quad \mathcal{I}(t) = \,]\varkappa_-(t), \varkappa_+(t)[, & (5.16a) \\ \partial_t \mathbf{v}^{\mathrm{e}} + \partial_x \mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) = \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \quad \text{in} \quad \mathcal{E}(t), & (5.16b) \\ \left.\begin{aligned} \eta^i &= \eta_{\text{lid}}, \\ \frac{\mathrm{d}\,\underline{q^i}}{\mathrm{d}t} &= -\Big(\int_{\mathcal{I}(t)}\frac{1}{H^i}\,\mathrm{d}x\Big)^{-1}\Big[\!\Big[\frac{1}{2}\Big(\frac{\underline{q^i}}{H^i}\Big)^2 + \mathrm{g}\eta^i\Big]\!\Big]_{\mathcal{I}(t)}, \end{aligned}\right\} \quad \text{in} \quad \mathcal{I}(t) & (5.16c) \\ \eta^e = \eta^i, \quad q^e = q^i \quad \text{and} \quad \underline{\mathrm{p}^i} = \mathrm{p}_{\text{atm}} \quad \text{at} \quad \varkappa_\pm(t) & (5.16d) \end{cases}$$

Remark 42. Equation (5.15) may be regarded as a solvability condition for problem (5.14). Hence, assuming that $\underline{q^i}'$ and $\mathcal{I}(t) = \big(\varkappa_-(t), \varkappa_+(t)\big)$ satisfying (5.15) are known, one can solve (5.14) for any given $x \in \mathcal{I}(t)$:

$$\underline{\mathrm{p}}^i(x,t) = \mathrm{p}_{\mathrm{atm}} - \rho\left\{\frac{\mathrm{d}\,\underline{q}^i(t)}{\mathrm{dt}}\int_{\varkappa_-(t)}^{x}\frac{\mathrm{d}x'}{H^i(x',t)} + \frac{1}{2}\underline{q}^i(t)^2\Big(\frac{1}{H^i(x,t)^2} - \frac{1}{H^i(\varkappa_-(t),t)^2}\Big) + \mathrm{g}\Big(\eta^i(x,t) - \eta^i(\varkappa_-(t),t)\Big)\right\}. \tag{5.17}$$

5.3.2 IBVP and existence result

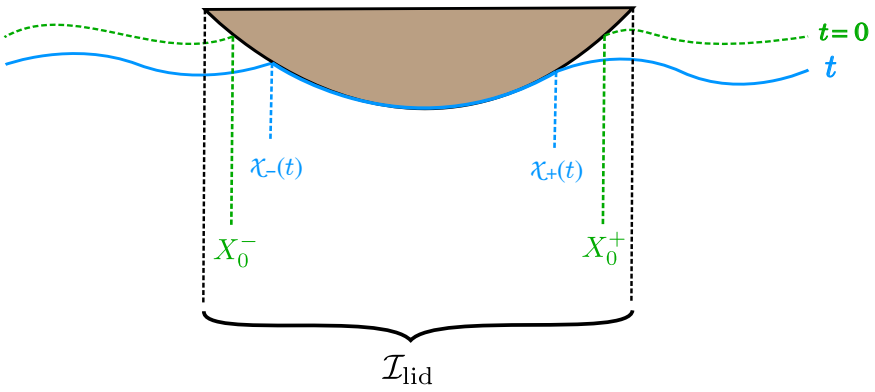

Figure 5.2: Water interacting with a surface-piercing body.

Let consider the initial partition of the computational domain: $\Omega_0 = \mathcal{E}_0 \cup \mathcal{I}_0$ with

$$\mathcal{I}_0 = \left]X_0^-\,,\,X_0^+\right[, \quad \mathcal{E}_0 = \mathcal{E}_0^- \cup \mathcal{E}_0^+, \quad \text{with} \quad \mathcal{E}_0^- = \left]-\infty, X_0^-\right[, \quad \mathcal{E}_0^+ = \left]X_0^+, +\infty\right[, \tag{5.18}$$

where $X_0^\pm$ are the initial locations of the contact points, see Fig. 5.2. Supplementing (5.16) with the following initial data:

$$\begin{cases} \mathbf{v}^{\mathrm{e}}_{|t=0} & := \mathbf{v}^{\mathrm{e}}_0 \in H^s(\mathcal{E}_0)^2, & \text{(5.19a)} \\ (\varkappa_-,\varkappa_+)_{|t=0} & := (X_0^-, X_0^+), & \text{(5.19b)} \\ \mathbf{v}^{\mathrm{i}}_{|t=0} & := (\eta_{\mathrm{lid}}, q_0^i) \in \big(\mathcal{C}^1(\mathcal{I}_{\mathrm{lid}}) \cap W^{s,\infty}(\mathcal{I}_{\mathrm{lid}})\big) \times \mathbb{R}, & \text{(5.19c)} \end{cases}$$

where $H^s(\mathcal{E})$ is the Sobolev space of functions $v \in L^2(\mathcal{E})$ such that their weak derivatives up to order s have a finite L^2-norm, a local well-posedness result is stated in [87] for the particular model of shallow-water flow with a stationary surface-piercing object, under additional assumptions on the data which aim at ensuring that: (i) no dry state occur in the vicinity of the partially immersed objects, (ii) the flow is initially sub-critical at the free boundaries, (iii) the first-order spatial derivative of the free surface is singular at the contact points:

$$(\eta_0^e - \eta_0^i)' \neq 0 \quad \text{on} \quad X_0^\pm, \tag{5.20}$$

and (iv) η_{lid} and its weak derivatives up to order s are uniformly bounded, then there exists a maximum time $T_{\max}$ and a unique solution of (5.16) such that $\mathbf{v}^{\mathrm{e}} \circ \chi \in \mathcal{C}^0\left([0,T_{\max}]; H^s(\mathcal{E}_0)\right) \cap \mathcal{C}^1\left([0,T_{\max}]; H^{s-1}(\mathcal{E}_0)\right)$, $\underline{q}^{i,n} H^{s+1}(0,T_{\max})$, $(\varkappa_-,\varkappa_+) \in (H^s(0,T_{\max}))^2$, where the smooth mapping χ, between the initial domain $\mathcal{E}_0$ and the current one $\mathcal{E}(t)$ is defined in (6.23).

5.4 A moving partly immersed object

We consider now the case of a moving object, and face the following alternative: (i) the motion of the object may be prescribed (and is therefore not influenced by the surface waves), (ii) the motion of the object may be free (and is therefore submitted to the flow motion).

5.4.1 Modeling and geometric description

Motion's laws

In both cases, for any time value, the spatial position of $\mathcal{O}_{\rm bj}$ is completely specified through the knowledge of the spatial coordinates $\mathbf{x}_{\rm G}(t) = (x_{\rm G}(t), z_{\rm G}(t))$ of $\mathcal{M}_G$ (where $x_{\rm G}, z_{\rm G}$ are respectively the horizontal and vertical coordinates), together with the (signed) value of the rotation (pitch) angle $\theta(t)$, see Fig. 5.3. In a similar way, the object's motion may be entirely defined through the knowledge of the velocity $\mathbf{v}_{\rm G}(t) = \big(u_{\rm G}(t), w_{\rm G}(t)\big) = \mathbf{x}_{\rm G}{}'(t)$ and the angular velocity $w(t) := -\theta'(t)$ (so that θ is oriented according to the standard trigonometric convention in the plane (Oxz)) of $\mathcal{M}_G$. For the sake of convenience, let introduce the vectors $\boldsymbol{\mathcal{X}}_{\rm G}$ and $\boldsymbol{\vartheta}_{\rm G}$ defined as follows:

$$\boldsymbol{\mathcal{X}}_{\rm G} := \begin{pmatrix} x_{\rm G} \\ z_{\rm G} \\ -\theta \end{pmatrix}, \qquad \boldsymbol{\vartheta}_{\rm G} := \begin{pmatrix} u_{\rm G} \\ w_{\rm G} \\ w \end{pmatrix}.$$

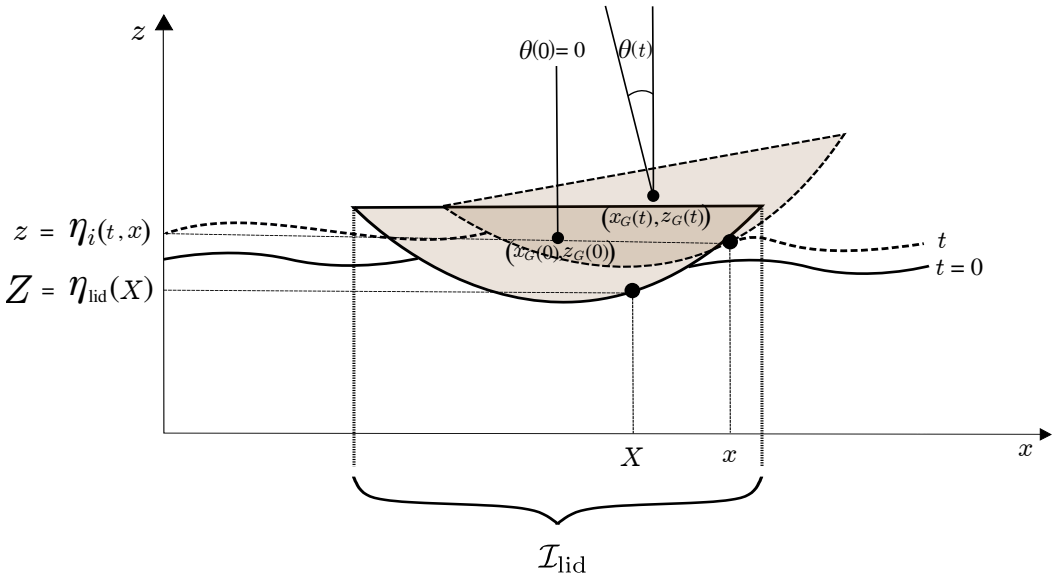

Figure 5.3: Translating and rotating body in water

We investigate the two situations:

♯ 1 $t \mapsto \boldsymbol{\mathcal{X}}_{\rm G}(t)$ and $t \mapsto \boldsymbol{\vartheta}_{\rm G}(t) := \boldsymbol{\mathcal{X}}'_{\rm G}(t)$ belong to the provided data,

♯ 2 $t \mapsto \boldsymbol{\vartheta}_{\rm G}(t)$ and $\mapsto \boldsymbol{\mathcal{X}}_{\rm G}(t) := \int_0^t \boldsymbol{\vartheta}_{\rm G}(s)ds + \boldsymbol{\mathcal{X}}_{\rm G}(0)$ both have to be explicitly computed from the object's response to external forces.

In order to specify such laws, several additional geometrical considerations should be stated:

– at $t = 0$, the initial location $\mathcal{M}_G$ is denoted by $(X_{\rm G}, Z_{\rm G}) := (x_G(0), z_G(0))$ and the initial pitch angle θ_0 is arbitrarily set to zero. We assume that the underside of the floating body is initially parameterized by a smooth function $\eta_{\rm lid}$ defined on an open interval $\mathcal{I}_{\rm lid} \subset \mathbb{R}$, with $\eta_{\rm lid} \in \mathcal{C}^1(\mathcal{I}_{\rm lid}) \cap W^{s,\infty}(\mathcal{I}_{\rm lid})$, $s \geq 1$ ($W^{s,\infty}(I)$ being the Sobolev space of functions which are uniformly bounded on I, together with their weak derivatives up to order s). We observe that:

$$\eta^i(\cdot, 0) = \eta_{\rm lid} \quad \text{on} \quad \mathcal{I}_0 := (X_0^-, X_0^+) \subset \mathcal{I}_{\rm lid},$$

and for any material point located on the underside of the object, which is identified by its horizontal and vertical coordinates (X, Z), we have $Z = \eta_{\rm lid}(X) = \eta^i(X, 0)$. For further use, let also define a normal vector on the underside of the body, with:

$$\mathbf{n}_{\rm lid}(x) := \begin{pmatrix} -\eta'_{\rm lid}(x) \\ 1 \end{pmatrix}.$$

– at any time value $t > 0$, we denote by $\mathbf{x} = (x, z)$, with $z = \eta^i(x, t)$, the coordinates of an arbitrary point belonging to the object's underside, and we note by $\mathbf{r}_{\rm G} := \mathbf{x} - \mathbf{x}_{\rm G}$ the translated coordinate vector of this point with respect to $\mathcal{M}_G$ and by $\mathbf{n}^{\rm i}$ a normal vector on the underside of the body, with:

$$\mathbf{n}^{\rm i}(x, t) := \begin{pmatrix} -\partial_x \eta^i(x, t) \\ 1 \end{pmatrix} = -\partial_x \mathbf{r}_{\rm G}(x, t)^\perp.$$

Object's dynamics

When a free motion is allowed for the floating object, its response to external forces and torques is ruled by Newton's second law for the conservation of linear and angular momentum, which are formulated as follows:

$$\left\{ \begin{array}{l} m_o \partial_t \mathbf{v}_{\rm G} = -m_o g \mathbf{e}_{\rm z} + \int_{\mathcal{I}(t)} \left(\underline{\rm p}^i - {\rm p}_{\rm atm}\right) \mathbf{n}^{\rm i}, \\ i_o \partial_t w = -\int_{\mathcal{I}(t)} \left(\underline{\rm p}^i - {\rm p}_{\rm atm}\right) \mathbf{r}_G^\perp \cdot \mathbf{n}^{\rm i}, \end{array} \right. \tag{5.21}$$

or equivalently,

$$\mathbb{M}_0 \boldsymbol{\vartheta}'_{\rm G} \;=\; -\begin{pmatrix} m_o \, {\rm g} \, \mathbf{e}_{\rm z} \\ 0 \end{pmatrix} - \int_{\mathcal{I}(t)} \left(\underline{\rm p}^i - {\rm p}_{\rm atm}\right) \begin{pmatrix} -\mathbf{n}^{\rm i} \\ \mathbf{r}_{\rm G}^\perp \cdot \mathbf{n}^{\rm i} \end{pmatrix}, \tag{5.22}$$

where the mass-inertia matrix is defined as:

$$\mathbb{M}_0 := \begin{pmatrix} m_o \, {\rm Id}_{2\times 2} & 0 \\ 0 & i_o \end{pmatrix}.$$

For further use, let also introduce the vector $\boldsymbol{\mathcal{T}}_{\rm G}$ defined as follows:

$$\boldsymbol{\mathcal{T}}_{\rm G}(x, t) := \begin{pmatrix} -\mathbf{r}_{\rm G}^\perp(x, t) \\ \frac{1}{2} \left|\mathbf{r}_{\rm G}(x, t)\right|^2 \end{pmatrix},$$

such that the following identities hold:

$$\partial_x \boldsymbol{\mathcal{T}}_{\rm G} = \begin{pmatrix} -\mathbf{n}^{\rm i} \\ \mathbf{r}_{\rm G}^\perp \cdot \mathbf{n}^{\rm i} \end{pmatrix}, \qquad \partial_t \boldsymbol{\mathcal{T}}_{\rm G} = \mathbb{M}_{\rm G} \boldsymbol{\vartheta}_{\rm G}, \tag{5.23}$$

with

$$\mathbb{M}_{\mathrm{G}} := \begin{pmatrix} \mathbf{e}_{\mathrm{x}} \cdot \mathbf{n}_{\mathrm{lid}} & 0 & -\mathbf{r}_{\mathrm{G}}^{\perp} \cdot \mathbf{n}_{\mathrm{lid}} \\ 1 & 0 & 0 \\ -\mathbf{r}_{\mathrm{G}}^{\perp} \cdot \mathbf{n}_{\mathrm{lid}} & 0 & -(\mathbf{e}_{\mathrm{z}} \cdot \mathbf{r}_{\mathrm{G}})(\mathbf{r}_{\mathrm{G}}^{\perp} \cdot \mathbf{n}_{\mathrm{lid}}) \end{pmatrix}.$$

Interior flow description

Then, we reformulate the flow equations in the interior domain. For any point (x, z) of the object's underside, the corresponding initial coordinates (X, Z), with $X \in \mathcal{I}_0$, can be traced back through the following identity:

$$\mathbf{r}_{\mathrm{G}}(x,t) = \begin{pmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{pmatrix} \mathbf{r}_{\mathrm{G}}(X,0), \tag{5.24}$$

and as a consequence, we have for any given $t > 0$ and $x \in \mathcal{I}(t)$:

$$\eta^i(x,t) = z_{\mathrm{G}}(t) + \sin\big(\theta(t)\big)\big(X - X_{\mathrm{G}}\big) + \cos\big(\theta(t)\big)\big(\eta_{\mathrm{lid}}(X) - Z_{\mathrm{G}}\big) =: \widetilde{F}(X, t, z_{\mathrm{G}}, X_{\mathrm{G}}, Z_{\mathrm{G}}, \theta, \eta_{\mathrm{lid}}), \tag{5.25}$$

where X satisfies the following nonlinear algebraic equation:

$$\frac{x - x_{\mathrm{G}}(t) + \sin(\theta(t))(\eta_{\mathrm{lid}}(X) - Z_{\mathrm{G}})}{\cos(\theta(t))} + X_{\mathrm{G}} - X = 0. \tag{5.26}$$

Remark 43. Using (5.24), the equation (5.25) can be written as:

$$\begin{aligned} &\big(\eta^i(x,t) - z_G(t)\big)\cos\theta(t) - (x - x_G(t))\sin\theta(t) + z_G(0) \\ &= \eta_{\mathrm{lid}}\left((x - x_G(t))\cos\theta(t) + \big(\eta^i(x,t) - z_G(t)\big)\sin\theta(t) + x_G(0)\right), \end{aligned} \tag{5.27}$$

which gives an expression of $\eta^i(x,t)$ implicitly in terms of x_G, z_G, θ and η_{lid}.

Under the additional assumptions that $\mathcal{M}_G$ remains close to its initial location, and that the pitch angle is small enough, in the following sense:

$$\forall\, t \in]0, T_{\max}],\ |\theta(t)| \le \theta_{\max},\ \text{with}\quad \theta_{\max} \in (0, \pi/2) \text{ such that } \left\|\eta'_{\mathrm{lid}}\right\|_{\infty} \tan(\theta_{\max}) < 1,$$

it is possible to show that: (i) there is a unique $X \in \mathcal{I}_0$ satisfying (5.26), (ii) the discharge can be expressed as:

$$q^i(x,t) = \boldsymbol{\vartheta}_{\mathrm{G}}(t) \cdot \boldsymbol{\mathcal{T}}_{\mathrm{G}}(x,t) + \underline{q^i}(t), \tag{5.28}$$

where $\underline{q^i}$ is the solution of the following BVP:

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{q^i} = -\left(\langle\!\langle f_1 \rangle\!\rangle_{\mathcal{I}(t)} + \langle\!\langle f_2 \rangle\!\rangle_{\mathcal{I}(t)} + \langle\!\langle f_3 \rangle\!\rangle_{\mathcal{I}(t)}\right), \tag{5.29a}$$

$$\underline{q^i}(0) := \underline{q^i_0}, \tag{5.29b}$$

with the following right-hand sides:

$$\begin{aligned} f_1 &:= \partial_x\big(u^i q^i\big) + \mathrm{g} H^i \partial_x \eta^i, \\ f_2 &:= \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\vartheta}_{\mathrm{G}} \cdot \boldsymbol{\mathcal{T}}_{\mathrm{G}}, \\ f_3 &:= \boldsymbol{\vartheta}_{\mathrm{G}} \cdot \partial_t \boldsymbol{\mathcal{T}}_{\mathrm{G}}. \end{aligned} \tag{5.30}$$

The reader is referred to [87] for the details of this formulation.

Remark 44. Additionally looking at the pressure in $\mathcal{I}(t)$, we observe that $\underline{\mathrm{p}}^i$ satisfies the following BVP:

$$\partial_x \underline{\mathrm{p}}^i = -\frac{\rho}{H^i}\left(\frac{\mathrm{d}}{\mathrm{d}t}\underline{q}^i + f_1 + f_2 + f_3\right) \quad \text{in } \mathcal{I}(t), \tag{5.31a}$$

$$\underline{\mathrm{p}}^i_{|\chi_\pm} := \mathrm{p}_{\mathrm{atm}}. \tag{5.31b}$$

Remark 45. It is known that, for partially immersed object, part of the force and torque applied on the object by the surrounding fluid acts as if the mass-inertia matrix in Newton's laws was modified through the addition of a positive matrix, which is the so-called *added-mass* effect. Hence, (5.22) may be reformulated in order to exhibit the corresponding added-mass, as follows:

$$\Big(\mathbb{M}_0 + \mathbb{M}_{\mathrm{a}}\left(H^i, \boldsymbol{\mathcal{T}}_{\mathrm{G}}\right)\Big)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\vartheta}_{\mathrm{G}} = \begin{pmatrix} -m_{\mathrm{o}}\mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho\int_{\mathcal{I}(t)}\left(f_1^\star + f_3^\star\right)\frac{\boldsymbol{\mathcal{T}}_{\mathrm{G}}^\star}{H^i}, \tag{5.32}$$

where the *added-mass-inertia matrix* $\mathbb{M}_{\mathrm{a}}$ is defined as:

$$\mathbb{M}_{\mathrm{a}}\left(H^i, \boldsymbol{\mathcal{T}}_{\mathrm{G}}\right) := \int_{\mathcal{I}(t)} \frac{\boldsymbol{\mathcal{T}}_{\mathrm{G}}^\star \otimes \boldsymbol{\mathcal{T}}_{\mathrm{G}}^\star}{H^i}. \tag{5.33}$$

see Appendix § F for more details about the calculation of (5.32).

Remark 46. We note that deriving the second equation of the geometric relation (5.24) with respect to t and x leads to the following identity for the time derivative of the interior free-surface:

$$\partial_t \eta_h^i(\cdot,t) = \Big(\mathbf{v}_{\mathrm{G}}(t) - \mathbf{r}_{\mathrm{G}}(\cdot,t)^\perp w(t)\Big)\cdot\mathbf{n}^{\mathrm{i}}(\cdot,t) = -\partial_x\Big(\boldsymbol{\vartheta}_{\mathrm{G}}(t)\cdot\boldsymbol{\mathcal{T}}_{\mathrm{G}}(x,t)\Big). \tag{5.34}$$

This identity is used in the next section to update the location of the contact-points.

Remark 47. The assumption that $\mathcal{M}_G$ *remains close to its initial location* helps to ensure that some singular configurations do not occur. In particular, one has to ensure that the law $t \mapsto \boldsymbol{\mathcal{X}}_{\mathrm{G}}(t)$ is such that the object is never entirely immersed and that for all time value $H^i(\cdot,t) > 0$. To achieve this, we typically require that the object's diameter $\mathscr{d}_o$ is small when compared to the mean water-depth: $\mathscr{d}_o \ll H_0$ (or $\mathscr{d}_o \ll \min H_b$ with $H_b := H_0 - b$ when the topography is varying), in order to ensure that $H^i > 0$ and that η^i remains close to η_{lid}.

5.4.2 IBVP and existence results

Prescribed motion

In the case of a prescribed object's motion, the coupled problem (5.8) may be particularized as follows: find $(\mathbf{v}^{\mathrm{e}}, \mathbf{v}^{\mathrm{i}}, \chi_-, \chi_+)$ such that

$$\begin{cases} \partial_t\mathbf{v}^{\mathrm{e}} + \partial_x\mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) = \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \ \text{ in } \ \mathcal{E}(t) = \left]-\infty, \chi_-(t)\right) \cup \left(\chi_+(t), +\infty\right[, & \text{(5.35a)} \\ \left.\begin{array}{l} \eta^i(x,t) = \widetilde{F}\left(X, t, z_{\mathrm{G}}, X_{\mathrm{G}}, Z_{\mathrm{G}}, \theta, \eta_{\mathrm{lid}}\right) \text{ where } X \text{ solves (5.26)}, \\ q^i(x,t) = \boldsymbol{\vartheta}_{\mathrm{G}}(t)\cdot\boldsymbol{\mathcal{T}}_{\mathrm{G}}(x,t) + \underline{q}^i(t), \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\underline{q}^i \quad = -\left\langle\!\left\langle f_1\right\rangle\!\right\rangle_{\mathcal{I}(t)} - \left\langle\!\left\langle f_2\right\rangle\!\right\rangle_{\mathcal{I}(t)} - \left\langle\!\left\langle f_3\right\rangle\!\right\rangle_{\mathcal{I}(t)}, \end{array}\right\} \text{ in } \mathcal{I}(t) = \left]\chi_-(t), \chi_+(t)\right[, & \text{(5.35b)} \\ \mathbf{v}^{\mathrm{e}} = \mathbf{v}^{\mathrm{i}} \quad \text{and} \quad \underline{\mathrm{p}}^i = \mathrm{p}_{\mathrm{atm}} \quad \text{on} \quad \partial\mathcal{I}(t). & \text{(5.35c)} \end{cases}$$

Supplementing (5.35) with the following initial data, for $s \geq 2$:

$$\mathbf{v}^{\mathrm{e}}_{|t=0} := \mathbf{v}^{\mathrm{e}}_0 \in (H^s(\mathcal{E}_0))^2, \tag{5.36a}$$

$$(\varkappa_{\text{-}}, \varkappa_{\text{+}})_{|t=0} := (X^-_0, X^+_0), \tag{5.36b}$$

$$\mathbf{v}^{\mathrm{i}}_{|t=0} := (\eta_{\mathrm{lid}}, q^i_0) \in \left(\mathcal{C}^1(\mathcal{I}_{\mathrm{lid}}) \cap W^{s,\infty}(\mathcal{I}_{\mathrm{lid}})\right) \times \mathbb{R}, \tag{5.36c}$$

together with the prescribed evolution law:

$$\boldsymbol{\mathcal{X}}_{\mathrm{G}} \in (H^{s+2}(0,T))^3, \tag{5.37a}$$

$$\boldsymbol{\mathcal{X}}_{\mathrm{G}}(0) := (X_{\mathrm{G}}, Z_{\mathrm{G}}, 0), \tag{5.37b}$$

a local well-posedness result is proved in [87]. Specifically, under further assumptions on the data, which can be summarized as: (i) there is no dry state in the vicinity of the floating structure, (ii) the flow is initially sub-critical at $\varkappa_{\pm}$, (iii) the first-order spatial derivative of the free-surface is initially discontinuous at contact points:

$$(\eta^e_0 - \eta^i_0)' \neq 0 \quad \text{on} \quad X^{\pm}_0, \tag{5.38}$$

then there exists a maximum time $T_{\max} \leq T$ and a solution of (5.35)-(5.36)-(5.37) such that $\mathbf{v}^{\mathrm{e}} \circ \chi \in \mathcal{C}^0([0,T_{\max}]; H^s(\mathcal{E}_0)) \cap \mathcal{C}^1([0,T_{\max}]; H^{s-1}(\mathcal{E}_0))$, $\underline{q^i} \in H^{s+1}(0,T_{\max})$, $(\varkappa_{\text{-}}, \varkappa_{\text{+}}) \in (H^s(0,T_{\max}))^2$ and χ is a smooth diffeomorphism defined from the initial exterior domain towards the current one at time t, defined later in (6.23).

Free motion

When an object's free motion is allowed, (5.35) has to be supplemented with Newton's law (5.32) for the object's motion, and $\boldsymbol{\mathcal{X}}_{\mathrm{G}}$ is part of the problem's unknowns. The global problem reads as: find $(\mathbf{v}^{\mathrm{e}}, \mathbf{v}^{\mathrm{i}}, \varkappa_{\text{-}}, \varkappa_{\text{+}}, \boldsymbol{\mathcal{X}}_{\mathrm{G}})$ such that:

$$\begin{cases} \partial_t \mathbf{v}^{\mathrm{e}} + \partial_x \mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) = \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \ \text{ in } \ \mathcal{E}(t) = \left]-\infty, \varkappa_{\text{-}}(t)\right) \cup \left(\varkappa_{\text{+}}(t), +\infty\right[, & (5.39a) \\ \left.\begin{array}{l} \eta^i(x,t) = \widetilde{F}\left(X, t, z_{\mathrm{G}}, X_{\mathrm{G}}, Z_{\mathrm{G}}, \theta, \eta_{\mathrm{lid}}\right) \text{ where } X \text{ solves (5.26)}, \\ q^i(x,t) = \boldsymbol{\vartheta}_{\mathrm{G}}(t) \cdot \boldsymbol{\mathcal{T}}_{\mathrm{G}}(x,t) + \underline{q^i}(t), \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\underline{q^i} \quad = -\left\langle\!\left\langle f_1 \right\rangle\!\right\rangle_{\mathcal{I}(t)} - \left\langle\!\left\langle f_2 \right\rangle\!\right\rangle_{\mathcal{I}(t)} - \left\langle\!\left\langle f_3 \right\rangle\!\right\rangle_{\mathcal{I}(t)}, \end{array}\right\} \text{ in } \mathcal{I}(t) = \left]\varkappa_{\text{-}}(t), \varkappa_{\text{+}}(t)\right[, & (5.39b) \\ \mathbf{v}^{\mathrm{e}} = \mathbf{v}^{\mathrm{i}} \quad \text{and} \quad \underline{\mathrm{p}^i} = \mathrm{p}_{\mathrm{atm}} \quad \text{on} \quad \partial\mathcal{I}(t), & (5.39c) \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{X}}_{\mathrm{G}} = \boldsymbol{\vartheta}_{\mathrm{G}}, \\ \left(\mathbb{M}_0 + \mathbb{M}_{\mathrm{a}}\left(H^i, \boldsymbol{\mathcal{T}}_{\mathrm{G}}\right)\right)\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\vartheta}_{\mathrm{G}} = \begin{pmatrix} -m_o \mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho \displaystyle\int_{\mathcal{I}(t)} \left(f_1^\star + f_3^\star\right) \frac{\boldsymbol{\mathcal{T}}^\star_{\mathrm{G}}}{H^i}. & (5.39d) \end{cases}$$

Supplementing (5.39) with some initial data as specified in (5.36), together with the initial data for the equations of the object's motion:

$$\boldsymbol{\mathcal{X}}_{\mathrm{G}}(0) := (X_{\mathrm{G}}, Z_{\mathrm{G}}, 0), \tag{5.40a}$$

$$\boldsymbol{\vartheta}_{\mathrm{G}}(0) := (u^0_{\mathrm{G}}, w^0_{\mathrm{G}}, \varpi_0), \tag{5.40b}$$

a local well-posedness result is obtained in [87] under the same assumptions as for the prescribed motion case: there exists a maximum time $T_{\max} \leq T$ and a unique solution of (5.36)-(5.39)-(5.40) such that $\mathbf{v}^{\mathrm{e}} \circ \chi \in \mathcal{C}^0([0, T_{\max}]; H^s(\mathcal{E}_0)) \cap \mathcal{C}^1([0, T_{\max}]; H^{s-1}(\mathcal{E}_0))$, $\underline{q^i} \in H^{s+1}(0, T_{\max})$, $(\varkappa_-, \varkappa_+) \in (H^s(0, T_{\max}))^2$ and $\boldsymbol{\mathcal{X}}_{\mathrm{G}} \in (H^{s+2}(0, T_{\max}))^3$.

Remark 48. In the next section, as for the numerical validations of §6.5, we consider these IBVPs on a bounded computational domain of the form

$$\Omega_t = \mathcal{E}^-(t) \cup \mathcal{I}(t) \cup \mathcal{E}^+(t) = \left]x_{\mathrm{left}},\ \varkappa_-(t)\right[\cup \left]\varkappa_-(t),\ \varkappa_+(t)\right[\cup \left]\varkappa_+(t),\ x_{\mathrm{right}}\right[,$$

so that the exterior domain's boundary is defined as $\partial\Omega = \{x_{\mathrm{left}}, x_{\mathrm{right}}\}$ and (5.16)-(5.35)-(5.39) has to be supplemented both with some initial data of the form (5.19) and with prescribed boundary conditions on $\mathbf{v}^{\mathrm{e}}_{|x_{\mathrm{left}}}$ and/or $\mathbf{v}^{\mathrm{e}}_{|x_{\mathrm{right}}}$ depending on the flow characteristics, see also Remark 57.

Chapter 6

A robust discrete formulation for the floating body problem

6.1 Discrete setting for DG-ALE on mesh elements and FV-ALE on subcells

Computational domain, sub-domains and mesh

We consider an open bounded computational domain $\Omega = \left]x_{\text{left}}, x_{\text{right}}\right[$, with boundary $\partial\Omega = \{x_{\text{left}}, x_{\text{right}}\}$ and for any time value $t \in [0, T_{\max}]$, we introduce a partition $\mathcal{P}_\Omega(t) = \{\mathcal{E}^-(t), \mathcal{I}(t), \mathcal{E}^+(t)\}$ of Ω into disjoint sub-domains, defined through the knowledge of the *contact points* $\varkappa_-(t) < \varkappa_+(t)$ such that $\mathcal{I}(t) = \left]\varkappa_-(t),\, \varkappa_+(t)\right[$ and we set

$$\mathcal{E}(t) := \mathcal{E}^-(t) \cup \mathcal{E}^+(t), \quad \Omega_t := \mathcal{E}(t) \cup \mathcal{I}(t).$$

We consider a conforming partition $\mathscr{T}_h(t) = \left\{\omega_i(t)\right\}_{1 \le i \le n_{\text{el}}}$ of Ω_t into $|\mathscr{T}_h(t)|$ disjoint segments, such that we have $\overline{\Omega_t} = \bigcup_{\omega(t) \in \mathscr{T}_h(t)} \overline{\omega(t)}$. We make the following additional assumptions:

- ♯1 n_{el} does not depend on time,
- ♯2 $\forall t \in [0, T_{\max}],\ \varkappa_-(t) \neq x_{\text{left}}(t)$ and $\varkappa_+(t) \neq x_{\text{right}}(t)$,
- ♯3 $\mathscr{T}_h(t)$ is compatible with $\mathcal{P}_\Omega(t)$: each mesh element $\omega(t) \in \mathscr{T}_h(t)$ is a subset of only one set of the partition $\mathcal{P}_\Omega(t)$.

in practice, the two first assumptions are not very influential for our purpose . As a consequence, we can write:

$$\mathscr{T}_h(t) = \mathscr{T}_h^e(t) \cup \mathscr{T}_h^i(t), \quad \text{with} \quad \overline{\mathcal{E}(t)} = \bigcup_{\omega(t) \in \mathscr{T}_h^e(t)} \overline{\omega(t)} \quad \text{and} \quad \overline{\mathcal{I}(t)} = \bigcup_{\omega(t) \in \mathscr{T}_h^i(t)} \overline{\omega(t)},$$

where $\mathscr{T}_h^e(t)$ and $\mathscr{T}_h^i(t)$ are respective partitions of the sub-domains $\mathcal{E}(t)$ and $\mathcal{I}(t)$, and at any time $t \in [0, T_{\max}]$, the contact points $\varkappa_-(t), \varkappa_+(t)$ are uniquely identified with some mesh interfaces. For some specified mesh element $\omega_i(t) \in \mathscr{T}_h(t)$, we note $\omega_i(t) := \left]x_{i-\frac{1}{2}}(t),\, x_{i+\frac{1}{2}}(t)\right[$ (with the convention that $x_{\frac{1}{2}} := x_{\text{left}}$, $x_{n_{\text{el}}+\frac{1}{2}} := x_{\text{right}}$), $x_i(t)$ its barycenter and $\partial\omega_i(t) := \{x_{i-\frac{1}{2}}(t), x_{i+\frac{1}{2}}(t)\}$ its boundary.

Mesh interfaces are collected in the sets $\partial\mathscr{T}_h^e$ and $\partial\mathscr{T}_h^i$, respectively defined as follows:

$$\partial\mathscr{T}_h^e(t) := \{\partial\omega(t),\ \omega(t) \in \mathscr{T}_h^e(t)\}, \quad \partial\mathscr{T}_h^i(t) := \left\{\partial\omega(t),\ \omega(t) \in \mathscr{T}_h^i(t)\right\},$$

such that we have

$$\partial\mathcal{I}(t) = \partial\mathscr{T}_h^e(t) \cap \partial\mathscr{T}_h^i(t), \qquad \partial\mathscr{T}_h(t) := \partial\mathscr{T}_h^e(t) \cup \partial\mathscr{T}_h^i(t) = \left\{x_{i+\frac{1}{2}}(t), 0 \le i \le n_{\text{el}}\right\}. \tag{6.1}$$

DG: approximation spaces, basis functions

For any integer $k \ge 0$ and for any $t \in [0, T_{\max}]$, we consider the broken-polynomials space defined on the exterior domain:

$$\mathbb{P}^k(\mathscr{T}_h^e(t)) := \left\{v(\cdot,t) \in L^2(\mathcal{E}(t)),\ \forall\omega(t) \in \mathscr{T}_h^e(t),\ v_{|\omega(t)} \in \mathbb{P}^k(\omega(t))\right\}.$$

In what follows, piecewise polynomial functions (and, more generally, any discrete counterpart computed from or acting on piecewise polynomial functions) are denoted with a subscript h. **Also, for any $\omega(t) \in \mathscr{T}_h^e(t)$ and $v_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$, we may use in this chapter the alleviated shortcut u_ω instead of $v_h^\omega := v_{h|\omega}$ for an easy reading**, and we also note $\mathbf{P}^k(\mathscr{T}_h^e(t)) := \left(\mathbb{P}^k(\mathscr{T}_h^e(t))\right)^2$.

Remark 49. In this section, some notations defined in § 1.3 are extended to account for the time dependency t accordingly to the mesh-grid displacement.

For any mesh element $\omega(t) \in \mathscr{T}_h^e(t)$ and any integer $k \ge 0$, we consider a basis for $\mathbb{P}^k(\omega(t))$ denoted by

$$\Psi_{\omega(t)} = \left\{\psi_j^\omega(\cdot,t)\right\}_{j\in[\![1,k+1]\!]}.$$

We observe that we have:

$$\forall t \in [0,T_{\max}], \quad \forall\omega(t) \in \mathscr{T}_h^e(t), \quad \forall j \in [\![1,k+1]\!], \quad \operatorname{supp}(\psi_j^\omega(\cdot,t)) \subset \overline{\omega(t)}.$$

A basis for the global space $\mathbb{P}^k(\mathscr{T}_h^e(t))$ is obtained by gathering the local basis functions:

$$\Psi_h^{\text{e}}(t) := \bigtimes_{\omega(t)\in\mathscr{T}_h^e(t)} \Psi_{\omega(t)} = \left\{\left\{\psi_j^\omega(\cdot,t)\right\}_{j\in[\![1,\,k+1]\!]}\right\}_{\omega(t)\in\mathscr{T}_h^e(t)}.$$

Remark 50. In what follows, we choose the set of monomials in the physical space as basis functions, defined as follows:

$$\forall\omega_i(t) \in \mathscr{T}_h^e(t),\ \forall j \in [\![1,\dots,k+1]\!],\ \forall x \in \omega_i(t), \qquad \psi_j^{\omega_i}(x,t) := \left(\frac{x - x_i(t)}{|\omega_i(t)|}\right)^j. \tag{6.2}$$

For any given time value, the degrees of freedom are chosen to be the functionals that map a given discrete unknown belonging to $\mathbb{P}^k(\mathscr{T}_h^e(t))$ to the coefficients of its expansion on the chosen basis functions. Specifically, the degrees of freedom applied to a given function $v_h \in \mathbb{P}^k(\mathscr{T}_h^e(t))$ return the real numbers

$$\left\{\underline{v}_j^\omega\right\}_{j\in[\![1,k+1]\!]}^{\omega\in\mathscr{T}_h^e(t)}, \quad \text{such that} \quad u_\omega = \sum_{j=1}^{k+1} \underline{v}_j^\omega \psi_j^\omega(\cdot,t), \quad \forall\omega \in \mathscr{T}_h^e(t). \tag{6.3}$$

With a little abuse, we refer hereafter to the real numbers (6.3) as the *degrees of freedom* associated with v_h and we note $\underline{v}_\omega \in \mathbb{R}^{k+1}$ the vector that gathers the degrees of freedom associated with v_ω.

In a similar way, the approximating space $\mathbb{P}^k(\mathscr{T}_h^i(t))$ may be defined for the interior domain, and the global approximation space $\mathbb{P}^k(\mathscr{T}_h(t))$ is obtained by gathering the contributions coming from both sub-domains. The product spaces $\mathbf{P}^k(\mathscr{T}_h^i(t))$ and $\mathbf{P}^k(\mathscr{T}_h(t))$ are defined accordingly. Reversely, any function $\phi_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h(t))$ may be regarded as the gathering of its contributions coming respectively from the exterior and interior domains:

$$\phi_h(\cdot,t)_{|\mathcal{E}(t)} = \phi_h^{\mathrm{e}}(\cdot,t), \quad \phi_h(\cdot,t)_{|\mathcal{I}(t)} = \phi_h^{\mathrm{i}}(\cdot,t), \quad \phi_h^{\mathrm{e}}(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t)), \quad \phi_h^{\mathrm{i}}(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^i(t)).$$

Projection, interpolation, averages and jumps

For $\omega(t) \in \mathscr{T}_h^e(t)$, we denote by $p^k_{\omega(t)}$ the L^2-orthogonal projector onto $\mathbb{P}^k(\omega(t))$ and $p^k_{\mathscr{T}_h^e(t)}$ the L^2-orthogonal projector onto $\mathbb{P}^k(\mathscr{T}_h^e(t))$. Similarly, we denote $i^k_{\omega(t)}$ the element nodal interpolator into $\mathbb{P}^k(\omega(t))$. The global interpolator into $\mathbb{P}^k(\mathscr{T}_h^e(t))$, denoted by $i^k_{\mathscr{T}_h^e(t)}$, is obtained by gathering the local interpolating polynomials defined on each elements. Similar projector $p^k_{\mathscr{T}_h^i(t)}$ and interpolator $i^k_{\mathscr{T}_h^i(t)}$ may be defined on $\mathcal{I}(t)$, and globally on Ω_t by gathering the sub-domains contributions.

For any $\phi_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h(t))$ defined on $\omega_i(t) \cup \omega_{i+1}(t)$, we introduce the following *interface-centered* average $\{\!\{\cdot\}\!\}$ and jump $[\![\cdot]\!]$ operators defined as follows:

$$\{\!\{\phi_h(\cdot,t)\}\!\}_{i+\frac{1}{2}} := \frac{1}{2}\left(\phi_{\omega_i}(\cdot,t)_{|x_{i+\frac{1}{2}}} + \phi_{\omega_{i+1}}(\cdot,t)_{|x_{i+\frac{1}{2}}}\right), \qquad [\![\phi_h(\cdot,t)]\!]_{i+\frac{1}{2}} := \phi_{\omega_{i+1}}(\cdot,t)_{|x_{i+\frac{1}{2}}} - \phi_{\omega_i}(\cdot,t)_{|x_{i+\frac{1}{2}}},$$

and this definition should be supplemented with suitable values for the averages and jumps at exterior boundaries, depending on the chosen type of boundary conditions. For any regular-enough scalar-valued function $v(\cdot,t)$ defined on $\omega_i(t)$, and extending the convenient notation $v_{\omega_i(t)}(\cdot) := v(\cdot,t)_{|\omega_i(t)}$, we also introduce the *cell-centered* jump value defined as:

$$[\![v(\cdot,t)]\!]_{\partial\omega_i(t)} := v_{\omega_i(t)|x_{i+\frac{1}{2}}} - v_{\omega_i(t)|x_{i-\frac{1}{2}}},$$

together with the following shortcuts for the *exterior scalar-products* of functions $v, w \in L^2(\mathscr{T}_h^e(t))$ and $\mu, \nu \in L^2(\partial\mathscr{T}_h^e(t))$:

$$\big(v,w\big)_{\mathscr{T}_h^e(t)} := \sum_{\omega(t)\in\mathscr{T}_h^e(t)} \int_{\omega(t)} v(x,t)w(x,t)\mathrm{d}x, \quad \big\langle \mu,\nu \big\rangle_{\partial\mathscr{T}_h^e(t)} := \sum_{\omega(t)\in\mathscr{T}_h^e(t)} [\![\mu\nu]\!]_{\partial\omega(t)},$$

the extension to vector-valued functions being straightforward.

Discrete derivation and integration

In what follows, we need a consistent and accurate discrete counterpart of the first-order derivative, which may be applied to the broken polynomial functions defined above, while accounting for the domain partition $\mathcal{P}_\Omega(t)$ and the jumps of the functions at interfaces. This may be achieved in the current setting by adapting the liftings and discrete gradient of [12] to the sub-domains partition. Let define the element-by-element first-order derivative of a broken-polynomial belonging to $\mathbb{P}^k(\mathscr{T}_h^e(t))$:

$$\partial_x^h \,:\, \mathbb{P}^k(\mathscr{T}_h^e(t)) \ni \phi_h^{\mathrm{e}}(\cdot,t) \mapsto \partial_x^h \phi_h^{\mathrm{e}}(\cdot,t) \in \mathbb{P}^{k-1}\big(\mathscr{T}_h^e(t)\big),$$

such that:

$$(\partial_x^h \phi_h^{\mathrm{e}})_{|\omega(t)} := \partial_x(\phi^{\mathrm{e}}_{\omega(t)}), \quad \forall \omega(t) \in \mathscr{T}_h^e(t).$$

Then, for any $\phi_h^{\mathrm{e}}(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$, we introduce the following global lifting of the jumps on the exterior mesh interfaces $\partial\mathscr{T}_h^e(t)$, defined as follows:

$$\mathcal{R}_h^k([\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]) := \sum_{x_{i+\frac{1}{2}}(t) \in \partial\mathscr{T}_h^e(t)} r^k_{i+\frac{1}{2}}([\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]),$$

where, for all $x_{i+\frac{1}{2}}(t) \in \partial\mathscr{T}_h^e(t)$, the local lifting operator $r^k_{i+\frac{1}{2}}$ applied to the jumps of $\phi_h^{\mathrm{e}}(\cdot,t)$ is defined as the unique solution in $\mathbb{P}^k(\mathscr{T}_h^e(t))$ of the following problem:

$$\big(r^k_{i+\frac{1}{2}}([\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]), \psi_h^{\mathrm{e}}(\cdot,t)\big)_{\mathscr{T}_h^e(t)} = [\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]_{i+\frac{1}{2}} \{\!\{\psi_h^{\mathrm{e}}(\cdot,t)\}\!\}_{i+\frac{1}{2}}, \qquad \forall \psi_h^{\mathrm{e}}(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t)). \tag{6.4}$$

In order to apply the definition (6.4) to the interfaces corresponding to the contact-points (which are boundaries for $\mathcal{E}(t)$), the definitions of the interface-centered jumps and averages on $\varkappa_\pm(t)$ have to be provided. Denoting by $\underline{i}$ and $\overline{i}$ the respective mesh element labels such that $\varkappa_-(t) = \omega_{\underline{i}}(t) \cap \omega_{\underline{i}+1}(t)$ and $\varkappa_+(t) = \omega_{\overline{i}}(t) \cap \omega_{\overline{i}+1}(t)$, we set:

$$\begin{aligned}
[\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]_{\varkappa_-(t)} &:= \phi^{\mathrm{e}}_{\omega_{\underline{i}}}(\cdot,t)_{|\varkappa_-} - \phi^{\mathrm{i}}_{\omega_{\underline{i}+1}}(\cdot,t)_{|\varkappa_-}, \\
[\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]_{\varkappa_+(t)} &:= \phi^{\mathrm{i}}_{\omega_{\overline{i}}}(\cdot,t)_{|\varkappa_+} - \phi^{\mathrm{e}}_{\omega_{\overline{i}+1}}(\cdot,t)_{|\varkappa_+}, \\
\{\!\{\phi_h^{\mathrm{e}}(\cdot,t)\}\!\}_{\varkappa_-(t)} &:= \frac{1}{2}\left(\phi^{\mathrm{e}}_{\omega_{\underline{i}}}(\cdot,t)_{|\varkappa_-} + \phi^{\mathrm{i}}_{\omega_{\underline{i}+1}}(\cdot,t)_{|\varkappa_-}\right), \\
\{\!\{\phi_h^{\mathrm{e}}(\cdot,t)\}\!\}_{\varkappa_+(t)} &:= \frac{1}{2}\left(\phi^{\mathrm{i}}_{\omega_{\overline{i}}}(\cdot,t)_{|\varkappa_+} + \phi^{\mathrm{e}}_{\omega_{\overline{i}+1}}(\cdot,t)_{|\varkappa_+}\right).
\end{aligned}$$

Following [54, Section 2.3], we define the discrete first-order derivative $\mathfrak{G}_h^k : \mathbb{P}^k(\mathscr{T}_h^e(t)) \to \mathbb{P}^k(\mathscr{T}_h^e(t))$ such that, for all $\phi_h^{\mathrm{e}}(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$,

$$\mathfrak{G}_h^k \phi_h^{\mathrm{e}}(\cdot,t) := \partial_x^h \phi_h^{\mathrm{e}}(\cdot,t) - \mathcal{R}_h^k([\![\phi_h^{\mathrm{e}}(\cdot,t)]\!]). \tag{6.5}$$

This operator has better consistency properties than the element-by-element derivative, as it accounts for the jumps of its argument through the second contribution; see [53, Theorem 2.2] for further insight into this point. In a similar way, a discrete derivative acting on functions of $\mathbb{P}^k(\mathscr{T}_h^i(t))$ may be defined, and a discrete gradient globally defined on $\mathbb{P}^k(\mathscr{T}_h(t))$ is obtained by gathering both contributions coming from the exterior and interior domains.

Let also introduce a discrete counterpart for the integration of a regular-enough function $\phi(\cdot,t)$ on $\mathcal{I}(t)$:

$$\mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(t)}[\phi] := \sum_{\omega(t) \in \mathscr{T}_h^i(t)} \sum_{1 \le r \le n_{\mathrm{g}}} \alpha_r^\omega(t) \phi(x_r^\omega(t), t), \tag{6.6}$$

where $(\alpha_r^\omega(t))_{1\le r\le n_{\mathrm{g}}}$ and $(x_r^\omega(t))_{1\le r\le n_{\mathrm{g}}}$ respectively refer to some suitable Gauss quadrature weights and nodes transported onto the transient mesh element $\omega(t) \in \mathscr{T}_h^i(t)$, and the degree n_{g} may be adapted to the polynomial degree of the integrand (or estimated from the regularity of non-polynomial integrands). From (6.6), we deduce a discrete counterpart of the $\mathcal{I}(t)$-averaging operator (5.12), as follows:

$$\langle\!\langle v \rangle\!\rangle_h := \mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(t)}\big[\frac{1}{H_h^i}\big]^{-1} \mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(t)}\big[\frac{v}{H_h^i}\big], \tag{6.7}$$

and we set

$$v_h^\star := v - \langle\!\langle v \rangle\!\rangle_h \,. \tag{6.8}$$

FV on subcells: sub-partitions, sub-resolution basis and sub-mean-values

For any mesh element $\omega_i(t) \in \mathscr{T}_h^e(t)$, we introduce again a sub-partition $\mathscr{T}_{\omega_i(t)}$ into $k+1$ open disjoint subcells:

$$\overline{\omega_i(t)} = \bigcup_{m=1}^{k+1} \overline{S}_m^{\omega_i}(t), \tag{6.9}$$

where the subcell $S_m^{\omega_i}(t) = \left[\widetilde{x}^{\,\omega_i}_{m-\frac{1}{2}}(t), \widetilde{x}^{\,\omega_i}_{m+\frac{1}{2}}(t)\right]$ is of size $\left|S_m^{\omega_i}(t)\right| = \left|\widetilde{x}^{\,\omega_i}_{m+\frac{1}{2}}(t) - \widetilde{x}^{\,\omega_i}_{m-\frac{1}{2}}(t)\right|$, with the convention $\widetilde{x}^{\,\omega_i}_{\frac{1}{2}}(t) = x_{i-\frac{1}{2}}(t)$ and $\widetilde{x}^{\,\omega_i}_{k+\frac{3}{2}}(t) = x_{i+\frac{1}{2}}(t)$, see Fig. 1.1. When considering a sequence of neighboring mesh elements $\omega_{i-1}, \omega_i, \omega_{i+1}$, the convenient convention $S_0^{\omega_i} := S_{k+1}^{\omega_{i-1}}$ and $S_{k+2}^{\omega_i} := S_1^{\omega_{i+1}}$ may be used. For any regular enough function $v(\cdot,t)$ defined on $S_m^\omega(t)$, we use the following shortcut:

$$\left[\!\left[v(\cdot,t)\right]\!\right]_{\partial S_m^\omega(t)} := v(\cdot,t)_{|\widetilde{x}^{\,\omega}_{m+\frac{1}{2}}(t)} - v(\cdot,t)_{|\widetilde{x}^{\,\omega}_{m-\frac{1}{2}}(t)}.$$

For $\omega(t) \in \mathscr{T}_h^e(t)$, we define the *subcell indicator* functions $\left\{\mathbb{1}_m^\omega(\cdot,t),\ m \in [\![1,\, k+1]\!]\right\}$ as follows:

$$\mathbb{1}_m^\omega(x,t) := \begin{cases} 1 & \text{if } x \in S_m^\omega(t), \\ 0 & \text{if } x \notin S_m^\omega(t), \end{cases} \quad \forall m \in [\![1,\, k+1]\!],$$

and the *sub-resolution* basis functions $\left\{\phi_m^{\omega(t)}(\cdot,t) \in \mathbb{P}^k(\omega(t)),\ m \in [\![1,\, k+1]\!]\right\}$ as follows:

$$\phi_m^{\omega(t)}(x(X,t),t) := \widetilde{\phi}_m^{\omega(0)}(X), \quad \forall X \in \omega(0),\ \forall t \geq 0, \tag{6.10}$$

with

$$\widetilde{\phi}_m^{\omega(0)} := p^k_{\omega(0)}(\mathbb{1}_m^{\omega(0)}), \quad \forall m \in [\![1,\, k+1]\!], \tag{6.11}$$

in other words:

$$\int_{\omega(0)} \widetilde{\phi}_m^{\omega(0)} \widetilde{\psi} \mathrm{d}x = \int_{S_m^\omega(0)} \widetilde{\psi} \mathrm{d}x, \quad \forall \widetilde{\psi} \in \mathbb{P}^k\left(\omega(0)\right). \tag{6.12}$$

One can easy show that:

$$\int_{\omega(t)} \phi_m^{\omega(t)} \psi \mathrm{d}x = \int_{S_m^\omega(t)} \psi \mathrm{d}x, \quad \forall \psi \in \mathbb{P}^k\left(\omega(t)\right). \tag{6.13}$$

Actually, for a good choice of the mapping function $x(X,t)$, as in (6.26), such that $x_{|_{\omega(t)}} \in \mathbb{P}^k\left(\omega(0)\right)$, $\psi(x(X,t))$ can be writen as $\widetilde{\psi}(X,t)$ with $\widetilde{\psi}(.,t) \in \mathbb{P}^k\left(\omega(0)\right)$, thus,

$$\int_{\omega(t)} \phi_m^{\omega(t)}(x,t)\psi(x)dx = \int_{\omega(0)} \phi_m^{\omega(t)}(x(X,t),t)\psi(x(X,t))\jmath(X,t)dX$$

$$= \int_{\omega(0)} \widetilde{\phi}_m^{\omega(0)}(X)\widetilde{\psi}(X,t)\jmath(X,t)dX = \int_{S_m^\omega(0)} \widetilde{\psi}(X,t)\jmath(X,t)dX = \int_{S_m^\omega(0)} \psi(x(X,t))\jmath(X,t)dX = \int_{S_m^\omega(t)} \psi(x)dx.$$

See sub-section § 6.1.1 for the definition of $\jmath$.

For any $\omega(t) \in \mathscr{T}_h^e(t)$, we introduce the set of piecewise constant functions on the sub-grid:

$$\mathbb{P}^0(\mathscr{T}_\omega(t)) := \left\{ v(\cdot,t) \in L^2(\omega(t)),\ \ v_{|S_m^\omega(t)} \in \mathbb{P}^0(S_m^\omega(t)),\ \ \forall\, S_m^\omega(t) \in \mathscr{T}_\omega(t) \right\}.$$

For $\omega(t) \in \mathscr{T}_h^e(t)$, and $v_\omega \in \mathbb{P}^k(\omega(t))$, let denote

$$\overline{v}_m^\omega \quad \text{with} \quad m \in [\![1,\, k+1]\!], \tag{6.14}$$

the lowest-order piecewise constant components defined as the mean-values of v_ω on the subcells belonging to the subdivision $\mathscr{T}_\omega(t)$, called *sub-mean-values* in the following, which may be gathered in a vector $\overline{v}_\omega \in \mathbb{R}^{k+1}$. Whenever a sequence of neighboring mesh elements $\omega_{i-1}, \omega_i, \omega_{i+1}$ and associated neighboring approximations is considered, the following convenient convention may be used: $\overline{v}_0^{\omega_i} := \overline{v}_{k+1}^{\omega_{i-1}}$ and $\overline{v}_{k+2}^{\omega_i} := \overline{v}_1^{\omega_{i+1}}$.

Remark 51. We observe that any polynomial function $v_\omega \in \mathbb{P}^k(\omega)$ can be expressed equivalently either in terms of the degrees of freedom $\underline{v}^\omega$, or the sub-means values $\overline{v}_\omega$. Indeed, the degrees of freedom $\left\{\underline{v}_m^\omega\right\}_{m\in[\![1,k+1]\!]}$ are uniquely defined through the sub-mean-values $\left\{\overline{v}_m^\omega\right\}_{m\in[\![1,k+1]\!]}$, and reversely. Considering the local transformation matrix $\Pi_\omega = \left(\pi_{m,p}^\omega\right)_{m,p}$ defined as:

$$\pi_{m,p}^\omega = \frac{1}{|S_m^\omega(t)|}\int_{S_m^\omega(t)} \psi_p^\omega(\cdot,t)\mathrm{dx}, \qquad \forall\,(m,p) \in [\![1,\, k+1]\!]^2, \tag{6.15}$$

the following identities hold:

$$\mathbf{\Pi}_\omega\, \underline{v}_\omega = \overline{v}_\omega \quad \text{and} \quad \mathbf{\Pi}_\omega^{-1}\, \overline{v}_\omega = \underline{v}_\omega.$$

From a practical viewpoint, such transformation matrices $\mathbf{\Pi}_\omega$ are initialized in a preprocessing step and it is therefore computationally inexpensive to locally switch from one representation to another.

Relying on the previous Remark, we introduce the (one-to-one) following projector onto the piecewise constant sub-grid space:

$$\begin{aligned} \pi_{\mathscr{T}_\omega}^k : \mathbb{P}^k(\omega(t)) &\longrightarrow \mathbb{P}^0(\mathscr{T}_\omega(t)) \\ v_\omega &\longmapsto \pi_{\mathscr{T}_\omega}^k(v_\omega) := \overline{v}_\omega. \end{aligned} \tag{6.16}$$

6.1.1 ALE description

In this section, an ALE description for the coupled problems (5.16), (5.35) and (5.39) is introduced. A central aspect of any ALE description is the construction of a continuous and regular coordinate transformation, allowing to recast the equations from the initial (stationary) domain Ω_0 to the current (moving) domain Ω_t:

$$\Omega_0 \times [0, T_{\max}] \ni (X,t) \mapsto x(X,t) \in \Omega_t, \tag{6.17}$$

where X refers to the *reference* coordinate (in the reference frame) and $x := x(X,t)$ the associated *physical* coordinate (in the current frame). Further assuming this mapping to be continuously differentiable with respect to time, piecewise continuously differentiable with respect to X, and denoting by $v_{\mathrm{g}}(x,t)$ the grid's velocity at the physical point $x := x(X,t)$, the following identity holds:

$$v_{\mathrm{g}}(x(X,t),t) = \partial_t x(X,t). \tag{6.18}$$

Now, for the sake of notations, considering any function $v(x,t)$, let introduce $\widetilde{v}(X,t)$ its counterpart defined on the referential frame as

$$v\left(x(X,t),t\right) =: \widetilde{v}(X,t). \tag{6.19}$$

Then, for any arbitrary and regular enough function $v(x,t)$, the fundamental ALE relation between the total time derivative, the Eulerian time derivative and the spatial derivative is

$$\frac{\mathrm{d}}{\mathrm{dt}} v\Big(x(X,t),t\Big) := \Big(\partial_t + v_{\mathrm{g}}\,\partial_x\Big) v(x(X,t),t) =: \partial_t \widetilde{v}(X,t). \tag{6.20}$$

Grid's motion

In order to build such a mapping, for any given time value, the velocity of the contact points may be deduced from the current flow configuration, deriving the free surface continuity condition (5.16d) with respect to time, as follows:

$$\Big(\partial_t + v_{\mathrm{g}}\partial_x\Big)\eta^e = \Big(\partial_t + v_{\mathrm{g}}\partial_x\Big)\eta^i \quad \text{on} \quad \varkappa_{\pm},$$

so that using the identity $\partial_t \eta^e = -\partial_x q^e$:

$$v_{\mathrm{g}|\varkappa_{\pm}} = \left(\frac{\partial_x q^e + \partial_t \eta^i}{\partial_x \eta^e - \partial_x \eta^i}\right)_{\big|\varkappa_{\pm}}. \tag{6.21}$$

Remark 52. For the case of a fixed body, the term $\partial_t \eta^i$ is equal to zero, so the velocity of the contact points may have the form:

$$v_{\mathrm{g}|\varkappa_{\pm}} = \left(\frac{\partial_x q^e}{\partial_x \eta^e - \partial_x \eta^i}\right)_{\big|\varkappa_{\pm}}. \tag{6.22}$$

Having such contact points velocity at hand, let consider the following smooth diffeomorphism $\chi(\cdot,t) : \mathcal{E}_0 \to \mathcal{E}(t)$, defined as:

$$\chi(X,t) := \begin{cases} X + \varphi\left(\frac{X - X_0^-}{\varepsilon}\right)\left(\varkappa_-(t) - X_0^-\right) & \text{for} \quad X \in \mathcal{E}_0^-, \\ X + \varphi\left(\frac{X - X_0^+}{\varepsilon}\right)\left(\varkappa_+(t) - X_0^+\right) & \text{for} \quad X \in \mathcal{E}_0^+, \end{cases} \tag{6.23}$$

where $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$ is a cut-off function satisfying $\varphi(x) = 1$ for $|x| \leq 1$ and $\varepsilon := \varepsilon_0 \ell$ (the reader is referred to Appendix.C C for the practical definition of φ, ε_0 and Remark 54 for additional considerations regarding the value of ℓ). Then, for any moving grid's interface $x_{i+\frac{1}{2}}(t) := x\big(X_{i+\frac{1}{2}}, t\big)$, we enforce the corresponding interface's velocity as follows:

$$v_{\mathrm{g}_{|i+\frac{1}{2}}}(t) := \widetilde{v_{\mathrm{g}}}(X_{i+\frac{1}{2}}, t),$$

with:

$$v_{g_{|i+\frac{1}{2}}}(t) = \widetilde{v}_g(X_{i+\frac{1}{2}},t) := \begin{cases} \partial_t\chi(\cdot,t)_{|X_{i+\frac{1}{2}}} = \begin{cases} \varphi\left(\frac{X_{i+\frac{1}{2}}-X_0^-}{\varepsilon}\right) v_{g|\chi_-} & \text{if} \quad X_{i+\frac{1}{2}} \in \mathcal{E}_0^-, \\ \varphi\left(\frac{X_{i+\frac{1}{2}}-X_0^+}{\varepsilon}\right) v_{g|\chi_+} & \text{if} \quad X_{i+\frac{1}{2}} \in \mathcal{E}_0^+, \end{cases} \\ \dfrac{(X_0^+ - X_{i+\frac{1}{2}})}{|\mathcal{I}_0|} v_{g|\chi_-} + \dfrac{(X_{i+\frac{1}{2}} - X_0^-)}{|\mathcal{I}_0|} v_{g|\chi_+} \quad \text{if} \quad X_{i+\frac{1}{2}} \in \mathcal{I}_0. \end{cases} \tag{6.24}$$

Once the grid's velocity is prescribed at the grid's interfaces, the updated locations of such interfaces may be obtained as the solutions of the following family of IVPs:

$$\begin{cases} \partial_t x(X_{i+\frac{1}{2}},t) = v_{g_{|i+\frac{1}{2}}}(t), \\ x(X_{i+\frac{1}{2}},0) \quad = X_{i+\frac{1}{2}}. \end{cases} \tag{6.25}$$

Gathering (6.21), (6.24) and solving (6.25), for any time value, one have available the following sets of discrete grid's interfaces velocities $\left(v_{g_{|i+\frac{1}{2}}}(t)\right)_{0\leq i\leq n_{el}}$ and locations $\left(x_{i+\frac{1}{2}}(t)\right)_{0\leq i\leq n_{el}}$.

Remark 53. The relations (6.21)-(6.22) are initially well-defined, thanks to the assumption (5.38) on the initial data. For $t > 0$, and under the assumptions recalled in §5.3.2, the solution of (5.16)-(5.35)-(5.39)-(5.19) may exist as long as $\partial_x(\eta^e - \eta^i)_{|\chi_\pm} \neq 0$.

Remark 54. Knowing $\chi_\pm(t)$, (6.23) offers a way to dispatch the mesh elements in the moving exterior sub-domain $\mathcal{E}(t)$, avoiding elements collapsing, distorting and related stability issues. We also emphasize that (6.23) allows to properly deal with the possible occurrence of dry areas, provided that such areas are initially far enough from the object to prevent the water-height from vanishing at contact points. Indeed, assuming that the distance between $\chi_\pm(t)$ and the nearest mesh interface where the water-height vanishes is greater than ℓ, then (6.23) ensures that this mesh interface location does not vary over time.

Mapping and geometric parameters

We are now able to provide a suitable definition for the mapping (6.17) and we consider a piecewise linear and globally continuous transformation:

$$\Omega_0 \times [0, T_{\max}] \ni (X,t) \mapsto x(X,t) \in \Omega_t,$$

is such that, for any $\omega_i(0) := \left]X_{i-\frac{1}{2}}, X_{i+\frac{1}{2}}\right[\in \mathscr{T}_h(0)$, $X \in \omega_i(0)$ and $t \in [0, T_{\max}]$:

$$x_{|\omega_i(0)}(X,t) := \frac{\left(X_{i+\frac{1}{2}} - X\right)}{|\omega_i(0)|} x_{i-\frac{1}{2}}(t) + \frac{\left(X - X_{i-\frac{1}{2}}\right)}{|\omega_i(0)|} x_{i+\frac{1}{2}}(t) \in \omega_i(t). \tag{6.26}$$

From this mapping, the frame's velocity can be deduced:

Proposition 55. The frame's velocity is such that, for all $t \in [0, T_{\max}]$ and all mesh element $\omega_i(t) =]x_{i-\frac{1}{2}}(t), x_{i+\frac{1}{2}}(t)[\in \mathscr{T}_h(t)$, we have:

$$\forall x \in \omega_i(t),\ v_{\mathrm{g}|\omega_i(t)}(x,t) = \frac{\left(x_{i+\frac{1}{2}}(t) - x\right)}{|\omega_i(t)|} v_{\mathrm{g}_{|i-\frac{1}{2}}}(t) + \frac{\left(x - x_{i-\frac{1}{2}}(t)\right)}{|\omega_i(t)|} v_{\mathrm{g}_{|i+\frac{1}{2}}}(t). \tag{6.27}$$

Proof. Deriving (6.26) with respect to time gives:

$$\widetilde{v}_{\mathrm{g}|\omega_i(0)}(X,t) = \frac{(X_{i+\frac{1}{2}} - X)}{|\omega_i(0)|} v_{\mathrm{g}_{|i-\frac{1}{2}}}(t) + \frac{(X - X_{i-\frac{1}{2}})}{|\omega_i(0)|} v_{\mathrm{g}_{|i+\frac{1}{2}}}(t). \tag{6.28}$$

The deformation gradient associated with the grid's motion is obtained as the Jacobian of this mapping. In particular, the following identities are satisfied:

$$\partial_X x(X,t)_{|\omega_i(t)} =: \mathcal{J}_{\omega_i(t)} = \frac{|\omega_i(t)|}{|\omega_i(0)|},$$

$$\partial_X^k x(X,t)_{|\omega_i(t)} = 0,\ \forall k \geq 2,$$

so that the mapping is invertible and orientation-preserving. Also, for any $(X_a, X_b) \in (\omega_i(0))^2$, we have:

$$x(X_b,t) = x(X_a,t) + (X_b - X_a)\mathcal{J}_{\omega_i(t)},$$

and in particular, we deduce (6.27). □

From (6.18), we observe that the deformation gradient $\jmath = |\mathcal{J}|$ satisfies the fundamental relation, generally referred to as Geometric Conservation Law (GCL):

$$\partial_t\, \jmath(X,t) = \jmath\, \partial_x v_{\mathrm{g}}(x(X,t),t). \tag{6.30}$$

We also state an important property concerning the basis and sub-resolution basis functions:

Proposition 56. The basis functions, as well as the sub-resolution basis functions, follow the trajectories:

$$\forall \omega(t) \in \mathscr{T}_h^e(t),\ \forall p \in [\![1, \ldots, k+1]\!], \qquad \frac{\mathrm{d}}{\mathrm{d}t}\psi_p^{\omega}(x(X,t),t) = 0, \tag{6.31}$$

$$\forall \omega(t) \in \mathscr{T}_h^e(t),\ \forall m \in [\![1, \ldots, k+1]\!], \qquad \frac{\mathrm{d}}{\mathrm{d}t}\phi_m^{\omega}(x(X,t),t) = 0. \tag{6.32}$$

Proof. We have:

$$\psi_p^{\omega_i}(x,t) = \psi_p^{\omega_i}(x(X,t),t) = \left(\frac{\mathcal{J}(X - X_i)}{\mathcal{J}|\omega_i(0)|}\right)^p = \left(\frac{X - X_i}{|\omega_i(0)|}\right)^p = \widetilde{\psi}_p^{\omega_i}(X),$$

thus (6.31) is ensured, and the basis function $\psi_p^{\omega_i}(.,t)$ follows the trajectory of $x(X,t)$ in $\omega_i(t)$:

$$\partial_t \psi_p^{\omega_i}(x(X,t),t) = -v_{\mathrm{g}} \partial_x \psi_p^{\omega_i}(x(X,t),t). \tag{6.33}$$

In a similar way, the property for the sub-resolution basis derives from the piecewise linearity of the mapping. Indeed, we have by definition $\phi_m^{\omega(t)}(x(X,t),t) = \widetilde{\phi}_m^{\omega(0)}(X)$, where

$$\widetilde{\phi}_m^{\omega(0)} := \mathit{I}_{\omega(0)}^k(\mathbb{1}_m^{\omega(0)}), \quad \forall m \in [\![1,\, k+1]\!], \tag{6.34}$$

which directly implies (6.32). □

Multiplying (5.8a) by any $\psi(.,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$ satisfying $\dfrac{\mathrm{d}}{\mathrm{dt}}\psi(x(X,t),t) = 0$, and integrating over $\omega_i(t)$ gives:

$$\int_{\omega_i(t)} \psi\, \partial_t \mathbf{v}^{\mathrm{e}} \mathrm{d}x + \int_{\omega_i(t)} \psi\, \partial_x \mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) \mathrm{d}x = \int_{\omega_i(t)} \psi\, \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \mathrm{d}x. \tag{6.35}$$

Using (6.20) and (6.30) one can write:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(t)} \mathbf{v}^{\mathrm{e}} \psi \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(0)} \mathbf{v}^{\mathrm{e}} \psi \jmath \mathrm{d}X &= \int_{\omega_i(0)} \psi \frac{\mathrm{d}\mathbf{v}^{\mathrm{e}}}{\mathrm{dt}} \jmath \mathrm{d}X + \int_{\omega_i(0)} \psi \mathbf{v}^{\mathrm{e}} \partial_t \jmath \mathrm{d}X, && (6.36)\\
&= \int_{\omega_i(0)} \psi \frac{\mathrm{d}\mathbf{v}^{\mathrm{e}}}{\mathrm{dt}} \jmath \mathrm{d}X + \int_{\omega_i(0)} \psi \mathbf{v}^{\mathrm{e}} \jmath \partial_x v_{\mathrm{g}} \mathrm{d}X && (6.37)\\
&= \int_{\omega_i(t)} \psi \frac{\mathrm{d}\mathbf{v}^{\mathrm{e}}}{\mathrm{dt}} \mathrm{d}x + \int_{\omega_i(t)} \psi \mathbf{v}^{\mathrm{e}} \partial_x v_{\mathrm{g}} \mathrm{d}x && (6.38)\\
&= \int_{\omega_i(t)} \psi \partial_t \mathbf{v}^{\mathrm{e}} \mathrm{d}x + \int_{\omega_i(t)} (\psi v_{\mathrm{g}} \partial_x \mathbf{v}^{\mathrm{e}} + \psi \mathbf{v}^{\mathrm{e}} \partial_x v_{\mathrm{g}}) \mathrm{d}x && (6.39)\\
&= \int_{\omega_i(t)} \psi \partial_t \mathbf{v}^{\mathrm{e}} \mathrm{d}x + \int_{\omega_i(t)} \psi \partial_x (\mathbf{v}^{\mathrm{e}} v_{\mathrm{g}}) \mathrm{d}x, && (6.40)
\end{aligned}$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(t)} \mathbf{v}^{\mathrm{e}} \psi \mathrm{d}x = \int_{\omega_i(t)} \psi \partial_t \mathbf{v}^{\mathrm{e}} \mathrm{d}x + \int_{\omega_i(t)} \psi \partial_x (\mathbf{v}^{\mathrm{e}} v_{\mathrm{g}}) \mathrm{d}x,$$

and (6.35) becomes:

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(t)} \mathbf{v}^{\mathrm{e}} \psi \mathrm{d}x + \int_{\omega_i(t)} \psi \partial_x \mathbf{G}(\mathbf{v}^{\mathrm{e}}, b, v_{\mathrm{g}}) \mathrm{d}x = \int_{\omega_i(t)} \psi \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \mathrm{d}x, \tag{6.41}$$

where we have set $\mathbf{G}(\mathbf{v}^{\mathrm{e}}, b, v_{\mathrm{g}}) = \mathbf{F}(\mathbf{v}^{\mathrm{e}}, b) - \mathbf{v}^{\mathrm{e}} v_{\mathrm{g}}$. Another integration by parts gives:

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i} \mathbf{v}^{\mathrm{e}} \psi \mathrm{d}x - \int_{\omega_i} \mathbf{G}(\mathbf{v}^{\mathrm{e}}, b, v_{\mathrm{g}}) \partial_x \psi \, \mathrm{d}x + \left[\!\left[\psi \mathbf{G}(\mathbf{v}^{\mathrm{e}}, b, v_{\mathrm{g}}) \right]\!\right]_{\partial\omega_i(t)} = \int_{\omega_i} \psi \mathbf{B}(\mathbf{v}^{\mathrm{e}}, b') \mathrm{d}x. \tag{6.42}$$

which is the formulation retained for the next sub-section. Note that the eigenvalues and eigenvectors of the Jacobian matrix associated with $\mathbf{G}(\mathbf{v}, b, v_{\mathrm{g}})$ are trivially obtained from the NSW system written in ALE description:

$$\frac{\partial(\mathbf{F}(\mathbf{v}, b) - \mathbf{v} v_{\mathrm{g}})}{\partial \mathbf{v}}(\mathbf{v}, b) = \begin{pmatrix} v_{\mathrm{g}} & 1 \\ -u^2 + \mathrm{g}H & 2u - v_{\mathrm{g}} \end{pmatrix},$$

leading to the eigenvalues that account for the frame velocity:

$$\lambda^{\pm} := u - v_{\mathrm{g}} \pm \sqrt{\mathrm{g}H}.$$

6.1.2 DG-ALE formulation for the fluid/stationary structure model

In this sub-section, we introduce a general DG formulation in ALE description for the fluid-stationary structure problem. Let consider the coupled problem (5.16), together with initial data as specified in (5.19) with the assumptions of §5.3.2. Then, the associated DG-ALE semi-discrete formulation reads:

Find $\mathbf{v}_h^{\mathrm{e}} \in \mathcal{C}^1\big([0,T_{\max}];\big(\mathbb{P}^k(\mathscr{T}_h^e(t))\big)^2\big)$ and $\big(\underline{q}^i,\varkappa_-,\varkappa_+\big) \in \big(H^s(0,T_{\max})\big)^3$ such that, $\forall \varphi_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$ with $\dfrac{\mathrm{d}}{\mathrm{dt}}\varphi_h(x(X,t),t)=0$ the following system is ensured:

$$
\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{dt}}\big(\mathbf{v}_h^{\mathrm{e}},\ \varphi_h\big)_{\mathscr{T}_h^e(t)} + \big(\mathcal{A}_h(\mathbf{v}_h^{\mathrm{e}}),\varphi_h\big)_{\mathscr{T}_h^e(t)} = 0, \\
\mathbf{v}_h^{\mathrm{e}}(\cdot,0) \ := p^k_{\mathscr{T}_h^{e,0}}(\mathbf{v}_0^{\mathrm{e}}), \\
\eta^e_{h|\varkappa\pm} \quad\ = \eta^i_{h|\varkappa\pm}, \\
q^e_h(\varkappa_\pm,\cdot) = \underline{q}^i,
\end{cases} \tag{6.43a}
$$

$$
\begin{cases}
\eta^i_h(\cdot,t) \ := p^k_{\mathscr{T}_h^i(t)}(\eta^i), \\
\dfrac{\mathrm{d}}{\mathrm{dt}}\underline{q}^i(t) = -\Big(\displaystyle\int_{\mathcal{I}(t)}\frac{\mathrm{d}x}{H_h^i}\Big)^{-1}\Big[\!\Big[\frac{1}{2}\Big(\frac{\underline{q}^i(t)}{H_h^i}\Big)^2 + \mathrm{g}\,\eta_h^i\Big]\!\Big]_{\mathcal{I}(t)}, \\
\underline{q}^i(0) \quad := q_0^i,
\end{cases} \tag{6.43b}
$$

$$
\begin{cases}
\widetilde{v}_{\mathrm{g}|X_0^\pm} = v_{\mathrm{g}|\varkappa\pm} := \Big(\mathfrak{G}_h^k q^e_{h|\varkappa\pm}\Big)\Big(\mathfrak{G}_h^k \eta^e_{h|\varkappa\pm} - \mathfrak{G}_h^k \eta^i_{h|\varkappa\pm}\Big)^{-1}, \\
\dfrac{\mathrm{d}}{\mathrm{d}t}\varkappa_\pm(t) \qquad = \widetilde{v}_{\mathrm{g}}(X_0^\pm,t), \\
\varkappa_\pm(0) \qquad := X_0^\pm,
\end{cases} \tag{6.43c}
$$

$$
b_h(\cdot,t) := i^k_{\mathscr{T}_h(t)}(b), \tag{6.43d}
$$

where:

(i) the discrete nonlinear operator $\mathcal{A}_h$ in (6.43a) is defined by

$$
\big(\mathcal{A}_h(\mathbf{v}_h^{\mathrm{e}}),\varphi_h\big)_{\mathscr{T}_h^e(t)} := -\big(\mathbf{G}(\mathbf{v}_h^{\mathrm{e}},b_h,v_{\mathrm{g}}),\partial_x^h\varphi_h\big)_{\mathscr{T}_h^e(t)} + \big\langle \mathbf{G}^\star,\varphi_h\big\rangle_{\partial\mathscr{T}_h^e(t)} - \big(\mathbf{B}(\mathbf{v}_h^{\mathrm{e}},b_h'),\varphi_h\big)_{\mathscr{T}_h^e(t)}, \tag{6.44}
$$

and $\mathbf{G}^\star$ is an interface numerical flux which aims at approximating $\mathbf{F}(\mathbf{v},b) - v_{\mathrm{g}}\mathbf{v}$ at an interior element boundary, which is moving with velocity v_{g},

(ii) we set $\mathbf{G}^\star := \mathbf{F}^\star - v_{\mathrm{g}}\mathbf{v}^\star$, where $\mathbf{F}^\star$ and $\mathbf{v}^\star$ are also interface numerical fluxes, respectively consistent with $\mathbf{F}$ and $\mathbf{v}$, and computed with the LF formula:

$$
\mathbf{F}^\star(\mathbf{v}_R,\mathbf{v}_L,b_R,b_L) := \frac{1}{2}\left(\mathbf{F}(\mathbf{v}_R,b_R) - \mathbf{F}(\mathbf{v}_L,b_L) - \sigma(\mathbf{v}_R-\mathbf{v}_L)\right), \tag{6.45}
$$

$$
\mathbf{v}^\star(\mathbf{v}_R,\mathbf{v}_L,b_R,b_L) := \frac{1}{2}\left(\mathbf{v}_R+\mathbf{v}_L - \frac{1}{\sigma}\left(\mathbf{F}(\mathbf{v}_R,b_R) - \mathbf{F}(\mathbf{v}_L,b_L)\right)\right), \tag{6.46}
$$

with

$$\sigma := \max_{\omega \in \mathscr{T}_h^e(t)} \left(|u^e - v_{\mathrm{g}}| + \sqrt{\mathrm{g} H^e} \right)_{|\partial\omega}. \tag{6.47}$$

(iii) we introduce the following projections:

$$\eta_h^i(\cdot,t) := p^k_{\mathscr{T}_h^i(t)}(\eta^i), \quad b_h^i(\cdot,t) := i^k_{\mathscr{T}_h^i(t)}(b), \quad H_h^i := \eta_h^i - b_h^i,$$

where the implicit time dependency, due to L^2-projections onto time-dependent sub-domains, is made explicit for the sake of clarity. The interpolation of b into $\mathbb{P}^k(\mathscr{T}_h(t))$ allows to preserve the continuity of b at the mesh interfaces, provided that the elements boundary is included into the set of interpolation nodes. It also allows to easily compute a polynomial approximation of b'.

6.1.3 DG-ALE formulation for the fluid/moving structure model

In this sub-section, we introduce a general DG-ALE semi-discrete formulation associated to the free-boundary problems for the fluid/moving structure model. We directly describe the discrete formulation for the more general model (5.39), together with initial data (5.36)-(5.40) in the case of an object's free motion. The formulation for a prescribed motion can be straightforwardly deduced by forgetting the discrete dynamic equations for the object's motion (6.48e), which are replaced by some prescribed data (5.37). Then, the associated DG-ALE semi-discrete formulation reads:

for all $t \leq T_{\max}$, find $\left(\mathbf{v}_h^{\mathrm{e}}(\cdot,t), \mathbf{v}_h^{\mathrm{i}}(\cdot,t)\right) \in \mathbf{P}^k(\mathscr{T}_h^e(t)) \times \mathbf{P}^k(\mathscr{T}_h^i(t))$, $(\varkappa_-(t), \varkappa_+(t)) \in \left]x_{\mathrm{left}}(t),\, x_{\mathrm{right}}(t)\right[^2$ and $\boldsymbol{\mathcal{X}}_{\mathrm{G}}(t) \in \left]\varkappa_-(t), \varkappa_+(t)\right[\times \mathbb{R}^2$, such that the following system holds:

$$
\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{v}_h^{\mathrm{e}}, \varphi_h\right)_{\mathscr{T}_h^e(t)} + \left(\mathcal{A}_h(\mathbf{v}_h^{\mathrm{e}}), \varphi_h\right)_{\mathscr{T}_h^e(t)} = 0, \quad \forall \varphi_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t)) \text{ s.t. } \dfrac{\mathrm{d}}{\mathrm{d}t}\varphi_h(x(t),t) = 0, \\
\mathbf{v}_h^{\mathrm{e}}(\cdot,0) := p^k_{\mathscr{T}_h^{\mathrm{e},0}}(\mathbf{v}_0^{\mathrm{e}}), \\
\mathbf{v}_{h|\varkappa_\pm}^{\mathrm{e}} = \mathbf{v}_{h|\varkappa_\pm}^{\mathrm{i}},
\end{cases} \tag{6.48a}
$$

$$
\begin{cases}
\eta_h^i(\cdot,t) := p^k_{\mathscr{T}_h^i(t)} \circ F_h(\cdot, t, \boldsymbol{\mathcal{X}}_{\mathrm{G}}(t), \boldsymbol{\mathcal{X}}_{\mathrm{G}}(0), \eta_{\mathrm{lid}}), \\
q_h^i(\cdot,t) := p^k_{\mathscr{T}_h^i(t)}\left(\boldsymbol{\vartheta}_{\mathrm{G}}(t) \cdot \boldsymbol{\mathcal{T}}_{\mathrm{G},h}(\cdot,t) + \underline{q}^i(t)\right), \\
\dfrac{\mathrm{d}}{\mathrm{d}t}\underline{q}^i(t) = -\left(\left\langle\!\left\langle f_{1,h}\right\rangle\!\right\rangle_h + \left\langle\!\left\langle f_{2,h}\right\rangle\!\right\rangle_h + \left\langle\!\left\langle f_{3,h}\right\rangle\!\right\rangle_h\right), \\
\underline{q}^i(0) := q_0^i,
\end{cases} \tag{6.48b}
$$

$$
\begin{cases}
\widetilde{v}_{\mathrm{g}|X_0^\pm} = v_{\mathrm{g}|\varkappa_\pm} := \left(\mathfrak{G}_h^k q_{h|\varkappa_\pm}^e + \left(\mathbf{v}_{\mathrm{G}} - w\mathbf{r}_{\mathrm{G},h}^\perp\right) \cdot \mathbf{n}_h^{\mathrm{i}}\right)\left(\mathfrak{G}_h^k \eta_{h|\varkappa_\pm}^e - \mathfrak{G}_h^k \eta_{h|\varkappa_\pm}^i\right)^{-1}, \\
\dfrac{\mathrm{d}}{\mathrm{d}t}\varkappa_\pm(t) = \widetilde{v}_{\mathrm{g}}(X_0^\pm, t), \\
\varkappa_\pm(0) := X_0^\pm,
\end{cases} \tag{6.48c}
$$

$$
b_h(\cdot,t) := i^k_{\mathscr{T}_h(t)}(b), \tag{6.48d}
$$

$$
\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{X}}_{\mathrm{G}} = \boldsymbol{\vartheta}_{\mathrm{G}}, \\
\left(\mathbb{M}_0 + \mathbb{M}_{\mathrm{a},h}\left[H_h^i, \boldsymbol{\mathcal{T}}_{\mathrm{G},h}\right]\right)\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\vartheta}_{\mathrm{G}} = \begin{pmatrix} -m_{\mathrm{o}}\mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho\mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(t)}\left[\left(f_{1,h}^\star + f_{3,h}^\star\right)\dfrac{\boldsymbol{\mathcal{T}}_{\mathrm{G},h}^\star}{H_h^i}\right], \\
\boldsymbol{\mathcal{X}}_{\mathrm{G}}(0) := (X_0, Z_0, 0), \\
\boldsymbol{\vartheta}_{\mathrm{G}}(0) := (u_{\mathrm{G}}^0, w_{\mathrm{G}}^0, w_0),
\end{cases} \tag{6.48e}
$$

where:

(i) the discrete nonlinear operator $\mathcal{A}_h$ in (6.48a) is defined as in (6.44),
(ii) the first equation in (6.48b) offers a way to compute a high-order broken polynomial approximation of the specific part of the object's underside, which projection along the horizontal line at

time t identifies to $\mathcal{I}(t)$. More precisely, $\digamma_h(\cdot, t, \boldsymbol{\mathcal{X}}_{\rm G}(t), \boldsymbol{\mathcal{X}}_{\rm G}(0), \eta_{\rm lid})$ is a discrete nonlinear operator such that, for all $x \in \mathcal{I}(t)$:

$$\digamma_h(x, t, \boldsymbol{\mathcal{X}}_{\rm G}(t), \boldsymbol{\mathcal{X}}_{\rm G}(0), \eta_{\rm lid}) := \widetilde{F}(X_h, t, \boldsymbol{\mathcal{X}}_{\rm G}(t), \boldsymbol{\mathcal{X}}_{\rm G}(0), \eta_{\rm lid}), \tag{6.49}$$

where X_h is an approximation of the following nonlinear equation's unique root, obtained by Newton iterations:

$$\frac{x - x_{\rm G}(t) + \sin(\theta(t))(\eta_{\rm lid}(X_h) - Z_{\rm G})}{\cos(\theta(t))} + X_{\rm G} - X_h = 0. \tag{6.50}$$

Having η_h^i in hands, one can compute

$$H_h^i := \eta_h^i - p^k_{\mathscr{T}_h^i(t)}(b), \tag{6.51}$$

(iii) the second and third equations in (6.48b) allow to compute an approximation of the discharge in the interior domain, through the evaluation of a purely geometrical term, together with a time-dependent term obtained as the solution of a nonlinear ordinary differential equation. Specifically, we set:

$$\mathbf{r}_{{\rm G},h}(\cdot, t) := \begin{pmatrix} \cdot - x_{\rm G}(t) \\ \eta_h^i(\cdot, t) - z_{\rm G}(t) \end{pmatrix}, \quad \boldsymbol{\mathcal{T}}_{{\rm G},h}(\cdot, t) := \begin{pmatrix} -\mathbf{r}^{\perp}_{{\rm G},h}(\cdot, t) \\ \frac{1}{2}\left|\mathbf{r}_{{\rm G},h}(\cdot, t)\right|^2 \end{pmatrix}, \tag{6.52}$$

and the discrete versions of the right-hand sides are defined as follows

$$\begin{aligned} f_{1,h} &:= \mathfrak{G}_h^k \circ p^k_{\mathscr{T}_h^i(t)}\left(u_h^i q_h^i\right) + {\rm g} H_h^i \mathfrak{G}_h^k \eta_h^i, \\ f_{2,h} &:= \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\vartheta}_{\rm G} \centerdot \boldsymbol{\mathcal{T}}_{{\rm G},h}, \\ f_{3,h} &:= \boldsymbol{\vartheta}_{\rm G}^T \, \mathbb{M}_{{\rm G},h} \boldsymbol{\vartheta}_{\rm G}, \end{aligned} \tag{6.53}$$

where we use a discrete version of (5.23) to evaluate the term $\partial_t \boldsymbol{\mathcal{T}}_{\rm G}$ that appears in f_3, with

$$\mathbb{M}_{{\rm G},h} := \begin{pmatrix} \mathbf{e}_{\rm x} \centerdot \mathbf{n}_{\rm lid} & 0 & -\mathbf{r}^{\perp}_{{\rm G},h} \centerdot \mathbf{n}_{\rm lid} \\ 1 & 0 & 0 \\ -\mathbf{r}^{\perp}_{{\rm G},h} \centerdot \mathbf{n}_{\rm lid} & 0 & -(\mathbf{e}_{\rm z} \centerdot \mathbf{r}_{{\rm G},h})(\mathbf{r}^{\perp}_{{\rm G},h} \centerdot \mathbf{n}_{\rm lid}) \end{pmatrix}.$$

and we recall that the discrete version of the $\mathcal{I}(t)$-averaging operator is provided in (6.7),

(iv) the BVPs (6.48c) allow to compute the time evolution of the contact-points $\varkappa_{\pm}(t)$, and therefore to re-define a new mesh-grid accordingly using (6.24) and (6.28).

(v) the discrete contact-points velocity (6.48c) is obtained from (6.21), using the expression of the time-derivative (5.34). Once this velocity is known, the the updated $\Omega_t = \mathcal{E}(t) \cup \mathcal{I}(t)$ may be computed,

(vi) the discrete counterpart of the added-mass-inertia matrix, denoted by $\mathbb{M}_{{\rm a},h}\left[H_h^i, \boldsymbol{\mathcal{T}}_{{\rm G},h}\right]$, is defined as follows:

$$\mathbb{M}_{{\rm a},h}\left[H_h^i, \boldsymbol{\mathcal{T}}_{{\rm G},h}\right] := \mathfrak{S}^{h,n_{\rm g}}_{\mathcal{I}(t)}\left(\frac{\boldsymbol{\mathcal{T}}^{\star}_{{\rm G},h} \otimes \boldsymbol{\mathcal{T}}^{\star}_{{\rm G},h}}{H_h^i}\right). \tag{6.54}$$

This matrix is simply denoted by $\mathbb{M}_{{\rm a},h}$ in what follows.

Remark 57. The boundary conditions on $\partial\Omega_t$ are weakly enforced through the numerical fluxes $\mathbf{G}^\star$. As far as boundary conditions are concerned on $\partial\Omega$, we may enforce any type of boundary conditions usually available for the NSW equations, including inflow and outflow conditions within subcritical or supercritical configurations relying on local Riemann invariants, periodic conditions or solid-wall conditions.

Remark 58. In practice, the positivity of H_h^i (defined in (6.51)) may be obtained from the sizing introduced in Remark 47. Indeed, considering a flat bottom for the sake of simplicity, assuming that $H_h^{\mathrm{i},0} \gg 0$, together with $\mathscr{d}_o \ll H_0$ and considering that the object can not be completely immersed, with an object's motion $\boldsymbol{\mathcal{X}}_{\mathrm{G}}$ such that $\mathcal{M}_G$ remains close to its initial location, then we necessarily have $H_h^i(\cdot,t) > 0$ for $t \in [0, T_{\max}]$. According to (6.49), this entails that $\digamma_h(\cdot, t, \boldsymbol{\mathcal{X}}_{\mathrm{G}}(t), \boldsymbol{\mathcal{X}}_{\mathrm{G}}(0), \eta_{\mathrm{lid}}) > 0$. This can be also verified when the topography is not flat by assuming that $\mathscr{d}_o \ll H_b$.

Remark 59. This general DG-ALE formulation has still to be supplemented with some specific treatments to ensure its robustness and to handle the topography variations in a well-balanced way. These issues are addressed in the remainder of this section. A global flowchart of the resulting general algorithm, detailing the processing order of these various numerical ingredients for the most complex case (*i.e.* case of a freely floating body), is also provided in §6.4.

6.1.4 Time-marching algorithms

For a given final computational time $T_{\max} > 0$, we consider a partition $(t^n)_{0\le n\le N}$ of the time interval $[0, T_{\max}]$ with $t^0 := 0$, $t^N := T_{\max}$ and $t^{n+1} - t^n =: \Delta t^n$. For any sufficiently regular function w depending on time , we set $w^n := w(t^n)$ and in what follows, such a "superscript n" notation may be used with any time-varying entity, evaluated at discrete time t^n. In particular, we note:

$$\mathcal{E}^n := \mathcal{E}(t^n),\ \mathcal{I}^n := \mathcal{I}(t^n),\ \varkappa_-^n := \varkappa_-(t^n),\ \varkappa_+^n := \varkappa_+(t^n),\ \mathscr{T}_h^n := \mathscr{T}_h(t^n),\ \mathscr{T}_h^{e,n} := \mathscr{T}_h^e(t^n),$$

and so on, together with similar notations for the main unknowns of the problem:

$$\mathbf{v}_h^{\mathrm{e},n} := \mathbf{v}_h^{\mathrm{e}}(\cdot,t^n),\quad \mathbf{v}_h^{\mathrm{i},n} := \mathbf{v}_h^{\mathrm{i}}(\cdot,t^n),\quad \varkappa_\pm^n := \varkappa_\pm(t^n),\quad \boldsymbol{\mathcal{X}}_{\mathrm{G}}^n := \boldsymbol{\mathcal{X}}_{\mathrm{G}}(t^n),\quad \boldsymbol{\vartheta}_{\mathrm{G}}^n := \boldsymbol{\vartheta}_{\mathrm{G}}(t^n).$$

When fully-discrete formulations are considered, the time-stepping is carried out with explicit SSP-RK schemes [74, 149]. For instance, writing the semi-discrete evolution equation of (6.48a) (or (6.43a)) in the operator form

$$\partial_t \mathbf{v}_h^{\mathrm{e}} + \mathcal{A}_h(\mathbf{v}_h^{\mathrm{e}}) = 0,$$

we advance the discrete solution $\mathbf{v}_h^{\mathrm{e},n} \in \mathbf{P}^k(\mathscr{T}_h^{e,n})$ from time-level n to level $(n+1)$, with $\mathbf{v}_h^{\mathrm{e},n+1} \in \mathbf{P}^k(\mathscr{T}_h^{e,n+1})$, through the third-order SSP-RK scheme as follows:

$$\begin{aligned}
\mathbf{v}_h^{e,n,(1)} &= \mathbf{v}_h^{e,n} - \Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n}),\\
\mathbf{v}_h^{e,n,(2)} &= \frac{1}{4}(3\mathbf{v}_h^{e,n} + \mathbf{v}_h^{e,n,(1)}) - \frac{1}{4}\Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n,(1)}),\\
\mathbf{v}_h^{e,n+1} &= \frac{1}{3}(\mathbf{v}_h^{e,n} + 2\mathbf{v}_h^{e,n,(2)}) - \frac{2}{3}\Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n,(2)}),
\end{aligned} \tag{6.55}$$

where $\mathbf{v}_h^{e,n,(i)}$, $1 \le i \le 2$, are the solutions obtained at intermediate stages and Δt^n is obtained from the CFL condition (6.56). Anticipating on the description of our stability-enforcement operator in the next section, which relies on both DG approximations on mesh elements $\omega^n \in \mathscr{T}_h^{e,n}$ and FV

schemes on the subcells $S_m^{\omega,n} \in \mathscr{T}_\omega^n$, the time-step Δt^n is computed adaptively using the following CFL condition:

$$\Delta t^n = \frac{\min\limits_{\omega^n \in \mathscr{T}_h^{e,n}} \left(\dfrac{h_\omega^n}{2k+1}, \min\limits_{S_m^{\omega,n} \in \mathscr{T}_\omega^n} |S_m^{\omega,n}| \right)}{\sigma}, \tag{6.56}$$

where σ is the constant previously introduced in (6.47). The same SSP-RK method for the discretization of (6.48c) leads to the following discrete algorithm:

$$\begin{aligned} \chi_\pm^{n,(1)} &= \chi_\pm^n + \Delta t^n v_{\mathrm{g}|\chi_\pm}^n, \\ \chi_\pm^{n,(2)} &= \frac{3\chi_\pm^n + \chi_\pm^{n,(1)}}{4} + \frac{\Delta t^n}{4} v_{\mathrm{g}|\chi_\pm}^{n,(1)}, \\ \chi_\pm^{n+1} &= \frac{\chi_\pm^n + 2\chi_\pm^{n,(2)}}{3} + \frac{2\Delta t^n}{3} v_{\mathrm{g}|\chi_\pm}^{n,(2)}. \end{aligned} \tag{6.57}$$

Also, we use the same SSP-RK method for the discretization of the EDOs equations as (6.18), (6.43b), (6.48b) and (6.48e).

6.1.5 DG-ALE as a FV-ALE scheme on subcells

It is well-established that the discrete formulation (6.43a) (or (6.48a)) needs some additional stabilization in order to ensure the positivity of H_h^e at the discrete level, and to avoid Gibbs phenomenon in the vicinity of spatial discontinuities, sharp gradients or smooth extrema. In order to design some suitable correction mechanisms, we show that the FV-Subcell reformulation of the DG method for the NSW equations developed in § 3, and initially introduced in [164] for general hyperbolic conservation laws, may be extended to the current DG-ALE framework. We follow the lines of § 3 while highlighting the differences due to the frame's motion. Let introduce the following projections onto $\mathbb{P}^k(\mathscr{T}_h^e(t))^2$:

$$\mathbf{F}_h^e := p^k_{\mathscr{T}_h^e(t)}(\mathbf{F}(\mathbf{v}_h^{\mathrm{e}}, b_h)) \quad \text{and} \quad \mathbf{B}_h^e := p^k_{\mathscr{T}_h^e(t)}(\mathbf{B}(\mathbf{v}_h^{\mathrm{e}}, b_h')), \tag{6.58}$$

together with the respective shortcuts $\mathbf{F}_{\omega_i(t)} = \mathbf{F}^e_{h|\omega_i(t)}$, $\mathbf{B}_{\omega_i(t)} = \mathbf{B}^e_{h|\omega_i(t)}$, $\mathbf{G}_{\omega_i} = \mathbf{F}_{\omega_i} - v_{\mathrm{g}}\mathbf{v}_{\omega_i}$. We substitute these projections into (6.43a) and integrate by parts the second term to obtain, for all $\psi(.,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t))$ satisfying $\dfrac{\mathrm{d}}{\mathrm{dt}}\psi(x(X,t),t) = 0$:

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(t)} \mathbf{v}^{\mathrm{e}}_{\omega_i} \psi \mathrm{d}x = -\int_{\omega_i(t)} \partial_x \mathbf{G}_{\omega_i} \psi \mathrm{d}x + \left[\!\left[(\mathbf{G}_{\omega_i} - \mathbf{G}^\star)\psi \right]\!\right]_{\partial\omega_i(t)} + \int_{\omega_i(t)} \mathbf{B}_{\omega_i} \psi \mathrm{d}x. \tag{6.59}$$

For a given mesh element $\omega(t) \in \mathscr{T}_h^e(t)$, we consider a sub-partition $\mathscr{T}_\omega(t)$ defined in (6.9), together with the sub-resolution basis functions (6.10). Substituting $\psi = \phi_m^{\omega_i}$ into (6.59), for all m in $[\![1, \ldots, k+1]\!]$, recalling the definition of the sub-mean-values $\overline{\mathbf{v}}_m^{\omega_i}$ in (6.14), recalling also that $\mathbf{v}_{\omega_i}$, $\partial_x(\mathbf{v}_{\omega_i} v_{\mathrm{g}})$, $\partial_x \mathbf{F}_{\omega_i}$ and $\mathbf{B}_{\omega_i}$ all belong to $\mathbb{P}^k\left(\omega_i(t)\right)^2$ and finally using identity (6.13), the following discrete formulation holds, for all m in $[\![1, \ldots, k+1]\!]$:

$$\frac{\mathrm{d}}{\mathrm{dt}} \left(|S_m^{\omega_i}(t)| \overline{\mathbf{v}}_m^{\omega_i} \right) = -\left[\!\left[\mathbf{G}_{\omega_i} \right]\!\right]_{\partial S_m^\omega(t)} + \left[\!\left[(\mathbf{G}_{\omega_i} - \mathbf{G}^\star)\phi_m^{\omega_i} \right]\!\right]_{\partial\omega_i(t)} + |S_m^{\omega_i}(t)| \overline{\mathbf{B}}_m^{\omega_i}. \tag{6.60}$$

We introduce the $k+2$ subcell's *reconstructed fluxes*, denoted by $\{\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}}\}_{m\in[\![0,\,k+1]\!]}$, and defined as the solution of the following linear system:

$$\begin{aligned}
\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}} &:= [\![\mathbf{G}_{\omega_i}]\!]_{\partial S^{\omega}_m(t)} - [\![(\mathbf{G}_{\omega_i} - \mathbf{G}^{\star})\phi^{\omega_i}_m]\!]_{\partial\omega_i(t)}, \quad \forall m \in [\![1,\,k+1]\!],\\
\widehat{\mathbf{G}}^{\omega_i}_{\frac{1}{2}} &:= \mathbf{G}^{\star}_{i-\frac{1}{2}},\\
\widehat{\mathbf{G}}^{\omega_i}_{k+3/2} &:= \mathbf{G}^{\star}_{i+\frac{1}{2}},
\end{aligned}$$

so that (6.43a) may be recast as a FV-ALE formulation on the sub-partition:

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(|S^{\omega_i}_m(t)|\overline{\mathbf{v}}^{\omega_i}_m\right) = -\left(\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}\right) + |S^{\omega_i}_m(t)|\overline{\mathbf{B}}^{\omega_i}_m, \quad \forall m \in [\![1,\,k+1]\!]. \tag{6.61}$$

Remark 60. For practical purpose, an explicit formula for the computation of the interior reconstructed fluxes for $m \in [\![1,\ldots,k]\!]$ is:

$$\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}} = \mathbf{G}_{\omega_i}\left(\widetilde{x}_{m+\frac{1}{2}}\right) - C^{i-\frac{1}{2}}_{m+\frac{1}{2}}\left(\mathbf{G}_{\omega_i}\left(x_{i-\frac{1}{2}}\right) - \mathbf{G}^{\star}_{i-\frac{1}{2}}\right) - C^{i+\frac{1}{2}}_{m+\frac{1}{2}}\left(\mathbf{G}_{\omega_i}\left(x_{i+\frac{1}{2}}\right) - \mathbf{G}^{\star}_{i+\frac{1}{2}}\right), \tag{6.62}$$

with

$$C^{i-\frac{1}{2}}_{m+\frac{1}{2}} = \sum_{p=m+1}^{k+1}\phi^{\omega_i}_p\left(x_{i-\frac{1}{2}}\right) \qquad \text{and} \qquad C^{i+\frac{1}{2}}_{m+\frac{1}{2}} = \sum_{p=1}^{m}\phi^{\omega_i}_p\left(x_{i+\frac{1}{2}}\right). \tag{6.63}$$

Remark 61. We require that the integrals and source term in (6.43a) are exactly computed at motionless steady states. This can be achieved, thanks to the *pre-balanced* formulation of the NSW equations, by using any quadrature rule that is exact for polynomials of degree up to $2k$. Let us recall that $2\,k$ is in any case the minimum requirement to reach the expected $k+1$ order of accuracy.

6.1.6 Subcell low-order corrected FV-ALE fluxes

In this sub-section, we show that the reconstructed fluxes may be locally corrected to enforce some required properties. As investigated in § 3 for the NSW equations, lowest-order FV fluxes may be introduced in order to: (i) prevent high-order approximations from spurious oscillations in the vicinity of discontinuities and sharp gradients, (ii) ensure the preservation of the water-height's positivity. Additionally, one needs to introduce some states reconstructions, inspired from [116] in order to ensure a well-balancing property. In what follows, we recall the definition of such corrected fluxes, highlighting the new terms associated with the frame's motion. The specification of suitable admissibility criteria is postponed to the next section.
For any time value $t \in [0, T_{\max}]$, $\omega_i(t) \in \mathscr{T}^e_h(t)$, and any marked subcell $S^{\omega_i}_m(t) \in \mathscr{T}_{\omega_i(t)}$, let define the sub-partition interface values for b, where the subscript ω_i and the time dependency are forgotten for the sake of simplicity:

$$\overline{b}_{m+\frac{1}{2}} := \max(\overline{b}_m, \overline{b}_{m+1}), \quad \overline{b}^{\pm}_m := \overline{b}_{m\pm\frac{1}{2}} - \max\left(0, \overline{b}_{m\pm\frac{1}{2}} - \overline{\eta}_m\right).$$

Subcell's interfaces reconstructions for the water-height are defined as follows:

$$\overline{H}_m^{\pm} := \max\left(0, \overline{\eta}_m - \bar{b}_{m\pm\frac{1}{2}}\right),$$

and the corresponding free surface elevation and discharge are deduced as follows:

$$\overline{\eta}_m^{\pm} := \overline{H}_m^{\pm} + \bar{b}_m^{\pm}, \qquad \overline{q}_m^{\pm} := \overline{H}_m^{\pm} \frac{\overline{q}_m}{\overline{H}_m}, \qquad \overline{\mathbf{v}}_m^{\pm} := (\overline{\eta}_m^{\pm}, \overline{q}_m^{\pm}), \tag{6.64}$$

where $\bar{b}_m^{\pm}$ refer to the trace of $\bar{b}_m$ at the subcell's interfaces. Related lowest-order numerical fluxes on subcell's $S_m(t)$ left and right interfaces are built accordingly:

$$\mathcal{F}^l_{m+\frac{1}{2}} := \mathbf{F}^{\star}\left(\overline{\mathbf{v}}_m^+, \overline{\mathbf{v}}_{m+1}^-, \bar{b}_m^+, \bar{b}_m^+\right) + \begin{pmatrix} 0 \\ \mathrm{g}\overline{\eta}_m^+\left(\bar{b}_m^+ - b_{\widetilde{x}_{m+\frac{1}{2}}}\right) \end{pmatrix}, \tag{6.65}$$

$$\mathcal{F}^r_{m-\frac{1}{2}} := \mathbf{F}^{\star}\left(\overline{\mathbf{v}}_{m-1}^+, \overline{\mathbf{v}}_m^-, \bar{b}_m^-, \bar{b}_m^-\right) + \begin{pmatrix} 0 \\ \mathrm{g}\overline{\eta}_m^-\left(\bar{b}_m^- - b_{\widetilde{x}_{m-\frac{1}{2}}}\right) \end{pmatrix}, \tag{6.66}$$

where $b_{\widetilde{x}_{m\pm\frac{1}{2}}}$ are respectively the interpolated polynomial values of b_h at $\widetilde{x}_{m+\frac{1}{2}}$ and $\widetilde{x}_{m-\frac{1}{2}}$. The associated numerical flux, in the ALE description, are deduced as follows:

$$\mathcal{G}^l_{m+\frac{1}{2}} := \mathcal{F}^l_{m+\frac{1}{2}} - v_{\mathrm{g}_{|m+\frac{1}{2}}} \mathbf{v}^{\star,l}_{m+\frac{1}{2}}, \quad \text{and} \quad \mathcal{G}^r_{m-\frac{1}{2}} := \mathcal{F}^r_{m-\frac{1}{2}} - v_{\mathrm{g}_{|m-\frac{1}{2}}} \mathbf{v}^{\star,r}_{m-\frac{1}{2}}, \tag{6.67}$$

with

$$\mathbf{v}^{\star,l}_{m+\frac{1}{2}} := \mathbf{v}^{\star}\left(\overline{\mathbf{v}}_m^+, \overline{\mathbf{v}}_{m+1}^-, \bar{b}_m^+, \bar{b}_m^+\right), \quad \text{and} \quad \mathbf{v}^{\star,r}_{m-\frac{1}{2}} := \mathbf{v}^{\star}\left(\overline{\mathbf{v}}_{m-1}^+, \overline{\mathbf{v}}_m^-, \bar{b}_m^-, \bar{b}_m^-\right). \tag{6.68}$$

Using such corrected FV-ALE fluxes, it is possible to modify the reconstructed fluxes $\widehat{\mathbf{G}}_{m+\frac{1}{2}}$ in a robust way, in some particular subcells, where the *uncorrected* DG scheme (6.61) has failed to produce an admissible solution. We are thus left with the issues of identifying the local subcells that may need some corrections and defining a robust correction procedure, which are respectively addressed in §6.2 and §6.3.

Remark 62. For the fully wet case, we show that our DG-ALE (for an arbitrary v_{g}) scheme with *a posteriori* LSC method, conserve the well-balanced property, see § 6.3.2. As near wet/dry regions, the mesh-grid velocity v_{g} is equal to zero, see Remark 54, so we come across the Eulerian framework, where we already shown the well-balanced property in § 3 for wet and wet/dry context. In addition, practically, we do not obtain negative water-height, since the ALE-moving grid ($v_{\mathrm{g}} \neq 0$) method is only activated in a "sufficiently" wet area far enough from dry regions. As for dry and almost dry regions, we have already shown in § 3 the preservation of water-height positivity for Eulerian method ($v_{\mathrm{g}} = 0$).

6.2 Admissibility criteria

A large number of sensors or detectors have been introduced in the literature in order to identify the particular cells/subcells in which some additional stabilization mechanisms are required. We use always the two admissibility criteria used in § 3: one for the *Physical Admissibility Detection*

(PAD) and the other to address the occurrence of spurious oscillations, called *Subcell Numerical Admissibility Detection* (SubNAD). This last criterion is supplemented with a relaxation procedure to exclude the smooth extrema from the troubled cells. These criteria, which definitions are not recalled in this chapter, are detailed in § 3.7.

6.3 *A posteriori* LSC method for DG-ALE scheme

Gathering all the previous ingredients, we introduce a global algorithm that ensures the stability and robustness of the flow's computation in $\mathcal{E}(t)$. This algorithm is adapted from § 3 and extended to the current DG-ALE framework. We only provide a qualitative description and focus on the steps that require further comments, due to the additional ALE description.
Starting from an admissible piecewise polynomial approximate solution $\mathbf{v}_h^{e,n} \in \left(\mathbb{P}^k(\mathscr{T}_h^{e,n})\right)^2$ at discrete time t^n, we first compute a *predictor candidate solution* $\mathbf{v}_h^{e,n+1} \in \left(\mathbb{P}^k(\mathscr{T}_h^{n+1})\right)^2$ at time t^{n+1} using the uncorrected DG-ALE scheme (6.44), together with the corresponding SSP-RK time discretization of §6.1.4. Then, for any mesh element $\omega_i^{n+1} \in \mathscr{T}_h^{e,n+1}$, we compute the predictor candidate sub-mean-values:

$$\mathbb{P}^0(\mathscr{T}_{\omega_i}^{e,n+1}) \ni \overline{\mathbf{v}}_{\omega_i}^{e,n+1} = \pi_{\mathscr{T}_{\omega_i}^{e,n+1}}(\mathbf{v}_{\omega_i}^{e,n+1}).$$

For any subcell $S_m^{\omega_i,n+1} \in \mathscr{T}_{\omega_i}^{n+1}$, we check admissibility of the associated sub-mean-values $\overline{\mathbf{v}}_m^{\omega_i,n+1}$ using the criteria of §6.2. For a given subcell $S_m^{\omega_i,n+1}$ that may need additional stabilization, the corresponding DG *reconstructed interface fluxes* $\widehat{\mathbf{G}}_{m\pm\frac{1}{2}}$ defined in (6.62), which were initially used to compute the predictor candidate $\mathbf{v}_h^{e,n+1}$, may be replaced by the FV *corrected fluxes* $\mathcal{G}_{m\pm\frac{1}{2}}^{l/r}$ of (6.67) into the update process to compute a new candidate subcell value through the local FV-ALE formulation (6.61). Both left and right interface fluxes, or only left or right interface flux, may be replaced depending on the admissibility of the neighboring subcells. The complete set of substituting rules is not recalled here (see § 3 for a complete description), but concisely, the new updating process for subcell value $\overline{\mathbf{v}}_m^{\omega_i,n+1}$ may fall into one of the following alternative:

$$i)\ \frac{\mathrm{d}}{\mathrm{dt}}\left(|S_m^{\omega_i}(t)|\overline{\mathbf{v}}_m^{\omega_i}\right) = -\left(\mathcal{G}_{m+\frac{1}{2}}^{l} - \mathcal{G}_{m-\frac{1}{2}}^{r}\right) + |S_m^{\omega_i}(t)|\overline{\mathbf{B}}_m^{\omega_i}, \tag{6.69}$$

$$ii)\ \frac{\mathrm{d}}{\mathrm{dt}}\left(|S_m^{\omega_i}(t)|\overline{\mathbf{v}}_m^{\omega_i}\right) = -\left(\mathcal{G}_{m+\frac{1}{2}}^{l} - \widehat{\mathbf{G}}_{m-\frac{1}{2}}^{\omega_i}\right) + |S_m^{\omega_i}(t)|\overline{\mathbf{B}}_m^{\omega_i}, \tag{6.70}$$

$$iii)\ \frac{\mathrm{d}}{\mathrm{dt}}\left(|S_m^{\omega_i}(t)|\overline{\mathbf{v}}_m^{\omega_i}\right) = -\left(\widehat{\mathbf{G}}_{m+\frac{1}{2}}^{\omega_i} - \mathcal{G}_{m-\frac{1}{2}}^{r}\right) + |S_m^{\omega_i}(t)|\overline{\mathbf{B}}_m^{\omega_i}. \tag{6.71}$$

For mesh elements $\omega_i(t)$ in which such fluxes corrections have occurred, leading to the computation of updated/limited sub-mean-values, a new high-order polynomial candidate solution, still denoted by $\mathbf{v}_h^{e,n+1}$ for the sake of simplicity, is built upon these updated sub-mean-values:

$$\mathbb{P}^k(\omega_i^n) \ni \mathbf{v}_{\omega_i}^{e,n+1} = \pi^{-1}_{\mathscr{T}_{\omega_i}^{n+1}}(\overline{\mathbf{v}}_{\omega_i}^{e,n+1}),$$

and the process may go further in time after checking that this new candidate is admissible.

The whole detection-correction-projection iterative process may be conveniently summarized through the application of a stabilization/correction operator denoted as follows:

$$\begin{array}{rccc} \Lambda_h^{k,n} & : & \left(\mathbb{P}^k\left(\mathscr{T}_h^{e,n}\right)\right)^2 & \to & \left(\mathbb{P}^k\left(\mathscr{T}_h^{e,n}\right)\right)^2, \\ & & \mathbf{v}_h^{e,n} & \mapsto & \Lambda_h^{k,n}(\mathbf{v}_h^{e,n}), \end{array} \tag{6.72}$$

where the resulting broken polynomial $\Lambda_h^{k,n}(\mathbf{v}_h^{e,n})$ satisfies all the admissibility criteria, see §6.2. Embedding such a stabilization operator into a fully discrete version of (6.43a), with for instance a third order SSP-RK method, would simply gives:

$$\begin{aligned}
\mathbf{v}_h^{e,n,(1)} &= \Lambda_h^{k,n,(1)}\Big(\mathbf{v}_h^{e,n} - \Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n})\Big), \\
\mathbf{v}_h^{e,n,(2)} &= \Lambda_h^{k,n,(2)}\Big(\frac{1}{4}(3\mathbf{v}_h^{e,n} + \mathbf{v}_h^{e,n,(1)}) - \frac{1}{4}\Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n,(1)})\Big), \\
\mathbf{v}_h^{e,n+1} &= \Lambda_h^{k,n+1}\Big(\frac{1}{3}(\mathbf{v}_h^{e,n} + 2\mathbf{v}_h^{e,n,(2)}) - \frac{2}{3}\Delta t^n \mathcal{A}_h(\mathbf{v}_h^{e,n,(2)})\Big),
\end{aligned} \tag{6.73}$$

and, for the sake of simplicity, this may be summarized within a semi-discrete notation as follows:

$$\begin{aligned}
&\frac{d}{dt}\big(\Lambda_h^k(\mathbf{v}_h^{\mathrm{e}}), \varphi_h\big)_{\mathscr{T}_h^e(t)} + \big(\mathcal{A}_h\big(\mathbf{v}_h^{\mathrm{e}}\big), \varphi_h\big)_{\mathscr{T}_h^e(t)} = 0, \quad \forall \varphi_h(\cdot,t) \in \mathbb{P}^k(\mathscr{T}_h^e(t)), \\
&\mathbf{v}_h^{\mathrm{e}}(\cdot,0) = p^k_{\mathscr{T}_h^{e,0}}(\mathbf{v}_0^{\mathrm{e}}),
\end{aligned} \tag{6.74}$$

where the shortcut semi-discrete notation $\frac{d}{dt}\Lambda_h^k(\mathbf{v}_h^{\mathrm{e}})$ simply means: *apply a posteriori LSC stabilization procedure to any fully discrete solution obtained through any chosen time-marching algorithm, at any updated discrete time or intermediate stage.*

6.3.1 A RK-DG-ALE fully-discrete formulation

In this sub-section, we describe the fully-discrete formulation obtained by considering the most complex case (*i.e.*, a case of a freely floating body) (6.48) together with a first-order Euler time-marching algorithm. Any higher-order RK-DG-ALE formulation based on §6.1.4 can be straightforwardly deduced from this lowest-order one by adapting accordingly the various time stages. Assuming that the needed quantities are available at discrete time t^n, the first-order in time fully-discrete formulation associated with (5.39) reads as follows:
find $\left(\mathbf{v}_h^{\mathrm{e},n+1}, \mathbf{v}_h^{\mathrm{i},n+1}\right) \in \mathbf{P}^k(\mathscr{T}_h^{e,n+1}) \times \mathbf{P}^k(\mathscr{T}_h^{i,n+1})$, $\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1} \in \left]\varkappa_-^{n+1}, \varkappa_+^{n+1}\right[\times \mathbb{R}^2$ and $(\varkappa_-^{n+1}, \varkappa_+^{n+1}) \in \left]x_{\mathrm{left}}^{n+1}, x_{\mathrm{right}}^{n+1}\right[^2$, such that the following system holds:

$$
\begin{cases}
v_{\mathrm{g}|\varkappa_\pm}^n := \left(\mathfrak{G}_h^k q_{h|\varkappa_\pm}^{\mathrm{e},n} + \left(\mathbf{v}_{\mathrm{G}}^n - w\mathbf{r}_{\mathrm{G},h}^{\perp,n}\right) \cdot \mathbf{n}_h^{\mathrm{i},n}\right)\left(\mathfrak{G}_h^k \eta_{h|\varkappa_\pm}^{\mathrm{e},n} - \mathfrak{G}_h^k \eta_{h|\varkappa_\pm}^{\mathrm{i},n}\right)^{-1}, \\
\varkappa_\pm^{n+1} - \varkappa_\pm^n = \Delta t^n \widetilde{v_{\mathrm{g}}}^n(X_0^\pm),
\end{cases} \tag{6.75a}
$$

$$
b_h^{n+1} := i_{\mathscr{T}_h^{n+1}}^k(b), \tag{6.75b}
$$

$$
\begin{cases}
\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1} - \boldsymbol{\mathcal{X}}_{\mathrm{G}}^n = \Delta t^n \boldsymbol{\vartheta}_{\mathrm{G}}^n, \\
\left(\mathbb{M}_0 + \mathbb{M}_{\mathrm{a},h}\right)\left(\boldsymbol{\vartheta}_{\mathrm{G}}^{n+1} - \boldsymbol{\vartheta}_{\mathrm{G}}^n\right) = \Delta t^n \left\{ \begin{pmatrix} -m_o \mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho \mathfrak{S}_{\mathcal{I}^n}^{h,n_{\mathrm{g}}} \left((f_{1,h}^{\star,n} + f_{3,h}^{\star,n}) \frac{\boldsymbol{\mathcal{T}}_{\mathrm{G},h}^{\star,n}}{H_h^{\mathrm{i},n}} \right) \right\},
\end{cases} \tag{6.75c}
$$

$$
\begin{cases}
\eta_h^{\mathrm{i},n+1} := \mathrm{p}_{\mathscr{T}_h^{i,n+1}}^k \circ \mathcal{F}_h^{n+1}(\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1}; \boldsymbol{\mathcal{X}}_{\mathrm{G}}^0, \eta_{\mathrm{lid}}), \\
q_h^{\mathrm{i},n+1} := \mathrm{p}_{\mathscr{T}_h^{i,n+1}}^k \left(\boldsymbol{\vartheta}_{\mathrm{G}}^{n+1} \cdot \boldsymbol{\mathcal{T}}_{\mathrm{G},h}^{n+1} + \underline{q}^{i,n+1} \right), \\
\underline{q}^{i,n+1} - \underline{q}^{i,n} = -\Delta t^n \left(\langle\!\langle f_{1,h}^n \rangle\!\rangle_h + \langle\!\langle f_{2,h}^n \rangle\!\rangle_h + \langle\!\langle f_{3,h}^n \rangle\!\rangle_h \right),
\end{cases} \tag{6.75d}
$$

$$
\begin{cases}
\mathbf{v}_h^{\mathrm{e},n+1} := \Lambda_h^{k,n+1}\left(\mathbf{v}_h^{\mathrm{e},n} - \Delta t^n \mathcal{A}_h(\mathbf{v}_h^{\mathrm{e},n})\right), \\
\mathbf{v}_{h|\varkappa_\pm}^{\mathrm{e},n+1} = \mathbf{v}_{h|\varkappa_\pm}^{\mathrm{i},n+1},
\end{cases} \tag{6.75e}
$$

and the first iteration is initialized with the following data:

$$
\mathbf{v}_h^{\mathrm{e},0} := p_{\mathscr{T}_h^{e,0}}^k(\mathbf{v}_0^{\mathrm{e}}), \text{ with } \mathbf{v}_0^{\mathrm{e}} \in (H^s(\mathcal{E}_0))^2, \tag{6.76a}
$$

$$
\boldsymbol{\mathcal{X}}_{\mathrm{G}}^0 := \boldsymbol{\mathcal{X}}_{\mathrm{G}}(0), \tag{6.76b}
$$

$$
\boldsymbol{\vartheta}_{\mathrm{G}}^0 := \boldsymbol{\vartheta}_{\mathrm{G}}(0), \tag{6.76c}
$$

$$
\varkappa_\pm^0 := X_0^\pm, \tag{6.76d}
$$

$$
\underline{q}^{i,0} := q_0^i. \tag{6.76e}
$$

Note that, for the sake of simplicity, the DG scheme (6.75e) is written in the operator form. In practice, it may be convenient either to express the corresponding scalar-products at the initial time, or to use the equivalent FV formulation on the subcells.

6.3.2 Some properties

Invertibility of the discrete *added-mass* matrix

Proposition 63. The matrix $\mathbb{M}_{\mathrm{a},h}\left[H_h^i, \boldsymbol{\mathcal{T}}_{\mathrm{G},h}\right]$ is symmetric and non-negative.

Proof. The matrix $\boldsymbol{\mathcal{T}}^\star_{\mathrm{G},h} \otimes \boldsymbol{\mathcal{T}}^\star_{\mathrm{G},h}$ is obviously symmetric and of rank one, with non-negative eigenvalues, thus a non-negative matrix. From the implicit additional assumption that $H_h^i \geq 0$ (some justifications are provided in Remarks 47 and 58), and since we are using quadrature rules with positive coefficients, it results that $\mathbb{M}_{\mathrm{a},h}\left[H_h^i, \boldsymbol{\mathcal{T}}_{\mathrm{G},h}\right]$ is a symmetric and non-negative matrix. □

Properties of the DG-ALE formulation with the *a posteriori* LSC method

In this section, we show that the resulting global fully-discrete DG-ALE scheme with *a posteriori* LSC is globally well-balanced for motionless steady states and satisfies the DGCL. To avoid repetition, we prove the properties for the stationary partly immersed structure case (5.16), considering the discrete formulation (6.43). We also show how the well-balanced property is also extended to the case of moving structure, see Remarks 65.

Well-balancing for motionless steady states

Let begin with the well-balanced property. Motionless steady states for problem (5.16) are trivially defined as follows:

$$\mathbf{v}^{\mathrm{e}}(\cdot,t) = \mathbf{v}^{\mathrm{c}} = \begin{pmatrix} \eta^{\mathrm{c}} \\ 0 \end{pmatrix}, \quad \underline{q}^i(t) = 0, \quad \varkappa_{\pm}(t) = X_0^{\pm}, \quad \forall t \geq 0. \tag{6.77}$$

We highlight that proving that the global semi-discrete formulation (6.43) preserves such steady states is equivalent to prove that the DG-ALE scheme (6.43a) is well-balanced on $\mathcal{E}(t) = \mathcal{E}^-(t) \cup \mathcal{E}^+(t)$, which again reduces to ensure the property on $\mathcal{E}^-(t)$ and $\mathcal{E}^+(t)$ separately. Indeed, it is straightforward to observe that at steady states, (6.43b)-(6.43c) lead to

$$\frac{d}{dt}\underline{q}^i(t) = 0, \qquad v_{\mathrm{g}|\varkappa_\pm}(t) = 0, \qquad \eta^i_{\varkappa_\pm} = \eta^{\mathrm{c}},$$

so that the coupling with the floating structure actually does not disturb the flow steady state. Hence, we have the following result for the first-order in time fully discrete formulation:

Proposition 64. The discrete formulation (6.43) with possible occurence of local corrected lowest-order fluxes in one of the three possible formulations (6.69)-(6.70)-(6.71), together with a first-order Euler time-marching algorithm, preserves the motionless steady states (6.77), provided that the integrals of (6.43a) are exactly computed at motionless steady states. Specifically, for all $n \geq 0$,

$$\left(\eta_h^n = \eta^c \text{ and } q_h^n = 0\right) \implies \left(\eta_h^{n+1} = \eta^c \text{ and } q_h^{n+1} = 0\right).$$

Proof. At steady states, for any given t and any mesh element $\omega(t)$, we have

$$\partial_x \mathbf{F}(\mathbf{v}_{\omega(t)}, b_{\omega(t)}) = \mathbf{B}(\mathbf{v}_{\omega(t)}, b'_{\omega(t)}). \tag{6.78}$$

Furthermore, both $\mathbf{F}(\mathbf{v}_h^{\mathrm{e}}, b_h)$ and $\mathbf{B}(\mathbf{v}_h^{\mathrm{e}}, b_h')$ belong to $\left(\mathbb{P}^k(\mathscr{T}_h^e(t))\right)^2$ so that we have:

$$\mathbf{F}_h := p^k_{\mathscr{T}_h^e(t)}(\mathbf{F}(\mathbf{v}_h^{\mathrm{e}}, b_h)) = \mathbf{F}(\mathbf{v}_h^{\mathrm{e}}, b_h), \tag{6.79}$$

$$\mathbf{B}_h := p^k_{\mathscr{T}_h^e(t)}(\mathbf{B}(\mathbf{v}_h^{\mathrm{e}}, b_h')) = \mathbf{B}(\mathbf{v}_h^{\mathrm{e}}, b_h'). \tag{6.80}$$

We also emphasize that it is equivalent to prove the property for the formulation (6.61) on the sub-partitions or for the formulation (6.43a) on $\mathscr{T}_h^{e,n}$. We choose to work with (6.61) and we want to show that the scheme is well-balanced at the subcell level:

$$\begin{aligned} &\forall \omega^n \in \mathscr{T}_h^{e,n}, \quad \forall m \in [\![1, \dots, k+1]\!], \ \overline{\eta}_m^{\omega,n} = \eta^c, \ \overline{q}_m^{\omega,n} = 0 \\ &\qquad \Longrightarrow \quad \forall \omega^{n+1} \in \mathscr{T}_h^{e,n+1}, \quad \forall m \in [\![1, \dots, k+1]\!], \ \overline{\eta}_m^{\omega,n+1} = \eta^c, \quad \overline{q}_m^{\omega,n+1} = 0. \end{aligned} \tag{6.81}$$

As stated in §6.3, investigating the various possibilities for the definition of the interface fluxes implies to investigate the "uncorrected" situation (6.61) (corresponding to high-order DG reconstructed fluxes at all subcells interfaces) plus three "corrected" situations enumerated in (6.69)-(6.70)-(6.71) (and corresponding to the occurrence of modified lowest-order FV fluxes at (some of) the subcells interfaces). As (6.70) and (6.71) boil down to the same situation with a permutation of left and right fluxes, we have, for any given value $m \in [\![1, \dots, k+1]\!]$, to distinguish three different situations:

case 1 - admissible subcell: $S_{m-1}^{\omega_i}$, $S_m^{\omega_i}$ and $S_{m+1}^{\omega_i}$ are all admissible. The local time-update formula (with reconstructed fluxes) is

$$|S_m^{\omega_i,n+1}|\overline{\mathbf{v}}_m^{\omega_i,n+1} = |S_m^{\omega_i,n}|\overline{\mathbf{v}}_m^{\omega_i,n} - \Delta t^n\big(\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}\big) + \Delta t^n |S_m^{\omega_i,n}|\overline{\mathbf{B}}_m^{\omega_i}. \tag{6.82}$$

where $\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}}$ and $\widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}$ are defined in (6.62). We observe that at steady state:

$$\eta^+_{i\pm\frac{1}{2}} = \eta^-_{i\pm\frac{1}{2}} = \eta^c, \qquad q^+_{i\pm\frac{1}{2}} = q^-_{i\pm\frac{1}{2}} = 0, \qquad \text{and} \qquad b^+_{i\pm\frac{1}{2}} = b^-_{i\pm\frac{1}{2}},$$

and therefore,

$$\mathbf{F}^\star_{i\pm\frac{1}{2}} = \begin{pmatrix} 0 \\ \frac{1}{2}\mathrm{g}\eta^c\big(\eta^c - 2b_{i\pm\frac{1}{2}}\big) \end{pmatrix} = \mathbf{F}\big(\mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}}, b_{\omega_i|x_{i\pm\frac{1}{2}}}\big), \tag{6.83}$$

and,

$$\mathbf{v}^\star_{i\pm\frac{1}{2}} = \begin{pmatrix} \eta^c \\ 0 \end{pmatrix} = \mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}},$$

resulting in

$$\mathbf{G}^\star_{i\pm\frac{1}{2}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}}, b_{\omega_i|x_{i\pm\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|i\pm\frac{1}{2}}}\mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}}. \tag{6.84}$$

Using (6.79), we also have:

$$\mathbf{G}_{\omega_i|x_{i\pm\frac{1}{2}}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}}, b_{\omega_i|x_{i\pm\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|i\pm\frac{1}{2}}}\mathbf{v}_{\omega_i|x_{i\pm\frac{1}{2}}}, \tag{6.85}$$

thus, using the definition (6.62) of $\widehat{\mathbf{G}}^{\omega_i}_{m\pm\frac{1}{2}}$, we obtain:

$$\widehat{\mathbf{G}}^{\omega_i}_{m\pm\frac{1}{2}} = \mathbf{G}_{\omega_i|\widetilde{x}_{m\pm\frac{1}{2}}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|\widetilde{x}_{m\pm\frac{1}{2}}}, b_{\omega_i|\widetilde{x}_{m\pm\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|i\pm\frac{1}{2}}}\mathbf{v}_{\omega_i|\widetilde{x}_{m\pm\frac{1}{2}}}, \tag{6.86}$$

allowing to compute the difference:

$$\widehat{\mathbf{G}}^{\omega_i}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}} = \int_{S^i_m} \partial_x \mathbf{F}(\mathbf{v}_{\omega_i}, b_{\omega_i})\mathrm{d}x - (v_{\mathrm{g}_{|m+\frac{1}{2}}} - v_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c. \tag{6.87}$$

Additionally, updating in time the frame's interfaces with a first-order Euler scheme leads to:

$$x^{n+1} = x^n + \Delta t^n v^n_{\mathrm{g}},$$

so that the geometric term may be simplified as follows:

$$(v_{\mathrm{g}_{|m+\frac{1}{2}}} - v_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c = \frac{\tilde{x}^{n+1}_{m+\frac{1}{2}} - \tilde{x}^{n+1}_{m-\frac{1}{2}} - (\tilde{x}^{n}_{m+\frac{1}{2}} - \tilde{x}^{n}_{m-\frac{1}{2}})}{\Delta t^n}\mathbf{v}^c = \frac{|S^{\omega_i,n+1}_m| - |S^{\omega_i,n}_m|}{\Delta t^n}\mathbf{v}^c.$$

Finally, (6.82) writes:

$$\begin{aligned}|S^{\omega_i,n+1}_m|\overline{\mathbf{v}}^{\omega_i,n+1}_m = |S^{\omega_i,n}_m|\mathbf{v}^c - \Delta t^n \left(\int_{S^{\omega_i,n}_m} \partial_x \mathbf{F}(\mathbf{v}^n_{\omega_i}, b_{\omega_i}) - \mathbf{B}(\mathbf{v}^n_{\omega_i}, \partial_x b_{\omega_i})\mathrm{d}x\right)\\ + |S^{\omega_i,n+1}_m|\mathbf{v}^c - |S^{\omega_i,n}_m|\mathbf{v}^c,\end{aligned}$$

and using (6.78), we obtain:

$$\overline{\mathbf{v}}^{\omega_i,n+1}_m = \mathbf{v}^c = \overline{\mathbf{v}}^{\omega_i,n}_m. \tag{6.88}$$

case 2 - neighbor of a non-admissible subcell: $S^{\omega_i}_m$, $S^{\omega_i}_{m-1}$ are admissible but $S^{\omega_i}_{m+1}$ is non-admissible (the symmetric situation of $S^{\omega_i}_m$, $S^{\omega_i}_{m+1}$ are admissible but $S^{\omega_i}_{m-1}$ is non-admissible may be treated in a similar way). The corresponding time-update formula is :

$$|S^{\omega_i,n+1}_m|\overline{\mathbf{v}}^{\omega_i,n+1}_m = |S^{\omega_i,n}_m|\overline{\mathbf{v}}^{\omega_i,n}_m - \Delta t^n\big(\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}\big) + \Delta t^n |S^{\omega_i,n}_m|\overline{\mathbf{B}}^{\omega_i}_m, \tag{6.89}$$

with $\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}}$ and $\widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}$ defined in (6.65)-(6.67) and (6.62). To evaluate $\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}}$ at steady state, we observe that $\overline{\eta}^+_m = \overline{\eta}^-_{m+1} = \eta^c$, leading to:

$$\mathbf{F}^\star\left(\overline{\mathbf{v}}^+_m, \overline{\mathbf{v}}^-_{m+1}, \overline{b}^+_m, \overline{b}^+_m\right) = \frac{1}{2}\begin{pmatrix} 0 \\ g\eta^c\left(\eta^c - 2\overline{b}^+_m\right)\end{pmatrix},$$

and

$$\mathcal{F}^{\omega_i,l}_{m+\frac{1}{2}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}, b_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}\Big). \tag{6.90}$$

As we also have

$$\mathbf{v}^{\star,l}_{m+\frac{1}{2}} = \mathbf{v}^\star\left(\overline{\mathbf{v}}^+_m, \overline{\mathbf{v}}^-_{m+1}, \overline{b}^+_m, \overline{b}^+_m\right) = \mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}},$$

we obtain

$$\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}, b_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|m+\frac{1}{2}}}\mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}.$$

The computation of $\widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}}$ is performed as in **case 1**, leading to (6.86) and we may evaluate the difference as follows:

$$\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} - \widehat{\mathbf{G}}^{\omega_i}_{m-\frac{1}{2}} = \int_{S^{\omega_i}_m} \partial_x \mathbf{F}(\mathbf{v}_{\omega_i}, b_{\omega_i})\mathrm{d}x - (v_{\mathrm{g}_{|m+\frac{1}{2}}} - v_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c, \tag{6.91}$$

so that,

$$\overline{\mathbf{v}}^{\omega_i,n+1}_m = \mathbf{v}^c = \overline{\mathbf{v}}^{\omega_i,n}_m.$$

case 3 - corrected subcell: $S^{\omega_i}_m$ is non-admissible. The time-update formula is:

$$|S^{\omega_i,n+1}_m|\overline{\mathbf{v}}^{\omega_i,n+1}_m = |S^{\omega_i,n}_m|\overline{\mathbf{v}}^{\omega_i,n}_m - \Delta t^n\big(\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} - \mathcal{G}^{\omega_i,r}_{m-\frac{1}{2}}\big) + \Delta t^n |S^{\omega_i,n}_m|\overline{\mathbf{B}}^{\omega_i}_m. \tag{6.92}$$

with $\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}}$ and $\mathcal{G}^{\omega_i,r}_{m-\frac{1}{2}}$ defined in (6.67). Reproducing the computation steps as in **case 2**, we obtain:

$$\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}, b_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|m+\frac{1}{2}}}\mathbf{v}_{\omega_i|\widetilde{x}_{m+\frac{1}{2}}}, \tag{6.93}$$

$$\mathcal{G}^{\omega_i,r}_{m-\frac{1}{2}} = \mathbf{F}\Big(\mathbf{v}_{\omega_i|\widetilde{x}_{m-\frac{1}{2}}}, b_{\omega_i|\widetilde{x}_{m-\frac{1}{2}}}\Big) - v_{\mathrm{g}_{|m-\frac{1}{2}}}\mathbf{v}_{\omega_i|\widetilde{x}_{m-\frac{1}{2}}}, \tag{6.94}$$

and

$$\mathcal{G}^{\omega_i,l}_{m+\frac{1}{2}} - \mathcal{G}^{\omega_i,r}_{m-\frac{1}{2}} = \int_{S^{\omega_i}_m} \partial_x \mathbf{F}(v_{\omega_i}, b_{\omega_i})\mathrm{d}x - (v_{\mathrm{g}_{|m+\frac{1}{2}}} - v_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c, \tag{6.95}$$

so that,

$$\overline{\mathbf{v}}^{\omega_i,n+1}_m = \mathbf{v}^c = \overline{\mathbf{v}}^{\omega_i,n}_m.$$

□

Remark 65. The well-balanced property may also be obtained for the moving-object problem (6.48). Motionless steady states for (6.48) are defined as follows:

$$\mathbf{v}^{\mathrm{e}}(\cdot,t) = \mathbf{v}^{\mathrm{c}}, \quad q^i(t) = 0, \quad \chi_{\pm}(t) = X^{\pm}_0, \quad \boldsymbol{\mathcal{X}}_{\mathrm{G}}(t) = \boldsymbol{\mathcal{X}}_{\mathrm{G}}(0) \ \text{ and } \ \boldsymbol{\vartheta}_{\mathrm{G}}(t) = 0, \quad \forall t \geq 0, \tag{6.96}$$

with the additional constraint that $\boldsymbol{\mathcal{X}}_{\mathrm{G}}(0)$ should be defined according to the mass m_o and inertia coefficient i_o of the object, so that Newton's laws are initially balanced too (or for the sake of simplicity, according to Appendix E, the mass m_o may be chosen such that the initial discrete acceleration is equal to zero, for a given value of $\boldsymbol{\mathcal{X}}_{\mathrm{G}}(0)$). Indeed, under the assumptions and configuration of the previous proposition, we consider the following flow/object equilibrium state, at time-step $t = t^n$:

$$\mathbf{v}^{\mathrm{e,n}}_h = \mathbf{v}^{\mathrm{c}}, \quad q^{\mathrm{i},n}_h = 0, \quad \chi^n_{\pm} = X^{\pm}_0, \quad \boldsymbol{\mathcal{X}}^n_{\mathrm{G}} = \boldsymbol{\mathcal{X}}^0_{\mathrm{G}}, \ \text{ and } \ \boldsymbol{\vartheta}^n_{\mathrm{G}} = 0, \tag{6.97}$$

together with the initial dynamic balance of Appendix E

$$\begin{pmatrix} -m_o \mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} = \rho \mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(0)} \begin{pmatrix} (f^{\star,0}_{1,h} + f^{\star,0}_{3,h}) \dfrac{\boldsymbol{\mathcal{T}}^{\star,0}_{\mathrm{G},h}}{H^{\mathrm{i},n}_h} \end{pmatrix},$$

then it is possible to show that this entails $\boldsymbol{\vartheta}_{\mathrm{G}}^{n+1} = 0$ and thus $\left\langle\!\left\langle f_{1,h}^n \right\rangle\!\right\rangle_h = \left\langle\!\left\langle f_{2,h}^n \right\rangle\!\right\rangle_h = \left\langle\!\left\langle f_{3,h}^n \right\rangle\!\right\rangle_h = 0$, leading to $q_h^{i,n+1} = q_h^{i,n} = 0$. The fact that $v_{\mathrm{g}|\chi_\pm}^n = 0$ can be deduced from the first and last assumptions of (6.97) and thus

$$\chi_\pm^{n+1} = \chi_\pm^n = X_0^\pm,$$

and the fact that $\boldsymbol{\vartheta}_{\mathrm{G}}^n = 0$ also implies that

$$\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1} = \boldsymbol{\mathcal{X}}_{\mathrm{G}}^n = \boldsymbol{\mathcal{X}}_{\mathrm{G}}^0.$$

We proceed in the same way as in Proposition 64 to show the well-balanceness property for the DG-ALE scheme (6.75e), and we get:

$$\mathbf{v}_h^{\mathrm{e,n+1}} = \mathbf{v}^{\mathrm{c}}.$$

Finally, the moving object problem (6.48) respect the well-balanced property:

$$\mathbf{v}_h^{\mathrm{e,n+1}} = \mathbf{v}^{\mathrm{c}}, \quad q_h^{\mathrm{i},n+1} = 0, \quad \chi_\pm^{n+1} = X_0^\pm, \quad \boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1} = \boldsymbol{\mathcal{X}}_{\mathrm{G}}^0, \quad \text{and} \quad \boldsymbol{\vartheta}_{\mathrm{G}}^{n+1} = 0.$$

Remark 66. This well-balanced property can be extended to any higher-order SSP-RK time discretization that can be expressed as a convex combination of first-order Euler schemes.

Discrete Geometric Conservation Law (DGCL)

In simulations of free surface flows involving free moving boundaries, it is important to ensure that the proposed numerical scheme in ALE description exactly preserves uniform flows. Such preservation property is called Geometric Conservation Law in the literature and simply states that the moving mesh procedure does not disturb the uniform flow configuration. Hence, considering $\Omega_t = \mathcal{E}(t)$ (no floating object) and $b = 0$, we inject a constant solution $\mathbf{v}_h^{\mathrm{e}}(\cdot, t) = (\eta^{\mathrm{c}}, q^{\mathrm{c}})$ into (6.43a), together with

$$\varphi_h(x,t) := \mathbb{1}^{\omega_i}(x,t) = \begin{cases} 1 & \text{if } x \in \omega_i(t), \\ 0 & \text{if } x \notin \omega_i(t), \end{cases} \qquad \forall m \in [\![1,\, k+1]\!],$$

to obtain:

$$\mathbf{v}^c \frac{\mathrm{d}}{\mathrm{dt}} \int_{\omega_i(t)} \mathrm{d}x = -\left[\!\left[\mathbf{F}(\mathbf{v}^c, 0) - v_{\mathrm{g}} \mathbf{v}^c\right]\!\right]_{\partial\omega_i(t)} = \mathbf{v}^c \left[\!\left[v_{\mathrm{g}}\right]\!\right]_{\partial\omega_i(t)},$$

and thus the GCL reduces to the following (automatically satisfied) property:

$$\frac{\mathrm{d}}{\mathrm{dt}} |\omega_i(t)| = \left[\!\left[v_{\mathrm{g}}\right]\!\right]_{\partial\omega_i(t)}. \tag{6.98}$$

At the fully discrete level, we show that a fully discrete formulation, relying on a third-order RK time discretization, satisfies the DGCL.

Proposition 67. The high-order DG-ALE semi-discrete scheme (6.43a), together with the third-order accurate SSP-RK time-marching algorithm (6.73) and the embedded stabilization operator with possible occurence of corrected lowest-order fluxes in one of the following formulations (6.69)-(6.70)-(6.71), preserves the Discrete Geometric Conservation Law. Specifically, assuming $b = 0$, we have, for any discrete time t^n:

$$\left(\mathbf{v}_h^{\mathrm{e},n} = \mathbf{v}^{\mathrm{c}}\right) \implies \left(\mathbf{v}_h^{\mathrm{e},n+1} = \mathbf{v}^{\mathrm{c}}\right).$$

Proof. Under the assumption $\mathbf{v}_h^{\mathrm{e},n} = \mathbf{v}^{\mathrm{c}}$, we have $\mathbf{F}(\mathbf{v}_h^{\mathrm{e}}, b_h) \in \left(\mathbb{P}^k(\mathscr{T}_h^{e,n})\right)^2$, and the following identity holds:

$$\mathbf{F}_h^n := p^k_{\mathscr{T}_h^{e,n}}(\mathbf{F}(\mathbf{v}_h^{\mathrm{e},n}, b_h)) = \mathbf{F}(\mathbf{v}_h^{\mathrm{e},n}, b_h). \tag{6.99}$$

As in the proof of Proposition 64, it is equivalent to show that the property holds at the subcell level, using formulation (6.61). Let denote by $|S_m^{\omega_i,(1)}|$, $|S_m^{\omega_i,(2)}|$, and $|S_m^{\omega_i,n+1}|$ the length of the subcell $S_m^{\omega_i}$ at the three different time stages of the SSP-RK algorithm (and whenever the RK stage dependency has to be specified, we apply the superscripts $(.)^{(1)}$, $(.)^{(2)}$ and $(.)^{n+1}$ to the concerned quantities). The 3^{rd} order SSP-RK discretization reads as follows:

$$\begin{cases} |S_m^{\omega_i,(1)}|\mathbf{v}_m^{\omega_i,(1)} = |S_m^{\omega_i,n}|\mathbf{v}_m^{\omega_i,n} + \Delta t^n \mathcal{R}_m^{\omega_i,n}, \\ |S_m^{\omega_i,(2)}|\mathbf{v}_m^{\omega_i,(2)} = \dfrac{3|S_m^{\omega_i,n}|\mathbf{v}_m^{\omega_i,n} + |S_m^{\omega_i,(1)}|\mathbf{v}_m^{\omega_i,(1)}}{4} + \dfrac{\Delta t^n}{4}\mathcal{R}_m^{\omega_i,(1)}, \\ |S_m^{\omega_i,n+1}|\mathbf{v}_m^{\omega_i,n+1} = \dfrac{|S_m^{\omega_i,n}|\mathbf{v}_m^{\omega_i,n} + 2|S_m^{\omega_i,(2)}|\mathbf{v}_m^{\omega_i,(2)}}{3} + \dfrac{2\Delta t^n}{3}\mathcal{R}_m^{\omega_i,(2)}. \end{cases} \tag{6.100}$$

We have, for any given value $m \in [\![1, \dots, k+1]\!]$, to distinguish three different cases:
case 1 - admissible subcell: $S_{m-1}^{\omega_i}$, $S_m^{\omega_i}$ and $S_{m+1}^{\omega_i}$ are all admissible. The residual $\mathcal{R}_m^{\omega_i,(j)}$ is:

$$\mathcal{R}_m^{\omega_i,(j)} = -\left(\widehat{\mathbf{G}}_{m+\frac{1}{2}}^{\omega_i,(j)} - \widehat{\mathbf{G}}_{m-\frac{1}{2}}^{\omega_i,(j)}\right), \tag{6.101}$$

As we assume $\mathbf{v}_h^{\mathrm{e},n} = \mathbf{v}^c$, using (6.99), we get :

$$\widehat{\mathbf{G}}_{m\pm\frac{1}{2}}^{\omega_i,n} = \mathbf{G}_{\omega_i|\widetilde{x}_{m\pm\frac{1}{2}}}^n = \mathbf{F}\left(\mathbf{v}^c, 0\right) - v^n_{\mathrm{g}_{|m\pm\frac{1}{2}}}\mathbf{v}^c,$$

so that

$$\widehat{\mathbf{G}}_{m+\frac{1}{2}}^{\omega_i,n} - \widehat{\mathbf{G}}_{m-\frac{1}{2}}^{\omega_i,n} = (v^n_{\mathrm{g}_{|m+\frac{1}{2}}} - v^n_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^{\mathrm{c}}.$$

Using the SSP-RK time update of the grid motion (as in (6.57)), we obtain:

$$|S_m^{\omega_i,(1)}|\mathbf{v}_m^{\omega_i,(1)} = |S_m^{\omega_i,n}|\mathbf{v}^{\mathrm{c}} + \Delta t^n\left(\frac{\widetilde{x}^{(1)}_{m+\frac{1}{2}} - \widetilde{x}^n_{m+\frac{1}{2}} - \widetilde{x}^{(1)}_{m-\frac{1}{2}} + \widetilde{x}^n_{m-\frac{1}{2}}}{\Delta t^n}\right)\mathbf{v}^{\mathrm{c}} \tag{6.102}$$

$$= |S_m^{\omega_i,n}|\mathbf{v}^{\mathrm{c}} + |S_m^{\omega_i,(1)}|\mathbf{v}^{\mathrm{c}} - |S_m^{\omega_i,n}|\mathbf{v}^{\mathrm{c}}, \tag{6.103}$$

$$= |S_m^{\omega_i,(1)}|\mathbf{v}^{\mathrm{c}}, \tag{6.104}$$

and therefore:

$$\mathbf{v}_m^{\omega_i,(1)} = \mathbf{v}^c.$$

In a similar way, we show that $\mathbf{v}_m^{\omega_i,(2)} = \mathbf{v}^c$ and $\mathbf{v}_m^{\omega_i,n+1} = \mathbf{v}^c$, leading to the desired conclusion.

case 2 - corrected subcell: $S_m^{\omega_i}$ is non-admissible. The residual $\mathcal{R}_m^{\omega_i,(j)}$ is:

$$\mathcal{R}_m^{\omega_i,(j)} = -\left(\mathcal{G}_{m+\frac{1}{2}}^{l,\omega_i,(j)} - \mathcal{G}_{m-\frac{1}{2}}^{r,\omega_i,(j)}\right), \tag{6.105}$$

and one can show that we have:

$$\mathcal{G}^{l,\omega_i,n}_{m+\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}^c,0\right) - v^n_{\mathrm{g}_{|m+\frac{1}{2}}}\mathbf{v}^c, \quad \text{and} \quad \mathcal{G}^{l,\omega_i,n}_{m-\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}^c,0\right) - v^n_{\mathrm{g}_{|m-\frac{1}{2}}}\mathbf{v}^c,$$

leading to

$$\mathcal{G}^{l,\omega_i,n}_{m+\frac{1}{2}} - \mathcal{G}^{r,\omega_i,n}_{m-\frac{1}{2}} = (v^n_{\mathrm{g}_{|m+\frac{1}{2}}} - v^n_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c,$$

and therefore

$$\mathbf{v}^{\omega_i,n+1}_m = \mathbf{v}^c.$$

case 3 - neighbor of a non-admissible subcell: $S^{\omega_i}_m$, $S^{\omega_i}_{m-1}$ are admissible but $S^{\omega_i}_{m+1}$ is non-admissible (the symmetric situation of $S^{\omega_i}_m$, $S^{\omega_i}_{m+1}$ are admissible but $S^{\omega_i}_{m-1}$ is non-admissible may be treated in a similar way). The residual $\mathcal{R}^{\omega_i,(j)}_m$ in a mixed DG/FV context is:

$$\mathcal{R}^{\omega_i,(j)}_m = -\left(\widehat{\mathbf{G}}^{\omega_i,(j)}_{m+\frac{1}{2}} - \mathcal{G}^{r,\omega_i,(j)}_{m-\frac{1}{2}}\right). \tag{6.106}$$

As in the two previous situations, we have:

$$\widehat{\mathbf{G}}^{\omega_i,n}_{m+\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}^c,0\right) - v^n_{\mathrm{g}_{|m+\frac{1}{2}}}\mathbf{v}^c,$$
$$\mathcal{G}^{r,\omega_i,,n}_{m-\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}^c,0\right) - v^n_{\mathrm{g}_{|m-\frac{1}{2}}}\mathbf{v}^c,$$

leading to

$$\widehat{\mathbf{G}}^{\omega_i,n}_{m+\frac{1}{2}} - \mathcal{G}^{r,\omega_i,n}_{m-\frac{1}{2}} = (v^n_{\mathrm{g}_{|m+\frac{1}{2}}} - v^n_{\mathrm{g}_{|m-\frac{1}{2}}})\mathbf{v}^c,$$

so that,

$$\mathbf{v}^{\omega_i,n+1}_m = \mathbf{v}^c.$$

□

Remark 68. One can also show the GCL property at the sub-cell level:

$$\frac{\mathrm{d}}{\mathrm{dt}}|S^{\omega}_m(t)| = [\![v_{\mathrm{g}}]\!]_{\partial S^{\omega}_m(t)}, \tag{6.107}$$

by considering the semi-discretised FV-like scheme (6.61) with the possible occurence of corrected lowest-order fluxes in one of the following formulations (6.69)-(6.70)-(6.61).

6.4 Flowchart

Let summarize the global numerical strategy associated with (6.48) for the simulation of free surface waves and a freely floating object interactions in shallow-water. For the sake of simplicity, a first-order Euler time-marching scheme is assumed to produce some fully discrete approximations. This procedure may be straightforwardly extended to higher-order time-marching algorithms.

Starting from available and admissible values of $\mathbf{v}^{\mathrm{e},n}_h$, $\chi^n_{\pm}$, $\boldsymbol{\mathcal{X}}^n_{\mathrm{G}}$, $\boldsymbol{\vartheta}^n_{\mathrm{G}}$, $\underline{q}^{i,n}$, $\eta^{i,n}_h$ and $b^n_h := i^k_{\mathscr{T}^n_h}(b)$,

1. locally compute the frame's velocity at contact points $v^n_{\mathrm{g}|\chi_{\pm}}$ (6.75a) that is obtained from (6.21), using the expression of the time-derivative (5.34) and the discrete first-order derivative (6.5),
2. globally compute the frame's velocity v^n_{g} for all mesh interfaces for $\mathscr{T}^{e,n}_h$ and $\mathscr{T}^{i,n}_h$, using (6.24),

3. globally compute the updated locations of mesh interfaces at discrete time $t = t^{n+1}$ using the frame's velocity v_{g}^{n} and the family of IVP (6.25). In particular, the updated positions of the contact points $x_{\pm}^{n+1}$ are obtained,

4. from the set of discrete values for grid points velocity and interfaces location, we reconstruct continuous profiles for both quantities inside mesh elements using (6.26) and (6.28), in order to qualify updated quadrature nodes or subcells interfaces and compute velocities at these points (used in the definition of numerical fluxes),

5. compute the new spatial-angular coordinates $\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n+1}$ from the knowledge of $\boldsymbol{\mathcal{X}}_{\mathrm{G}}^{n}$ and $\boldsymbol{\vartheta}_{\mathrm{G}}^{n}$ and then compute $\boldsymbol{\vartheta}_{\mathrm{G}}^{n+1}$ by solving Newton's second law system for the conservation of linear and angular momentum, see (6.75c),

6. compute the updated solution $\mathbf{v}_h^{\mathrm{e},n+1}$ and $\mathbf{v}_h^{\mathrm{i},n+1}$ while accounting for the geometric terms related to the displacement of the mesh, position and velocity of the floating body, as well as for the stabilization procedure (in the *exterior* region),

7. Increment time and cycle to step 1.

6.5 Numerical validations

In this section, we provide several numerical assessments of the DG-ALE discrete formulations for the water-body interaction model, associated with a third-order SSP-RK time-marching scheme. In what follows, unless stated otherwise, we choose to display sub-mean-values instead of point-wise values of the polynomial approximations, as it allows to precisely illustrate the subcell resolution of the scheme. In the following numerical validations, we also choose to consider floating objects with elliptic shapes. Such a choice is of course arbitrary and may be adapted to alternative object's profiles. The reader is referred to Appendix § D for explicit formulae.

Remark 69. In order to minimize the total number of figures, we sometimes choose to display both the free-surface η and the discharge q on the same graphics. As the magnitudes of these two flow quantities are generally not similar, instead of directly plotting q, we choose to display the rescaled and translated quantity $\widetilde{q} := \frac{q}{H_0} + H_0$.

6.5.1 Dam-break

This first test-case is dedicated to the numerical assessment of the DG-ALE implementation with *a posteriori* LSC method for the NSW model and we consider classical dam-break problems over a flat bottom, without incorporating any partially immerged structure. Hence we only consider the formulation (6.43a) associated with the shallow-water equation, considering that $\mathcal{E}(t) = \Omega$, and replacing the coupling conditions at the free boundaries by classical homogeneous Neumann boundary conditions for the NSW equations. The computational domain is defined as $\Omega = [0, 1]$ and the initial data are defined as follows:

$$\eta_0(x) = \begin{cases} 1 & \text{if} \quad x \leq 0.5, \\ 0.5 & \text{elsewhere,} \end{cases}, \qquad q_0 = 0, \qquad b = 0.$$

We set $T_{\max} = 0.075\,s$, $n_{\text{el}} = 50$, $k = 3$ and for this particular test-case, the frame velocity is defined in a *pseudo-Lagrangien* way, directly connected to the fluid's local velocity as follows:

$$v_{g_{|i+\frac{1}{2}}} = \frac{1}{2}\left(u^{+}_{i+\frac{1}{2}} + u^{-}_{i+\frac{1}{2}} - \frac{1}{\sigma}\left(\mathbf{F}^{q}(\mathbf{v}^{+}_{i+\frac{1}{2}}, b_{i+\frac{1}{2}}) - \mathbf{F}^{q}(\mathbf{v}^{-}_{i+\frac{1}{2}}, b_{i+\frac{1}{2}})\right)\right), \tag{6.108}$$

where $u^{\pm}_{i+\frac{1}{2}}$ are the left and right traces of the fluid's velocity at the mesh interface $x_{i+\frac{1}{2}}$ and σ is defined as $\sigma := \max_{\omega\in\mathscr{T}_h} \sigma_\omega$, with

$$\sigma_\omega := \max_m \left(\sqrt{g\overline{H}^{\omega}_{m}}\right).$$

We show on Fig. 6.1-left a snapshot of the free surface at $t = 0.075\,s$ and we highlight the *corrected* and *uncorrected* subcells on the right. This illustrates very clearly that even if the correction has been activated on in a very sharp area in the vicinity of the discontinuity, the solution has still been cleansed from its spurious oscillations. We see also that, applying the ALE framework does not affect our *a posteriori* LSC method efficiency. In Fig. 6.2 we show the fourth-order *a posteriori* LSC solution without applying a moving grid. We notice that, by applying the ALE framework, the interfaces of the cells to the left of the shock, are the most active, which is expected, since in this zone, the velocity of the fluid is the highest.

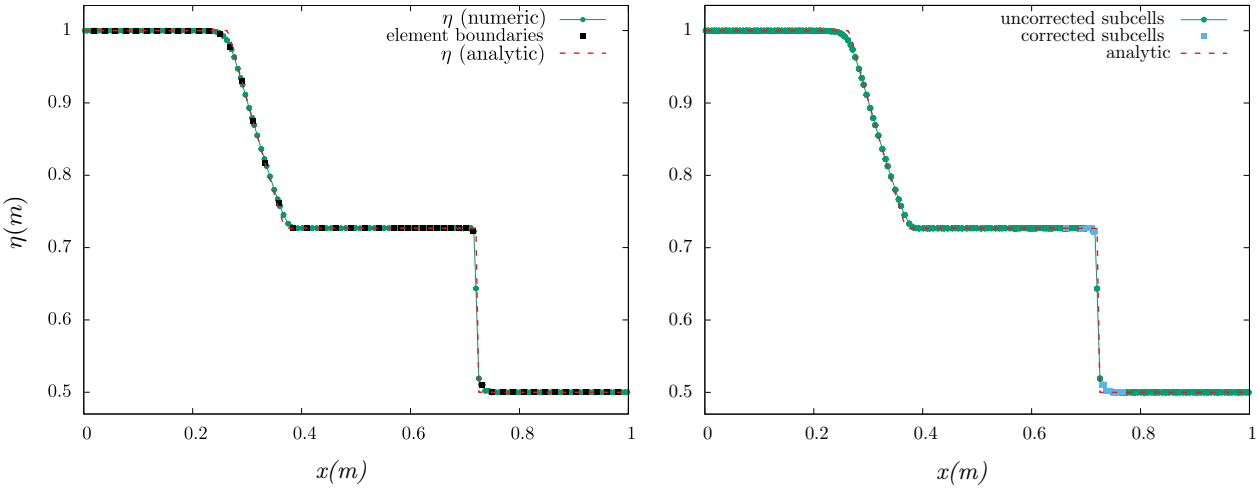

Figure 6.1: Test 24 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ for $k = 3$ and $n_{\text{el}} = 50$, with the *a posteriori* LSC with the ALE (6.108) framework, by marking the *corrected* and *uncorrected* subcells (right).

6.5.2 Well-balancing property

In this second test-case, we aim to assess the motionless steady-states preservation property for the model (6.43). This property has already been studied for the DG method with *a posteriori* LSC method in chapter 3. As a consequence, we only consider: (i) the case of a moving grid without incorporated floating object, (ii) the case of a partially immersed fixed object § 6.1.2 over a varying bottom (and later on, in § 6.5.8, we study the case of a partially immersed freely-floating object

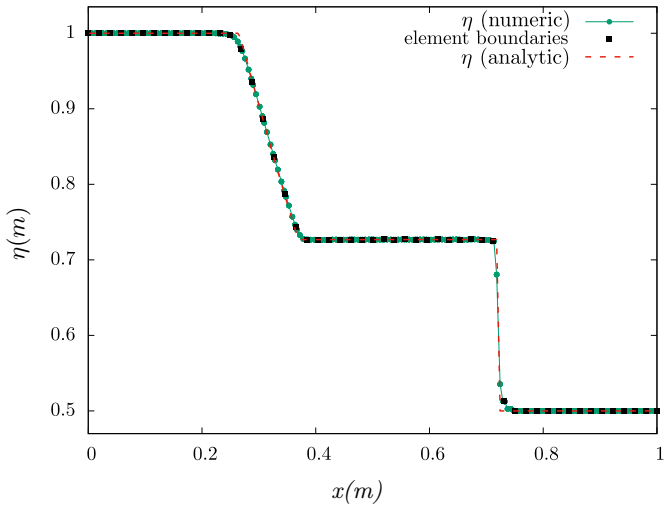

Figure 6.2: Test 24 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ for $k = 3$ and $n_{\text{el}} = 50$, with the *a posteriori* LSC without applying ALE.

§ 6.1.3 over a varying bottom). Firstly, we consider the computational domain $\Omega = [0, 1]$, together with a varying bottom defined as follows:

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{\pi\,(x - x_1)}{75}\right)\right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere,} \end{cases} \tag{6.109}$$

where $A = 4.75$, $x_1 = 0.125\,m$ and $x_2 = 0.875\,m$. The initial data is defined as:

$$\eta_0 := 10, \qquad q_0 := 0,$$

see Fig. 6.3. We set $k := 3$, $n_{\text{el}} := 50$ and the frame's velocity is defined in a *pseudo-Lagrangien* way (6.108). We evolve this initial configuration in time up to 4×10^5 time iterations ($50\,s$) and we observe that this initial configuration is preserved up to the machine accuracy, see again Fig. 6.3. Here the grid is not moving since the velocity of the grid is related to the velocity of the fluid, which is zero (steady state), see (6.108). As shown in the well-balanced property proof, the grid can move freely with any arbitrary velocity, while always retaining the well-balanced property of the scheme, as long as we are considering the totally wet context. In particular, we consider a uniform grid velocity with $v_{\text{g}} = 0.01\ \text{m} \cdot \text{s}^{-1}$. We can observe that the mesh grid translates with a distance of $0.5\,m$ after $50\,s$, while always respecting the well-balanced property of the scheme, see Fig. 6.4. A similar behavior is reported for other values of k, n_{el} and v_{g}.
In a second configuration, we introduce a partially immersed fixed object § 6.1.2, together with a dry area, corresponding to a plane sloping beach. The computational domain is $\Omega = [-50, 200]$ and the object is located at $(x_G, z_G) = (50, H_0 + 2.5)$. The topography profile is defined as follows:

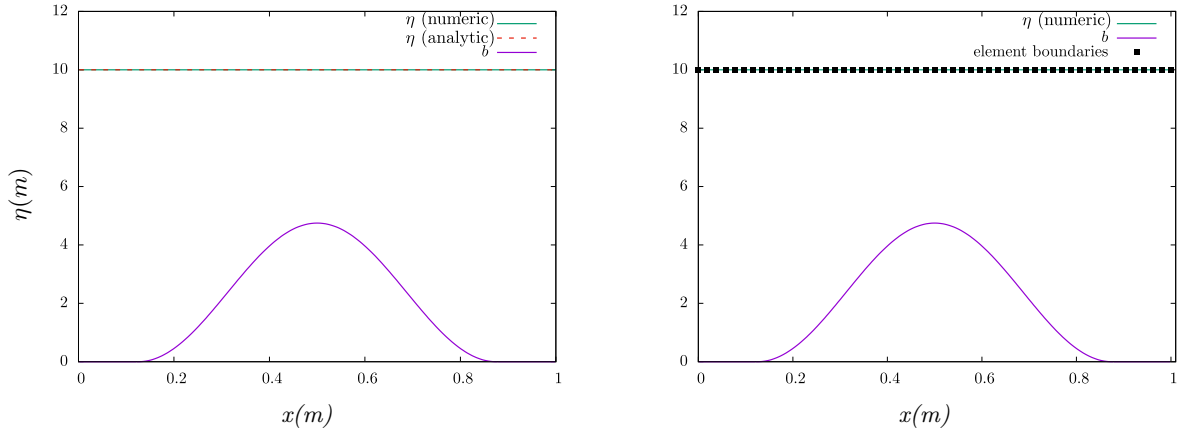

Figure 6.3: Test 25 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$ (left), showing element boundaries (right).

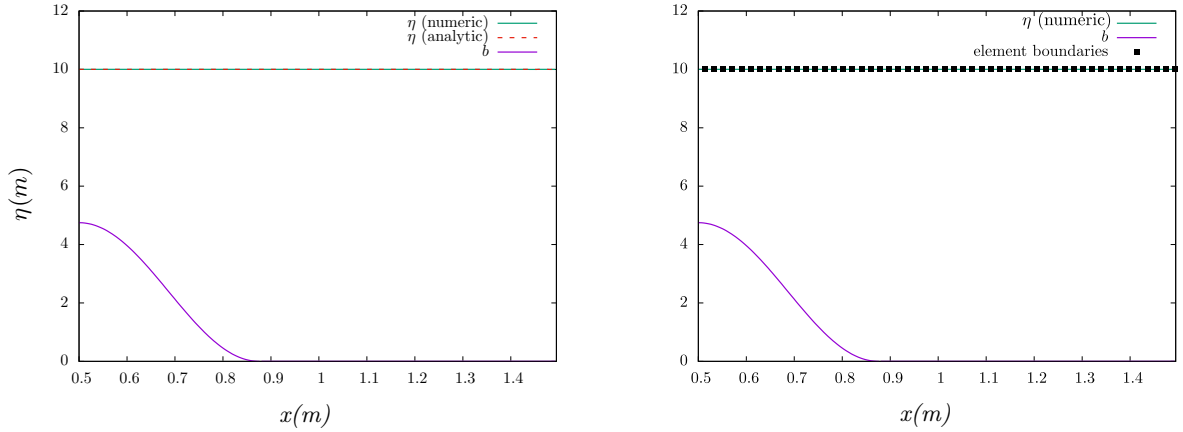

Figure 6.4: Test 25 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$ (left), showing element boundaries (right).

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{\pi\,(x - x_1)}{75}\right)\right)^2 & \text{if} \quad x_1 \le x \le x_2, \\ \dfrac{1}{\beta}\,(x - x_3) & \text{if} \quad x \ge x_3, \\ 0 & \text{elsewhere,} \end{cases} \tag{6.110}$$

where $A = 1.5\,m$, $\beta = 11$, $x_1 = 12.5\,m$, $x_2 = 87.5\,m$ and $x_3 = 90\,m$. The initial data in $\mathcal{E}_0$ is defined as

$$\eta_0^e(x) := \max\left(5, b(x)\right) \qquad \text{and} \qquad q_0^e = 0,$$

while in the interior domain $\mathcal{I}_0$, we have

$$\eta_0^i := p^k_{\mathscr{T}_h^{i,0}}(\eta_{\text{lid}}) \qquad \text{and} \qquad q_0^i = 0.$$

We evolve this initial configuration up to $T_{\max} = 50\,s$, with $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$. The numerical results obtained with the DG-ALE scheme using the *a posteriori* LSC method are shown on Fig. 6.5, with a zoom in the vicinity of the immersed structure and the shoreline. The corrected and uncorrected subcells are exhibited on Fig. 6.6.

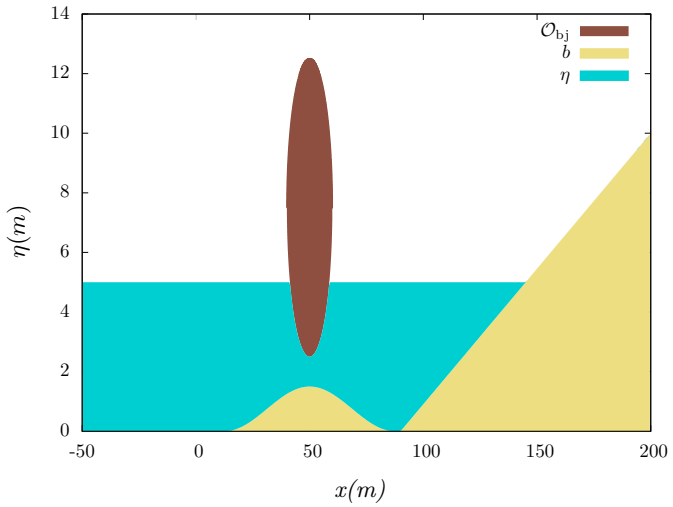

Figure 6.5: Test 26 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$ for $k = 3$ and $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$.

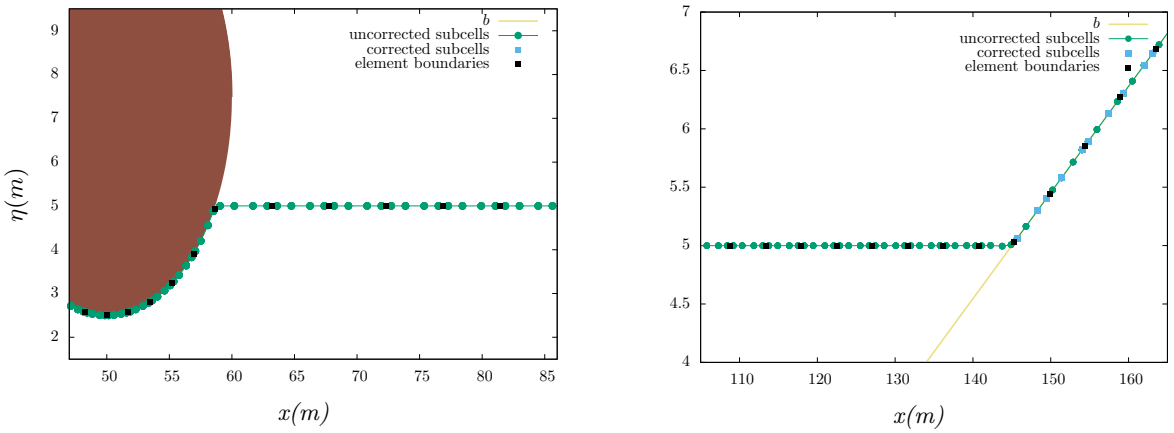

Figure 6.6: Test 26 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$, with a zoom near the floating body (left) and near shore (right), showing the corrected and the uncorrected subcells, for $k = 3$ and $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$.

6.5.3 A solitary wave interacting with a stationary partially immersed object

In this third test-case, we focus on the interactions between a weakly nonlinear solitary wave propagating towards a stationary partially immersed objectover a varying topography made of a bump

followed by a sloping beach. We consider the computational domain $\Omega = [0, 100]$ and $H_0 = 5\,m$. The topography profile is defined as follows:

$$b(x) := \begin{cases} A_b \left(\sin\left(\dfrac{\pi\,(x - x_1)}{75} \right) \right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere,} \end{cases} \tag{6.111}$$

where $A_b = 1.5\,m$, $x_1 = 12.5\,m$ and $x_2 = 87.5\,m$. A stationary partially immersed object is placed over the bump and the initial data is prescribed as follows:

$$\eta_0^e(x) := H_0 + A_w \operatorname{sech}\left(\gamma\,(x - x_0)\right), \qquad q_0^e(x) := \mathrm{g} c_{q_2} (\eta_0^e(x) - H_0) H_0^e,$$

and

$$\eta_0^i := p^k_{\mathscr{T}_h^{i,0}}(\eta_{\text{lid}}), \qquad q_0^i := 0,$$

where $A_w = 0.35\,m$, $c_{q_1} = 1$, $c_{q_2} = 0.5$, $\gamma := c_{q_1}\sqrt{\dfrac{3A_w}{4H_0}}$ and $x_0 = 20\,m$ stands for the initial location of the solitary wave's center. The elliptic object is defined with respective horizontal and vertical radius $a = 10\,m$ and $b = 5\,m$, and its center of mass is located at $(x_{\mathrm{G}}, z_{\mathrm{G}}) = (50, H_0 + 2.5)$ (see § D for the explicit definition of the object and the parameterization of its underside). We set $n_{\text{el}}^e = 50$, $n_{\text{el}}^i = 10$ and $k = 3$. Snapshots of the free surface at various times during the propagation are shown on Fig. 6.7. Interestingly, we observe a partial run-up, run-down and reflection on the object's left side. This reflected wave goes back towards the inlet boundary, while the transmitted wave propagates further beyond the object into the right exterior domain. Finally, both reflected and transmitted waves are evacuated from the computational domain. In Table 6.1, we gather the global L^2-errors obtained for several numbers of elements with $k = 3$ for inner pressure at $t_{max} = 100s$. We show in Fig. 6.8 the shape of the error curve for $n_{\text{el}} = 100$ in the range of time $[0\,s, 100\,s]$.

h	$E^{\mathrm{p_i}}_{L_2}$
100/20	1.72E-3
100/50	8.07E-4
100/100	1.25E-5
100/150	4.40E-6
100/200	1.33E-7

Table 6.1: Test 27 - A solitary wave interacting with a stationary partially immersed object: L^2-errors between numerical and initial solutions for inner pressure $\underline{\mathrm{p}}_{\mathrm{i}}$, $k = 3$ at time $t_{max} = 100\,s$.

To complete the picture, we also investigate the conservation of total mass

$$M_{\text{total}}(t) = \int_\Omega H(x,t)\,dx \tag{6.112}$$

and total energy

$$E_{\text{total}}(t) = E_c(t) + E_p(t) = \int_\Omega \frac{H(x,t)u^2(x,t)}{2}\,dx + \int_\Omega \mathrm{g} H(x,t)(b(x) + \frac{H(x,t)}{2})\,dx \tag{6.113}$$

over time. We compute the time evolution of their relative errors

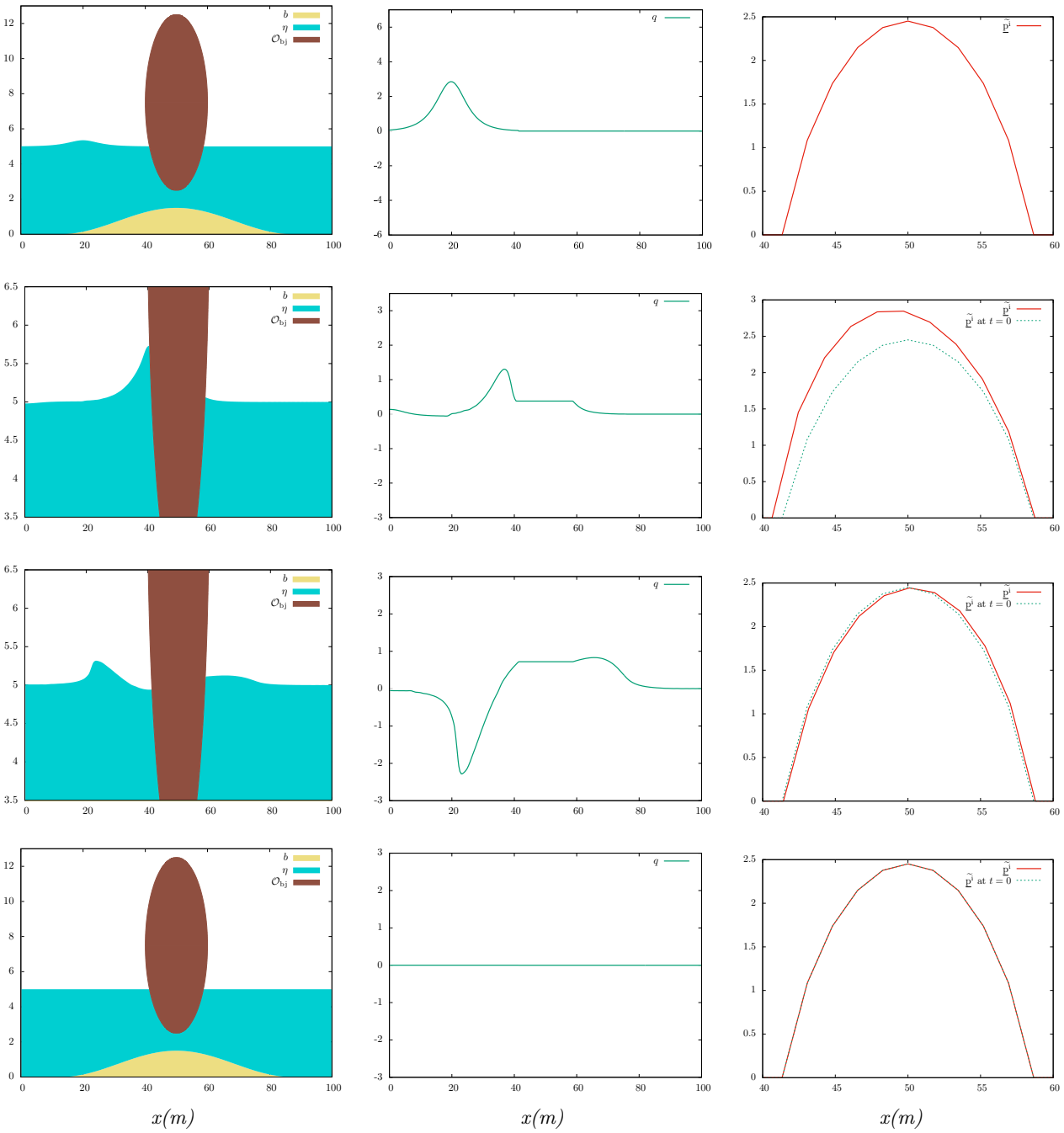

Figure 6.7: Test 3 - A solitary wave interacting with a stationary partially immersed object: surface elevation η, discharge q and inner pressure $\widetilde{\underline{p}^{i}} = \underline{p}^{i}/\rho g$ at the underside of the object for $n_{el}^{e} = 50$, $n_{el}^{i} = 10$ and $k = 3$.

$$E_r^{M_{\text{total}}} = |\frac{M_{\text{total}}(t) - M_{\text{total}}(0)}{M_{\text{total}}(0)}| \quad \text{and} \quad E_r^{E_{\text{total}}} = |\frac{E_{\text{total}}(t) - E_{\text{total}}(0)}{E_{\text{total}}(0)}|, \tag{6.114}$$

considering always the same test-case, but this time we impose "wall" boundary conditions for both

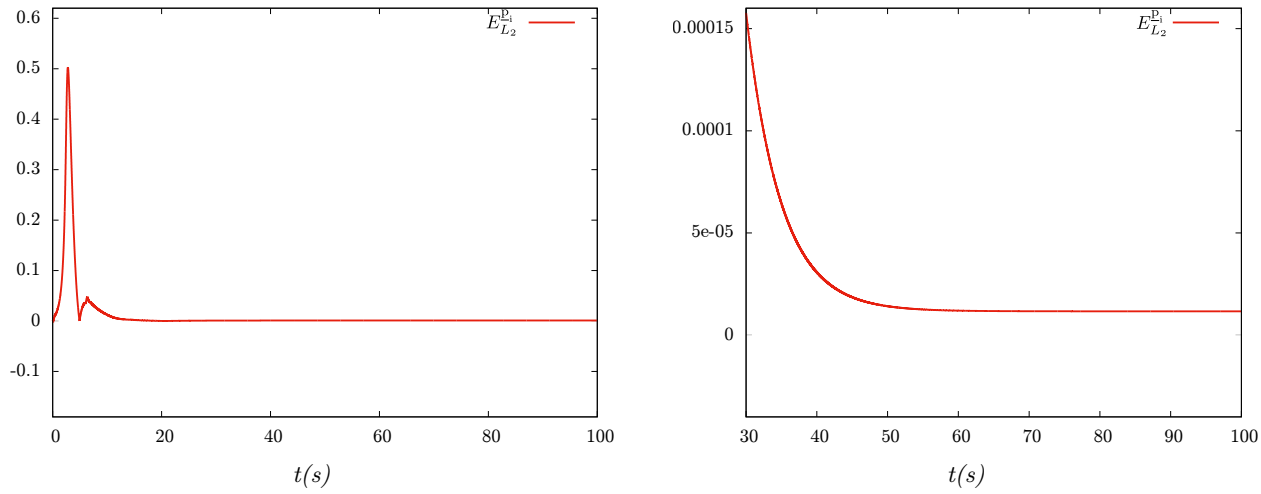

Figure 6.8: Test 27 - A solitary wave interacting with a partially immersed stationary object: L^2-errors between numerical and initial solutions for inner pressure $\underline{p}_i$ for $n^e_{\text{el}} = 80$, $n^i_{\text{el}} = 20$ and $k = 3$ in the range $[0\,s, 100\,s]$.

left and right exterior boundaries, in order to prevent water from getting out of the computational domain, see Fig. 6.9 and 6.10. As expected, the total mass and the total energy are preserved up to the machine accuracy. In table 6.2, we gather the total mass relative errors for several orders of approximation at $t = 20s$. Similar results are obtained for the total energy relative errors.

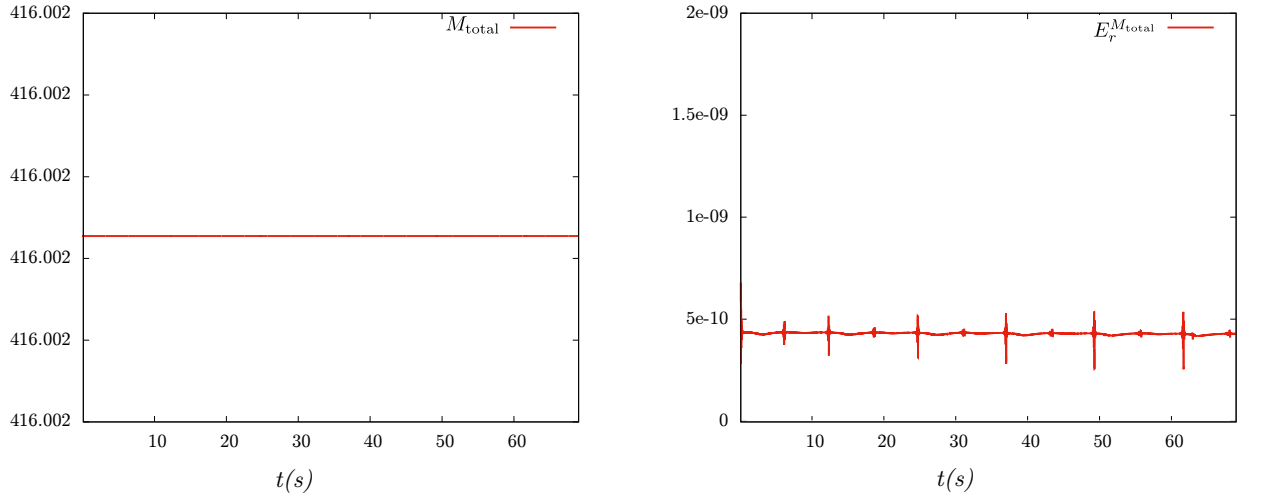

Figure 6.9: Test 27 - A solitary wave interacting with a partially immersed stationary object: total mass and relative error, for $k = 3$, $n_{\text{el}} = 200$ in the range $[0\,s, 70\,s]$.

6.5.4 A shock-wave interacting with a partially immersed stationary object

In this fourth test-case, we consider a shock-wave propagating over a flat bottom, with a stationary object partially immersed in the middle of the computational domain, in order to emphasize the

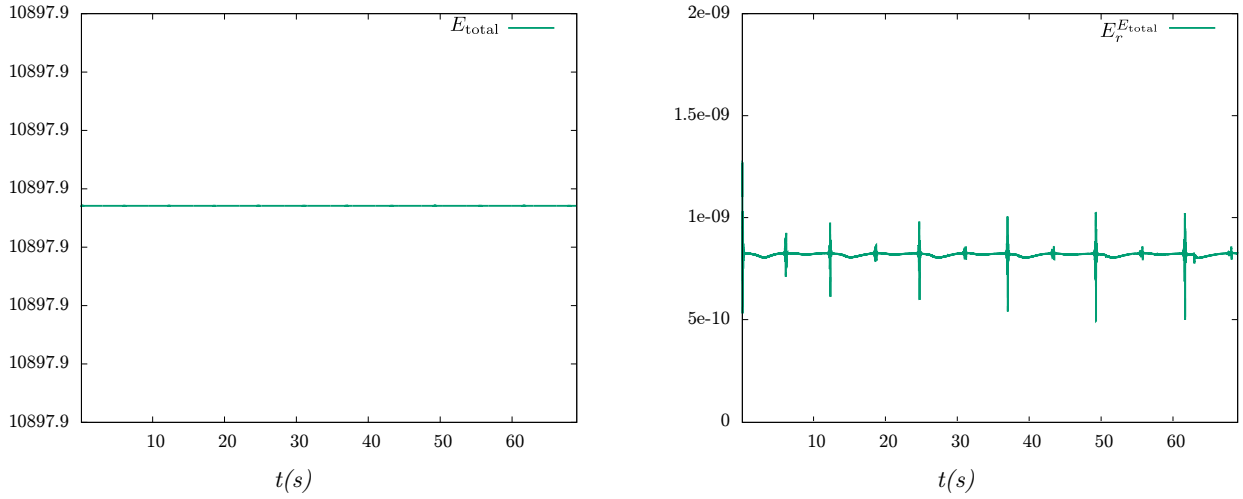

Figure 6.10: Test 27 - A solitary wave interacting with a partially immersed stationary object: total energy and relative error, for $k = 3$, $n_{\text{el}} = 200$ in the range $[0\,s, 70\,s]$.

k	1	2	3
h	$E_r^{M_{\text{total}}}$	$E_r^{M_{\text{total}}}$	$E_r^{M_{\text{total}}}$
100/50	4.0523E-5	8.7551E-7	9.9301E-8
100/100	5.5104E-6	1.5817E-7	1.0848E-8
100/200	4.0289E-7	7.2318E-9	4.3317E-10
100/400	2.9140E-08	3.1734E-9	2.8571E-11

Table 6.2: Test 27 - A solitary wave interacting with a partially immersed stationary object: total mass relarive error for several order of approximations at time t = 20 s.

robustness of the discrete formulation. We set $\Omega := [-20, 120]$, $n_{\text{el}}^e := 70$, $n_{\text{el}}^i := 10$ and $k := 3$. The initial data is defined as follows:

$$\eta_0^e(x) := \begin{cases} 6.5 & \text{if} \quad x \leq 0, \\ 5 & \text{elsewhere,} \end{cases}, \quad \eta_0^i := p^k_{\mathscr{T}_h^{i,0}}(\eta_{\text{lid}}), \qquad q_0^e = q_0^i := 0.$$

The discontinuity, initially located in $\mathcal{E}_0^-$, propagates towards the object, generating some interesting interactions. We emphasize that the proposed configuration simultaneously involves a shock wave reflection, some displacement of the frame, a partial run-up over the surface piercing object associated with the collision between the shock wave and the object profile and a partial transmission of the wave towards the right side of the exterior domain, with the formation of an interesting pattern that shares some similarities with a rarefaction wave. Snapshots of the free surface elevation at several time are shown on Fig. 6.11. In Fig. 6.12, we show the corrected and uncorrected subcells which are respectively plotted with blue squares and green dots. This clearly highlights the high robustness of the discrete DG-ALE formulation, and in particular of the *a posteriori* LSC method , as the dynamic of the free boundaries is computed in a very stable way, without any spurious oscillations or further time-step restriction. Additionally, a zoom on the surface discontinuity is shown (in Fig. 6.12), highlighting that the correction has been activated on in a very sharp area in the vicinity of the discontinuity.

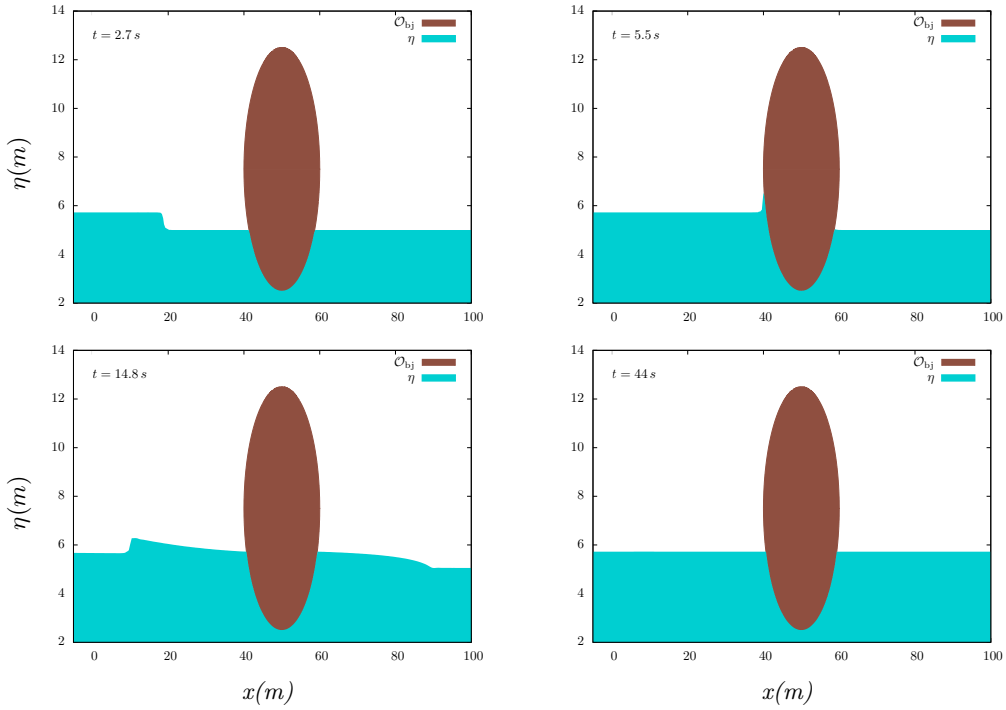

Figure 6.11: Test 28 - A shock-wave interacting with a partially immersed stationary object - Free surface elevation computed for different values of time $t = 2.7\,s$, $5.5\,s$, $14.8\,s$ and $44\,s$ respectively for $k = 3$, $n^e_{\text{el}} = 70$ and $n^i_{\text{el}} = 10$.

6.5.5 Run-up of a solitary wave partially reflected by a stationary object

In this test-case, we follow the propagation and run-up of a solitary wave over a plane beach, with a stationary partially immersed object placed on the way. The computational domain is set to $\Omega = [-200, 150]$ and the topography is made of a constant depth area followed by a sloping beach of constant slope $1/11$. We set $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$. The initial data is defined as follows:

$$\eta^e_0(x) := H_0 + A_w \operatorname{sech}(\gamma(x - x_0)), \quad q^e_0(x) := c_{q_2} \mathrm{g}(\eta^e_0 - H_0) H^e_0, \tag{6.115}$$

where $A_w := 0.55\,m$, $x_0 := -80\,m$, $c_{q_1} := 0.1$, $c_{q_2} := 0.5$, $\gamma := c_{q_1}\sqrt{\dfrac{3A_w}{4H_0}}$ and

$$\eta^i_0 := p^k_{\mathscr{T}^{i,0}_h}(\eta_{\text{lid}}), \qquad q^i_0 := 0.$$

We show on Fig. 6.13 some snapshots of the free surface elevation at several discrete times in the range $[0.57\,s, 300\,s]$. We observe a partial run-up and a reflection of the wave on the object, while the remaining part of the wave is transmitted beyond the object, propagating further in $\mathcal{E}(t)$. This secondary wave subsequently reaches the shore, generating a run-up on the beach, followed by a full reflection. The reflected wave is itself again partially reflected by the object, generating a third sequence of run-up and reflection, while the transmitted wave propagates back in $\mathcal{E}^-(t)$ towards the domain's left boundary. In Fig. 6.14, we zoom on the shoreline area, highlighting the corrected and uncorrected subcells which are respectively plotted with blue squares and green dots. Again, we

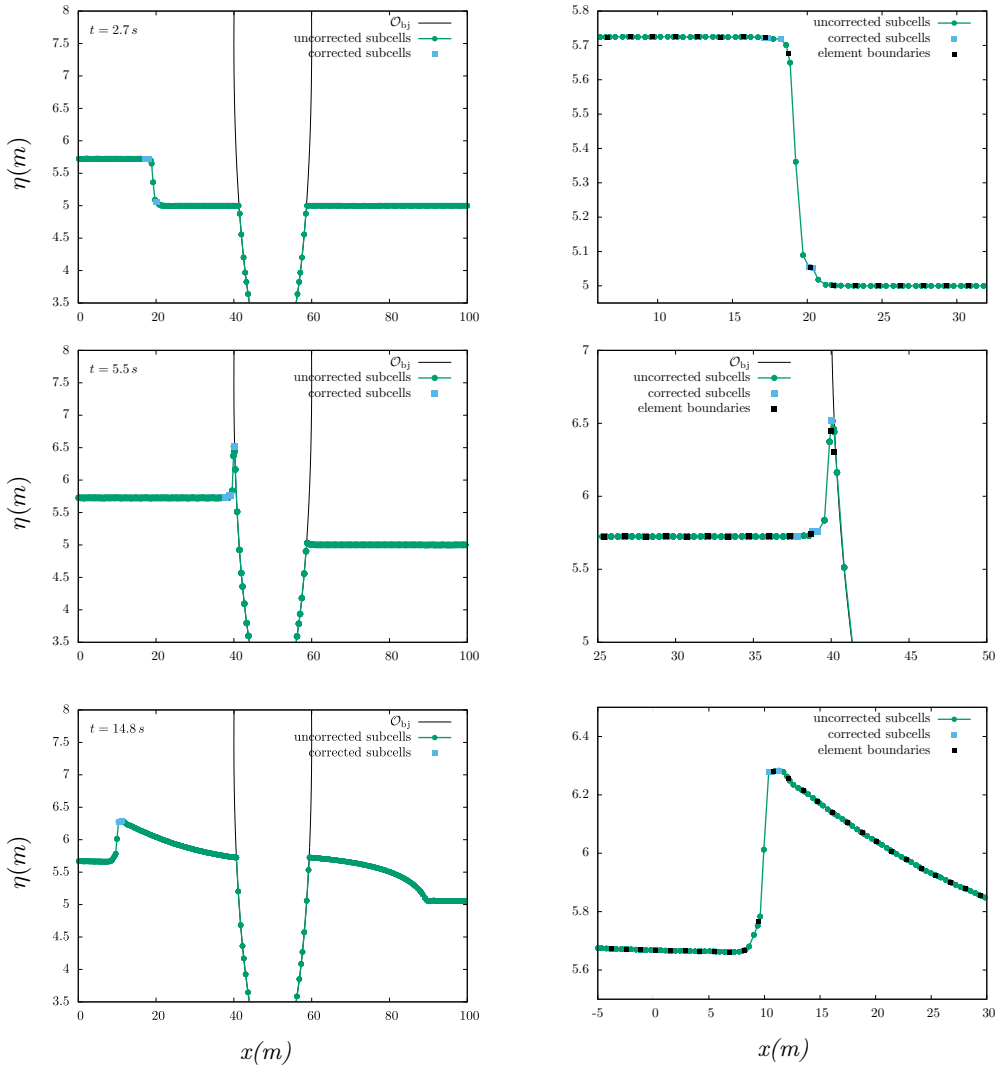

Figure 6.12: Test 28 - A shock-wave interacting with a stationary partially immersed object- Free surface elevation computed for different values of time $t = 2.7\,s$, $5.5\,s$ and $14.8\,s$, respectively: corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on discontinuity, for $k = 3$, $n_{\text{el}}^e = 70$ and $n_{\text{el}}^i = 10$.

observe that the *a posteriori* LSC method is only activated in a very thin area in the vicinity of the wet/dry front. We also display on Fig. 6.15 a comparison between the maximum run-up observed with the partially immersed object on place, and without the object, in order to highlight the impact of the object presence on the wave height.

6.5.6 Validations of prescribed motions

Let consider now model (6.48) for a moving object and some prescribed motions which are validated separately along each motion's direction.

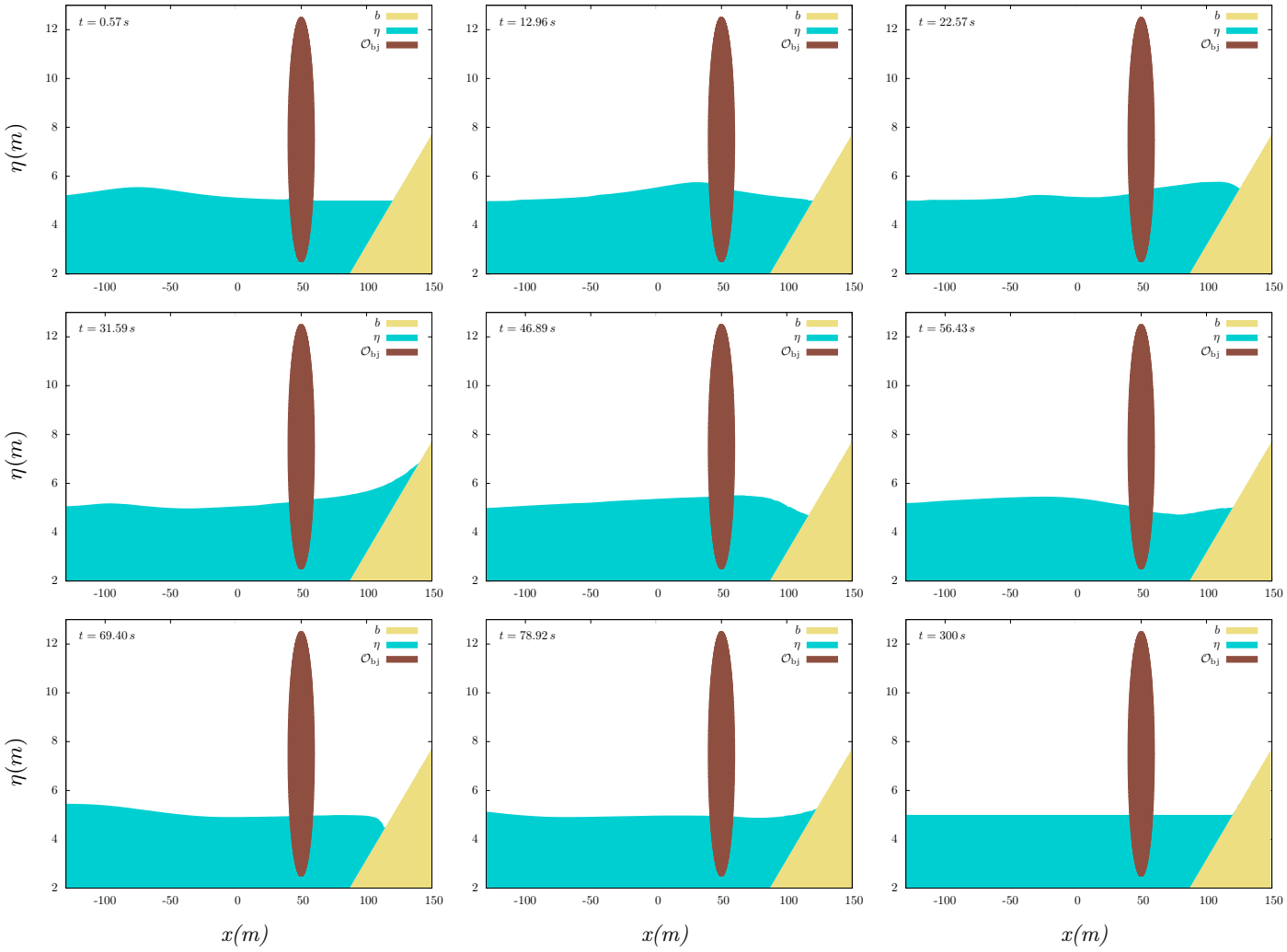

Figure 6.13: Test 29 - Run-up of a solitary wave partially reflected by a stationary object - Free-surface obtained at several times in the range $[0.57\,s,\ 300\,s]$, with $k = 3$ and $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$.

Pure heaving

This test-case focus on a prescribed periodic vertical motion. $\mathcal{M}_G$ is initially placed at $(x_G(0), z_G(0)) = (50, H_0 + 2.5)$ and the flow configuration is a lake at rest that extends for $180\,m$ in both directions, with a constant depth $H_0 = 5$. The initial data in the *interior* and *exterior* regions are defined by:

$$\eta^e_0 = H_0,\ \ \eta^i_0 = p^k_{\mathscr{T}^{i,0}_h}(\eta_{\text{lid}}) \qquad \text{and} \qquad q^e_0 = q^i_0 = 0.$$

The parametric equations describing the motion of $\mathcal{M}_G$ are defined as follows:

$$x_G(t) = 50 \ \text{ and } \ z_G(t) = H_0 + 3 - \frac{1}{2}\cos(\frac{2\pi t}{15}).$$

We set $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$, and $k = 3$. The numerical solution is shown in Fig. 6.16 for several time-steps, $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ respectively, where $T = 15\,s$ is the time-period of

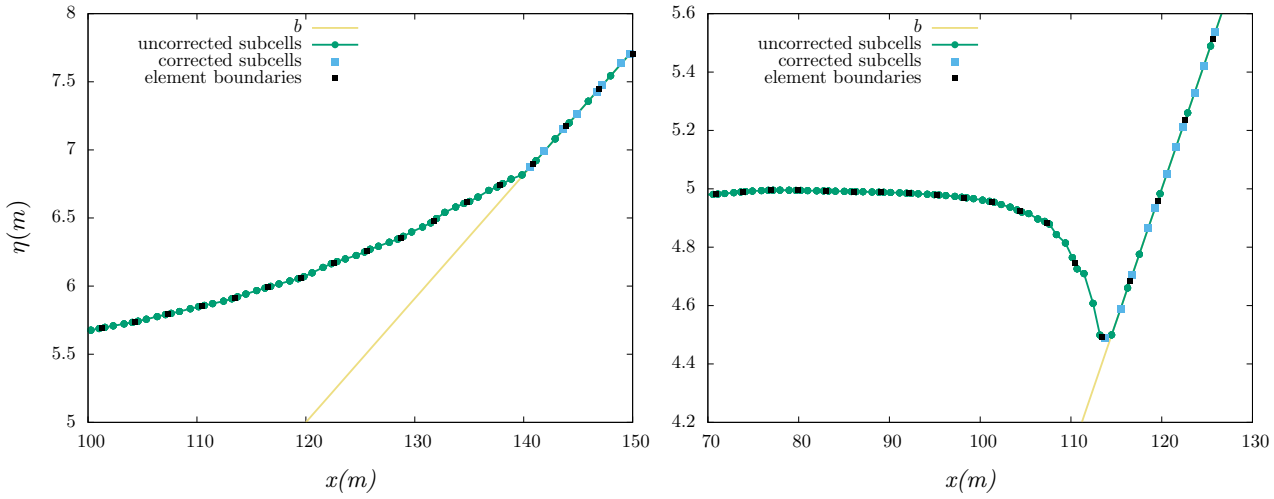

Figure 6.14: Test 29 - Run-up of a solitary wave partially reflected by a stationary object - A zoom on the shoreline showing the free surface at $t = 31.59\,s$ (left) and $t = 69.40\,s$ (right), where corrected and uncorrected subcells are respectively plotted with blue squares and green dots, for $k = 3$ and $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$.

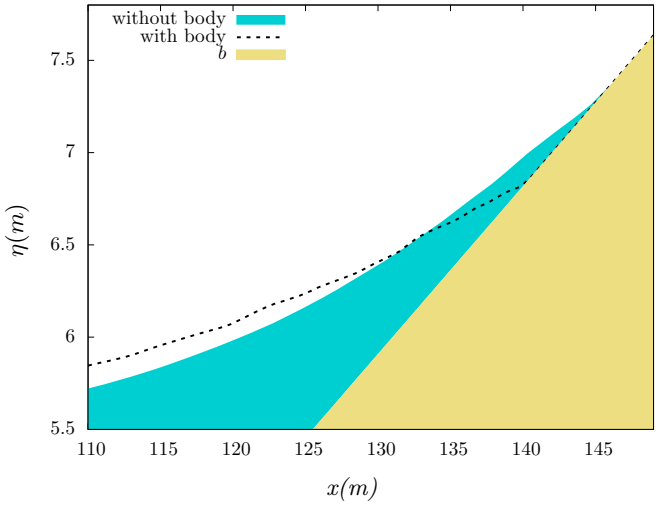

Figure 6.15: Test 29 - Run-up of a solitary wave - Snapshot of the free surface corresponding to the maximum observed run-up with the embedded surface-piercing object (in dashed-line) and without (in blue).

the structure's motion. The Figure shows the variation of the free surface and the discharge in the *interior* and *exterior* regions (left), with a zoom on the displacement of the mesh nodes near $\chi_{\pm}$ (right).

We also investigate the conservation of total mass and total energy over time, and we compute the time evolution of their relative errors (see (6.112)-(6.113)-(6.114) for their definitions), see Fig. 6.17 and 6.18. The total mass and the total energy are preserved over time up to order 10^{-10} and 10^{-8}

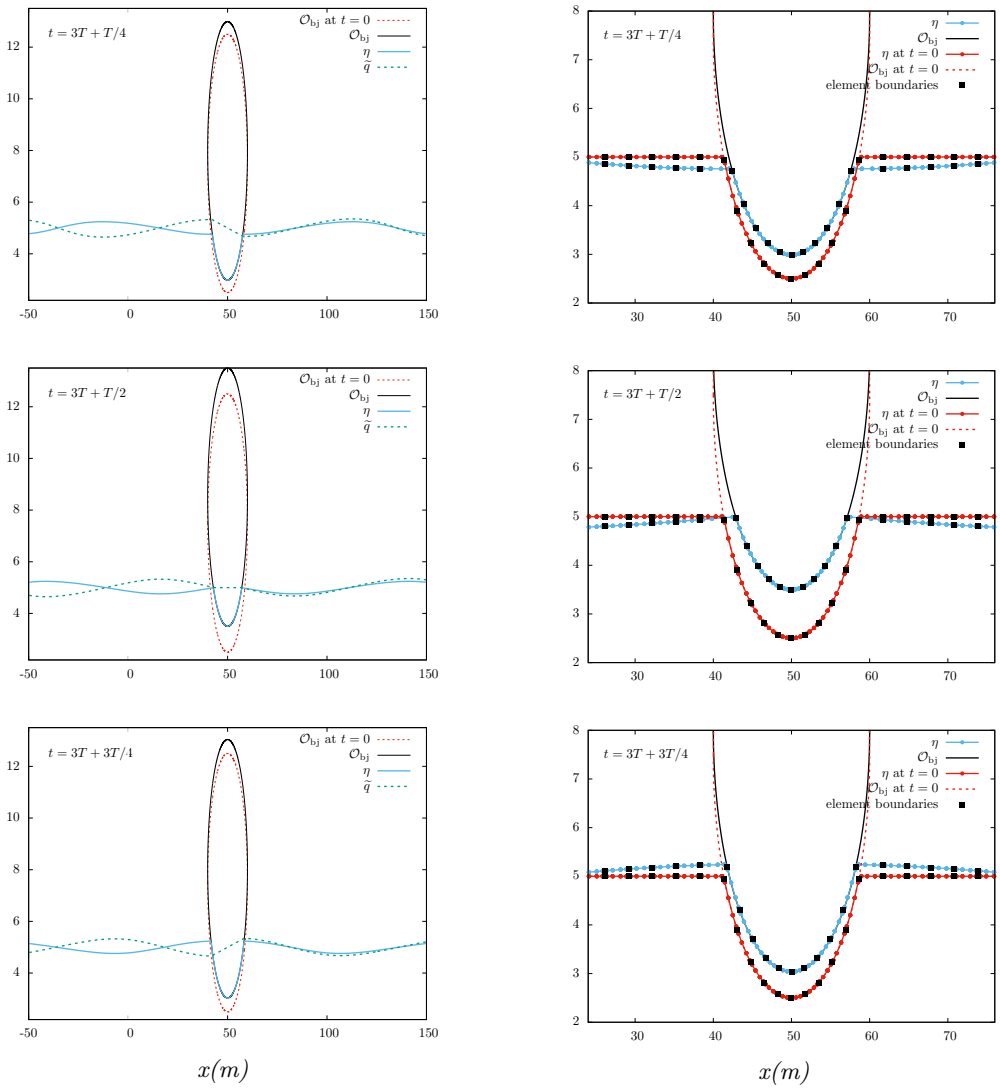

Figure 6.16: Test 30 - Prescribed motion: heaving - Free surface elevation and discharge computed for different values of time $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ (left) with a zoom showing the displacement of the mesh nodes near contact points $\chi_{\pm}$ (right), for $k = 3$, $n_{\text{el}}^e = 50$ and $n_{\text{el}}^i = 10$.

respectively.

Pure surging

Now, a purely periodic horizontal motion is enforced. The evolution equations of $\mathcal{M}_G$ are:

$$x_G(t) = 48 + 2\cos(\frac{2\pi t}{10}) \quad \text{and} \quad z_G(t) = H_0 + 2.5.$$

The solution is shown in Fig. 6.19 for several time-steps, $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ respectively, where $T = 10\,s$ is the time-period of the body motion. The figure shows the variations of the free surface and the discharge in the *interior* and *exterior* regions (left), with a zoom on the displacement

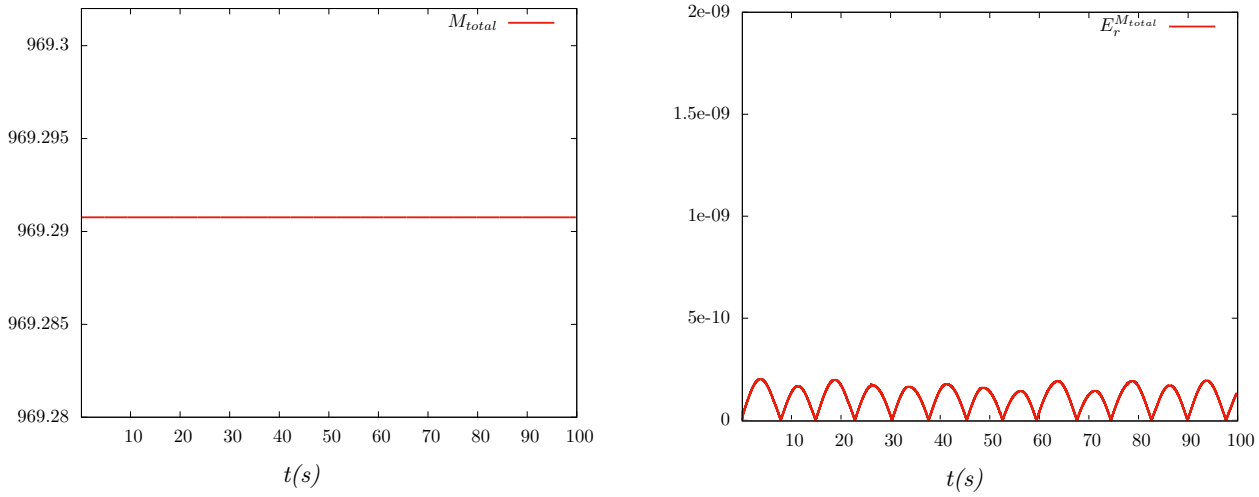

Figure 6.17: Test 30 - Prescribed motion: heaving - total mass and relative error, for $k = 3$, $n_{\text{el}} = 200$ in the range $[0\,s, 100\,s]$.

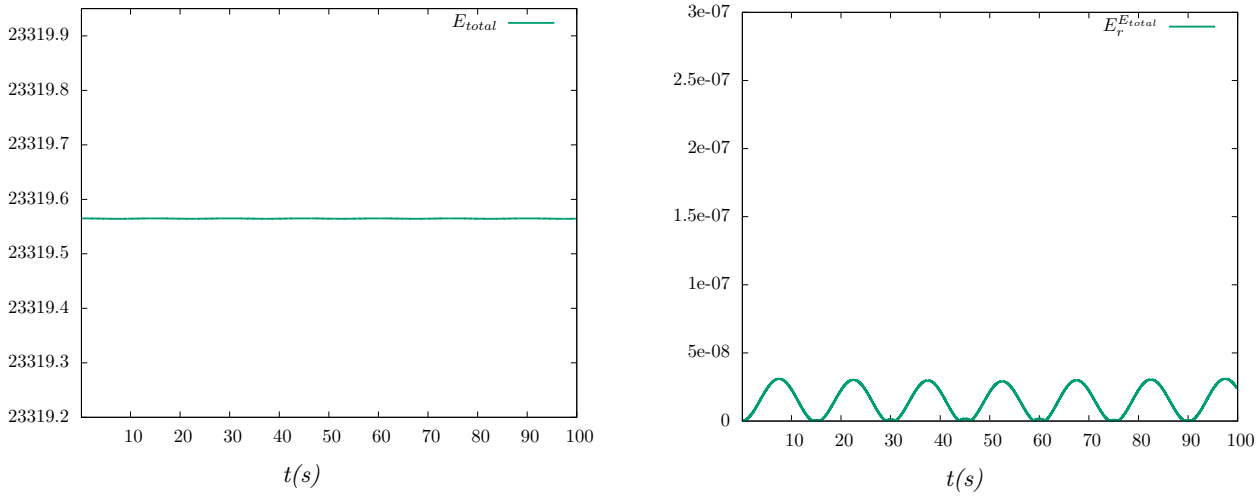

Figure 6.18: Test 30 - Prescribed motion: heaving - total energy and relative error, for $k = 3$, $n_{\text{el}} = 200$ in the range $[0\,s, 100\,s]$.

of the mesh nodes in the vicinity of $\chi_{\pm}$ (right).

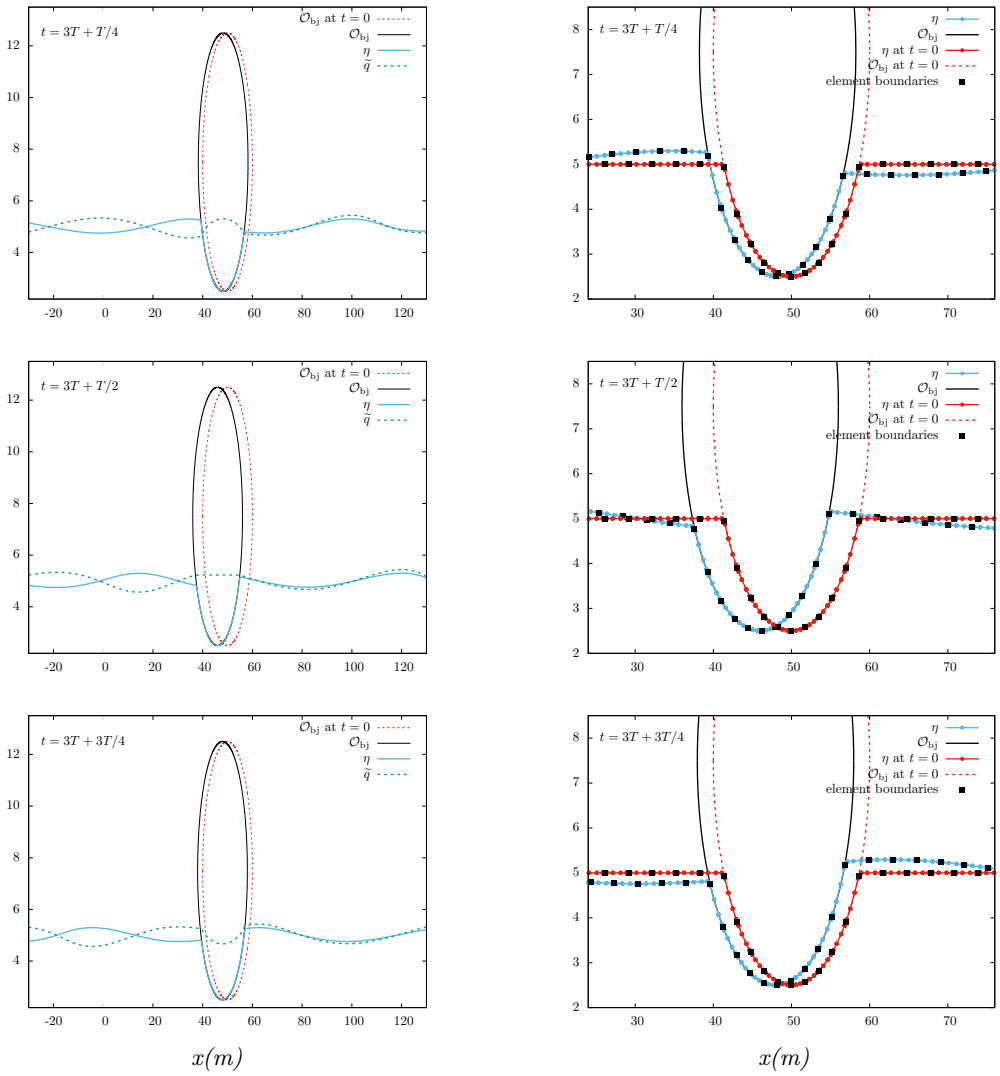

Figure 6.19: Test 31 - Prescribed motion: surging - Free surface elevation and discharge computed for different values of time $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ (left) with a zoom showing the displacement of the mesh nodes near contact points $\chi_{\pm}$ (right), for $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

Pure pitching

Now, we consider a prescribed periodic rotational motion.The angle's evolution law is given by:

$$\theta(t) = \frac{\pi}{25}\sin(\frac{2\pi t}{8}).$$

The solution is shown in Fig. 6.20 for several time-steps, $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ respectively, with $T = 8\,s$ is the time-period of the body motion. In Fig. 6.21, we show the horizontal coordinates of the contact points $\chi_{-}(t)$ and $\chi_{+}(t)$ in the time range $[0, 6T]$.

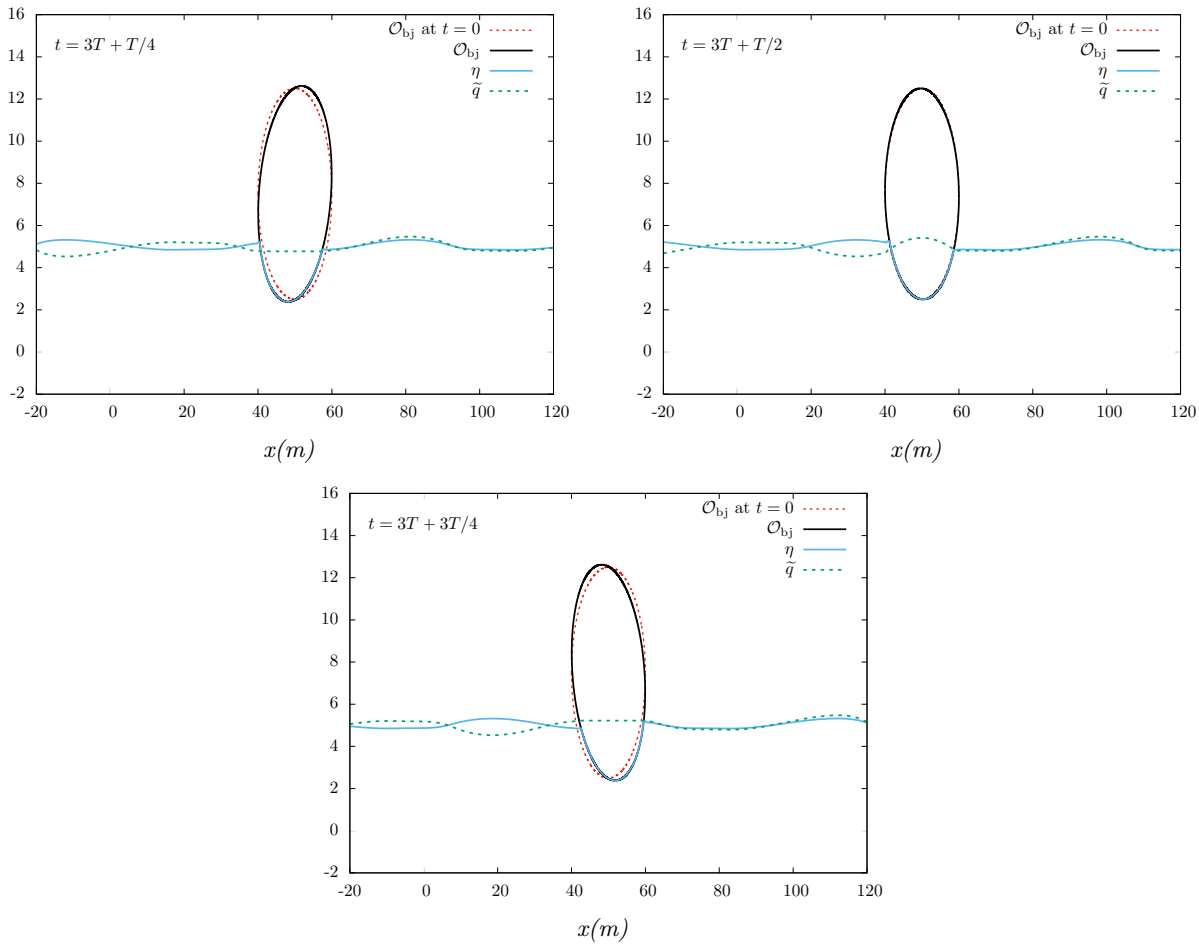

Figure 6.20: Test 32 - Prescribed motion: pitching - Free surface elevation and discharge computed for different values of time $t = 3T + \frac{T}{4}$, $3T + \frac{T}{2}$ and $3T + \frac{3T}{4}$ for $k = 3$, $n^e_{\rm el} = 50$ and $n^i_{\rm el} = 10$.

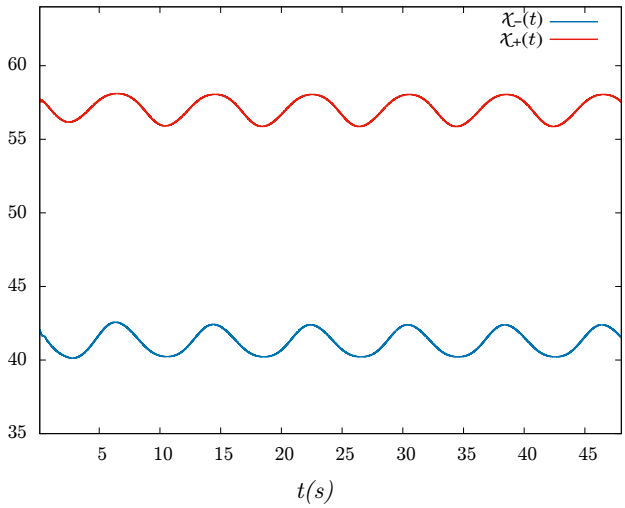

Figure 6.21: Test 32 - Prescribed motion: pitching - Variation of the horizontal coordinate of the contact points $\chi_-(t)$ and $\chi_+(t)$

6.5.7 Prescribed motion: heaving with a wet-dry transition

In this test-case, we qualitatively analyze the water-body interactions in the situation of a pure prescribed heaving, but this time also with a dry area. The purpose of this test-case is to highlight the scheme's ability to deal with wet-dry transitions in presence of a floating body, hence validating the DG-ALE formulation with *a posteriori* LSC method. The topography profile is defined as follows:

$$b(x) = \begin{cases} \frac{1}{\beta}(x - x_\beta) & \text{if } x < x_\beta, \\ 0 & \text{elsewhere}, \end{cases}$$

where $\beta = 11$ and $x_\beta = 65$. The right boundary condition is transmissive. As for the left boundary, we impose $\eta^e = 5\,m$ and $q^e = 0\ m^2.s^{-1}$. The initial condition is defined as follows:

$$\eta_0^e = \max(H_0 - b, 0) + b,\ \eta_0^i = p^k_{\mathscr{T}_h^{i,0}}(\eta_{\text{lid}}) \qquad \text{and} \qquad q_0^e = q_0^i = 0.$$

The elliptic object's location is initialized with $(x_G(0), z_G(0)) = (50, H_0 + 2.5)$ and we set $H_0 = 5$. We prescribe a periodic heave motion as follows:

$$x_G(t) = 50 \ \text{ and } \ z_G(t) = H_0 + 2.75 - \frac{1}{4}\cos(\frac{2\pi t}{20}).$$

We set $n^e_{\text{el}} = 60$, $n^i_{\text{el}} = 10$, and $k = 3$. The corresponding numerical results are shown on Fig. 6.22 for several time-steps, $t = 53\,s$, $60.5\,s$ and $66\,s$ respectively. This figure shows the variation of the free-surface (left), and the associated corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the shoreline (right).
We observe that the periodic heaving generates free-surface waves that propagate seaward and shoreward, with some run-up and run-down on the shore. After a transient phase, the coupled system converges towards a quasi-stationary periodic regime. The run-up is computed in a very stable way, thank's to the *a posteriori* LSC method.

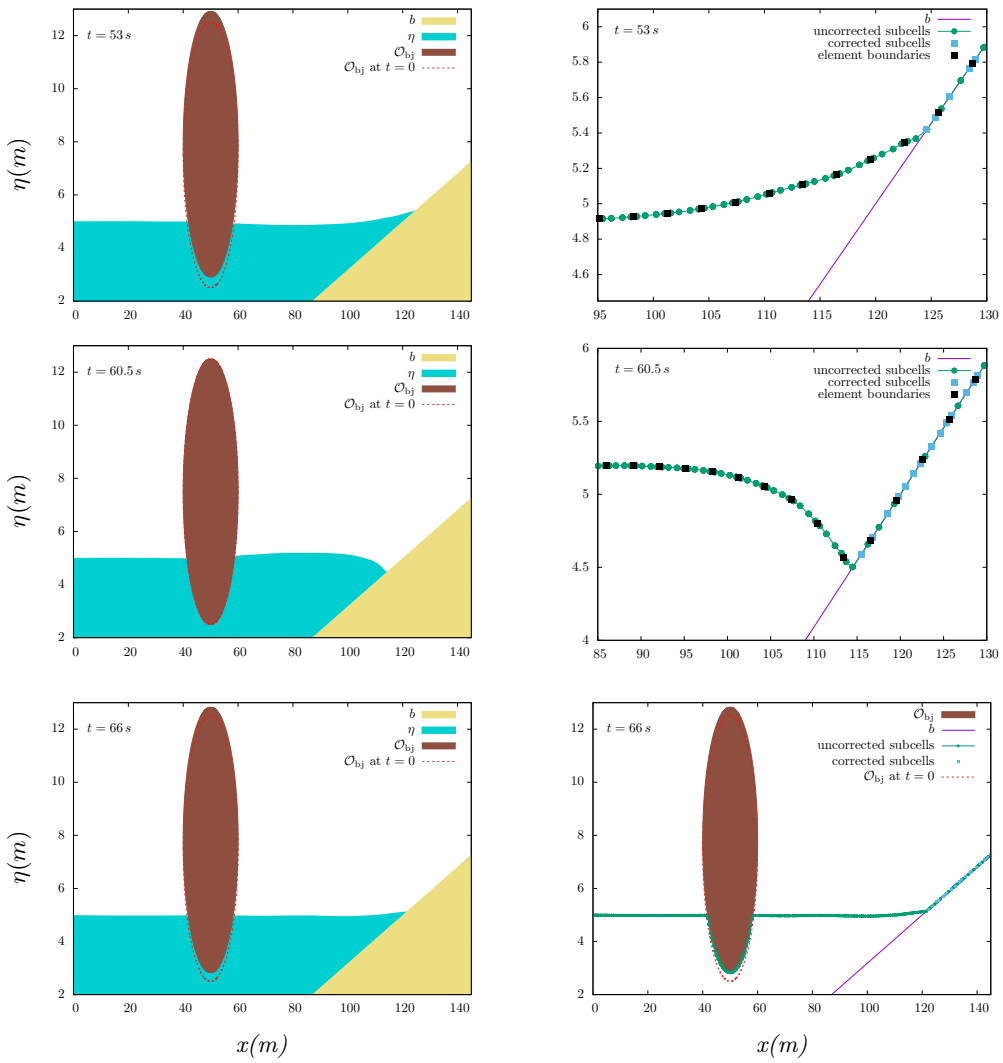

Figure 6.22: Test 33 - Prescribed motion: heaving with a wet-dry transition - Free surface elevation computed for different values of time $t = 53\,s$, $60.5\,s$ and $66\,s$ respectively (left): corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the shoreline (right), for $k = 3$ and $n^e_{\rm el} = 60$ and $n^i_{\rm el} = 10$.

6.5.8 Free-motion: well-balanced property

In this test-case, we aim to assess the motionless steady-states preservation property for the case of a floating body partially immersed in water over a varying bottom. The computational domain is $\Omega = [-50, 200]$. The object is initially placed such that $(x_G, z_G) = (50, H_0 + 2.5)$ and its mass m_o is defined according to Appendix E. The topography profile is defined as follows:

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{\pi\,(x - x_1)}{75}\right)\right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ \dfrac{1}{\beta}\,(x - x_3) & \text{if} \quad x \geq x_3, \\ 0 & \text{elsewhere,} \end{cases} \tag{6.116}$$

where $A = 1.5\,m$, $\beta = 11$, $x_1 = 12.5\,m$, $x_2 = 87.5\,m$ and $x_3 = 90\,m$. The initial data in $\mathcal{E}_0$ is defined as

$$\eta_0^e(x) := \max\left(5, b(x)\right) \qquad \text{and} \qquad q_0^e = 0,$$

while in the interior domain $\mathcal{I}_0$, we set

$$\eta_0^i := p^k_{\mathcal{T}_h^{i,0}}(\eta_{\text{lid}}) \qquad \text{and} \qquad q_0^i = 0.$$

We evolve this initial configuration up to $T_{\max} = 50\,s$, with $k = 3$, $n_{\text{el}}^e = 50$ and $n_{\text{el}}^i = 10$. The numerical results obtained with the DG-ALE scheme using the *a posteriori* LSC method are shown in Fig. 6.23. A zoom in the vicinity of the immersed structure and the shoreline, shows the corrected and uncorrected subcells in Fig. 6.24. The well-balanced steady state is preserved up to machine accuracy.

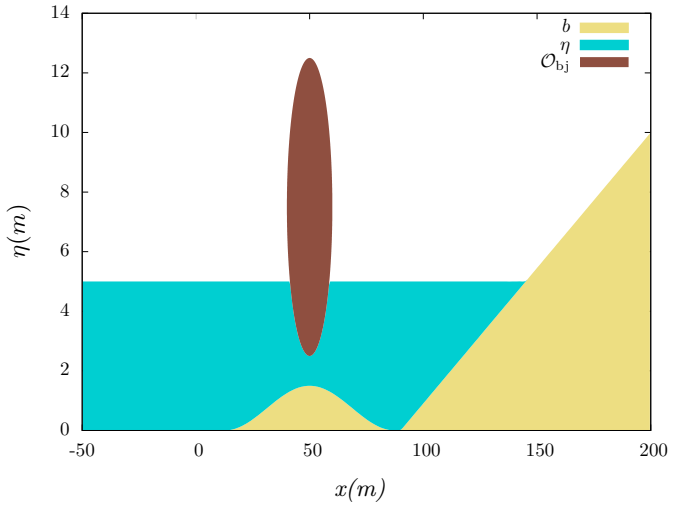

Figure 6.23: Test 34 - Free-motion: well-balanced property - Free surface elevation at $t = 50s$ for $k = 3$ and $n_{\text{el}}^e = 50$, $n_{\text{el}}^i = 10$.

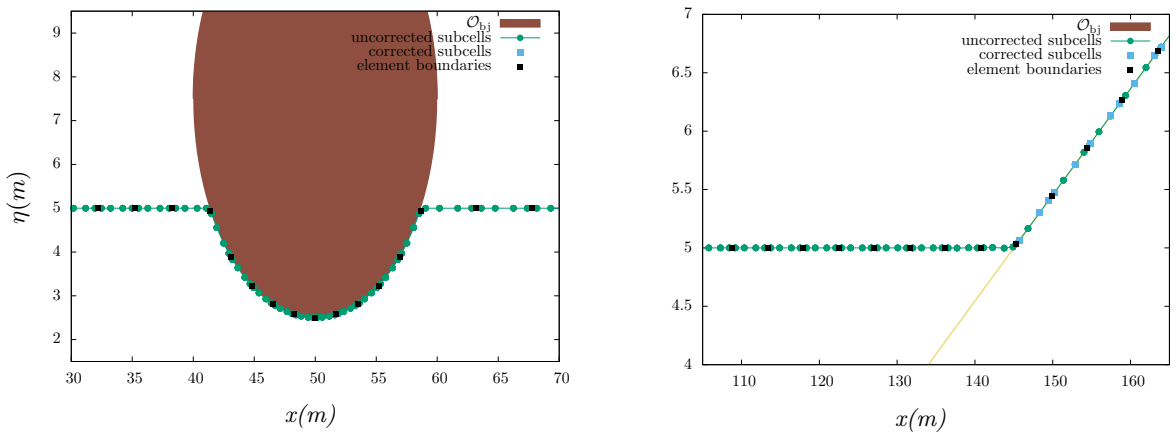

Figure 6.24: Test 34 - Free-motion: well-balanced property - Free surface elevation at $t = 50s$, with a zoom near the floating body (left) and near shore (right), showing the corrected and the uncorrected subcells, for $k = 3$ and $n^e_{\text{el}} = 50$, $n^i_{\text{el}} = 10$.

6.5.9 Free motion: convergence towards a motionless steady-state

Here, we analyze the case of a freely floating body returning to equilibrium. We consider the domain $\Omega =]-50, 150[$ and the topography is defined as

$$b(x) = \begin{cases} A \sin\left(\dfrac{\pi\,(x - x_1)}{75}\right) & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere.} \end{cases} \tag{6.117}$$

where $x_1 = 12.5$, $x_2 = 87.5$ and $A = 2.5$. The object is initially placed at $z_G(0) = H_0 + 2.5\ m$, with $H_0 = 8\,m$, and the initial conditions are defined as follows:

$$\eta^e_0 = H_0,\ \eta^i_0 = p^k_{\mathscr{T}^{i,0}_h}(\eta_{\text{lid}}) \qquad \text{and} \qquad q^e_0 = q^i_0 = 0.$$

However, we consider a modified (heavier) mass m_o such that the Newton's law is not initially balanced:

$$m_o = 2.19 \times \int_{\mathcal{I}(0)} \rho(H_0 - \eta^i_0(x))\, dx. \tag{6.118}$$

With such a choice, $z_G = H_0 + 0.83\ m$ would be the equilibrium elevation for $\mathcal{M}_G$ and when the object is released at $t = 0$ from an upper position, it returns to the targeted lower equilibrium location.
We set $n^e_{\text{el}} = 60$ and $n^i_{\text{el}} = 10$, and $k = 3$. The solution is shown in Fig. 6.25 for several time-steps in the range $[0\,s, 25\,s]$. We see how the release of the object creates some shock waves, due the sudden fall of the object. Thanks to the *a posteriori* LSC, the solution is free from spurious oscillations. In Fig. 6.26, we show the variation of $z_G(t)$ in the range of time $[0\,s, 25\,s]$, describing the return of the floating body to equilibrium state (left) and showing water-body equilibrium state at $t = 25$ (right). In Fig. 6.28, corrected and uncorrected subcells are respectively plotted with blue squares and green dots (left), and we zoom on the right wave (right). We also show, in Fig. 6.27, the variation of the horizontal coordinate of the contact points $\chi_-(t)$ and $\chi_+(t)$ in the same range of time.

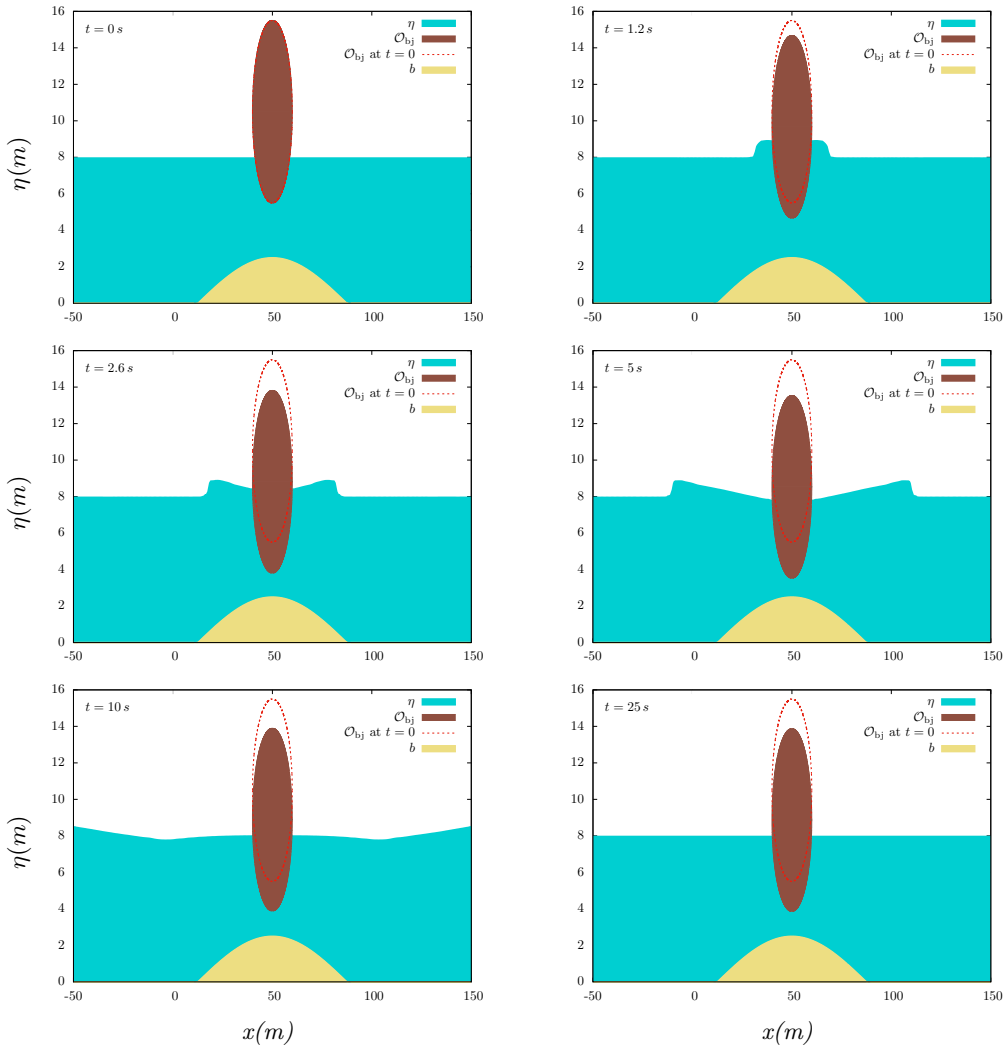

Figure 6.25: Test 35 - Free motion: convergence towards motionless steady-state - Free surface elevation computed for different values of time in the range $[0\,s, 25\,s]$, for $k = 3$, $n^e_{\rm el} = 60$ and $n^i_{\rm el} = 10$.

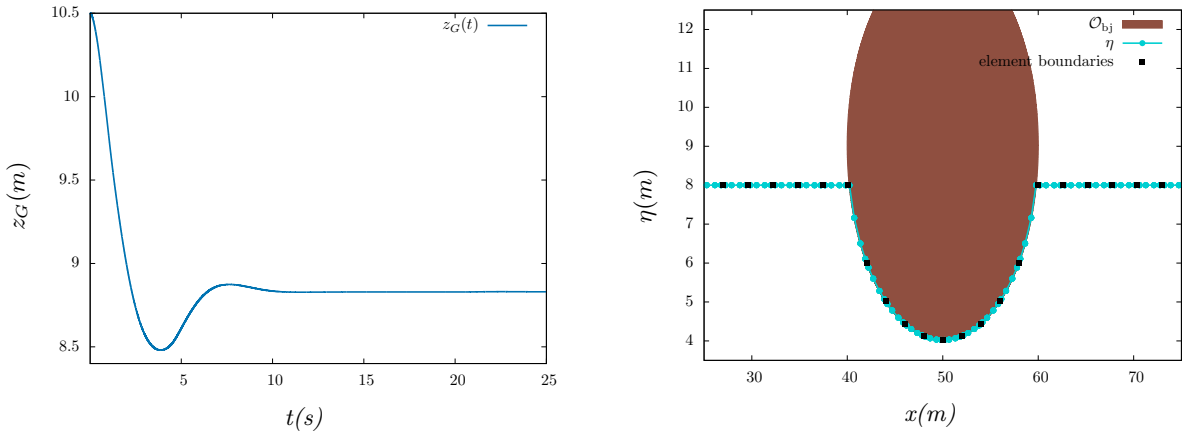

Figure 6.26: Test 35 - Free motion: convergence towards motionless steady-state - showing the variation of $z_G(t)$ in the range of time $[0\,s, 25\,s]$ (left), and showing water-body equilibrium state at $t = 25$ (right), for $k = 3$, $n^e_{\rm el} = 60$ and $n^i_{\rm el} = 10$.

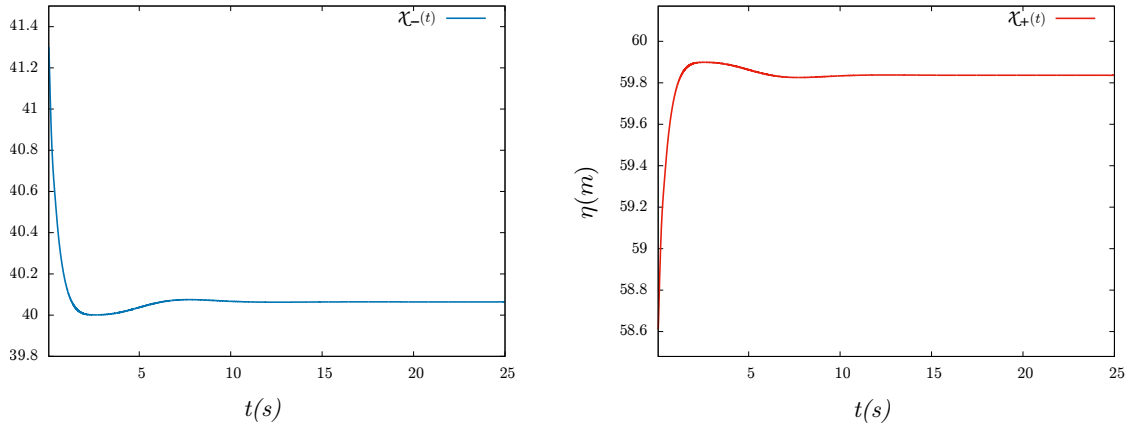

Figure 6.27: Test 35 - Free motion: convergence towards motionless steady-state - showing the variation of the horizontal coordinate of the contact points $\chi_-(t)$ and $\chi_+(t)$ in the range of time $[0\,s, 25\,s]$, for $k = 3$, $n^e_{\rm el} = 60$ and $n^i_{\rm el} = 10$.

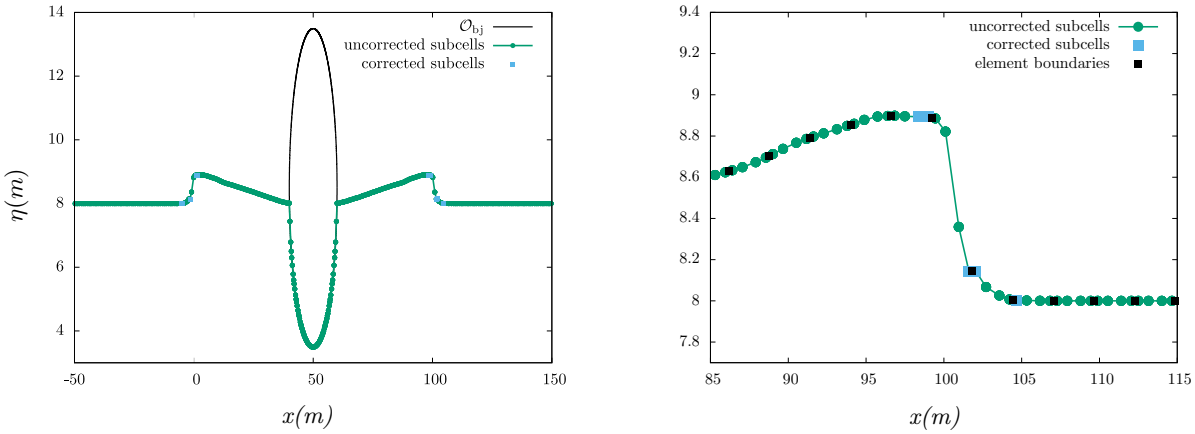

Figure 6.28: Test 35 - Free motion: convergence towards motionless steady-state - Free surface elevation computed at $t = 4.9\,s$: corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the right wave, for $k = 3$ and $n^e_{\text{el}} = 60$ and $n^i_{\text{el}} = 10$.

6.5.10 Free motion: interactions with a solitary wave

In this test-case, we aim to study the totally free response of a floating object to a propagating surface wave, with heaving, surging and pitching. The computational domain is set to $\Omega = [-150, 150]$. The structure is placed at $(x_G(0), z_G(0)) = (50, H_0 + 2.5)$, in a lake at rest with a flat bottom. The mass m_o of the body is defined as in (E.2) and is thus in equilibrium with the Archimedean force. A propagating wave is initially defined as follows

$$\eta^e_0(x) = H_0 + \frac{A_w}{\cosh(\gamma(x - x_0))^2} \quad \text{and} \quad q^e_0 = c_{q_2}\frac{\sqrt{\mathrm{g}}(\eta^e_0 - H_0)}{c_{q_1}}H^e_0, \tag{6.119}$$

and the initial water elevation and discharge beneath the floating body are defined by:

$$\eta^i_0 = p^k_{\mathscr{T}^{i,0}_h}(\eta_{\text{lid}}) \qquad \text{and} \qquad q^i_0 = 0,$$

with $A_w = 0.92\,m$, $x_0 = -80\,m$, $c_{q_1} = 0.1$, $c_{q_2} = 0.05$ and $\gamma = \frac{c_{q_1}}{\sqrt{\frac{4H_0}{3A_w}}}$. This wave propagates and meets the structure, which consequently starts to move, see Fig. 6.29. First, we sequentially isolate the three possible motions (only one degree of freedom is allowed: pure heaving, pure surging and pure pitching), see Fig. 6.30-6.31, Fig. 6.32-6.33 and Fig. 6.34-6.35 respectively. Next, we consider a totally free motion in which heaving, surging and pitching are naturally allowed, see Fig. 6.36- 6.37-6.38.

6.5.11 Free motion with a wet-dry transition

In this test-case, we study the propagation and run-up of a surface wave over a plane beach, with a partially immersed object placed on the way. The computational domain is set to $\Omega = [-300, 150]$ and the topography is made of a constant depth area and a plane sloping beach of constant slope $\frac{1}{\alpha}$ such that $\alpha = 11$, see Fig. 6.39. The mass of the body m_o is defined as in (E.2), and the initial condition for the flow in the *exterior* region is defined as follows:

$$\eta^e_0(x) = H_0 + \frac{A_w}{\cosh(\gamma(x - x_0))} \quad \text{and} \quad q^e_0 = c_{q_2}\frac{\sqrt{\mathrm{g}}(\eta^e_0 - H_0)}{c_{q_1}}H^e_0, \tag{6.120}$$

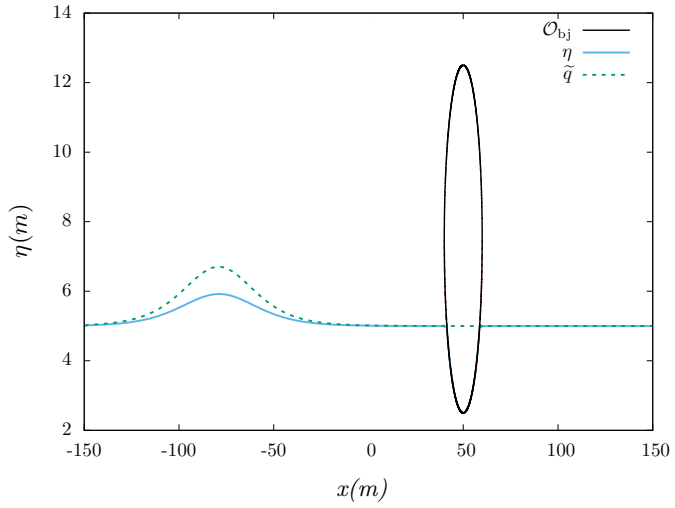

Figure 6.29: Test 36 - Free motion: interactions with a solitary wave - Free surface elevation and discharge computed at initial time for $k = 3$, $n^e_{\rm el} = 50$ and $n^i_{\rm el} = 10$.

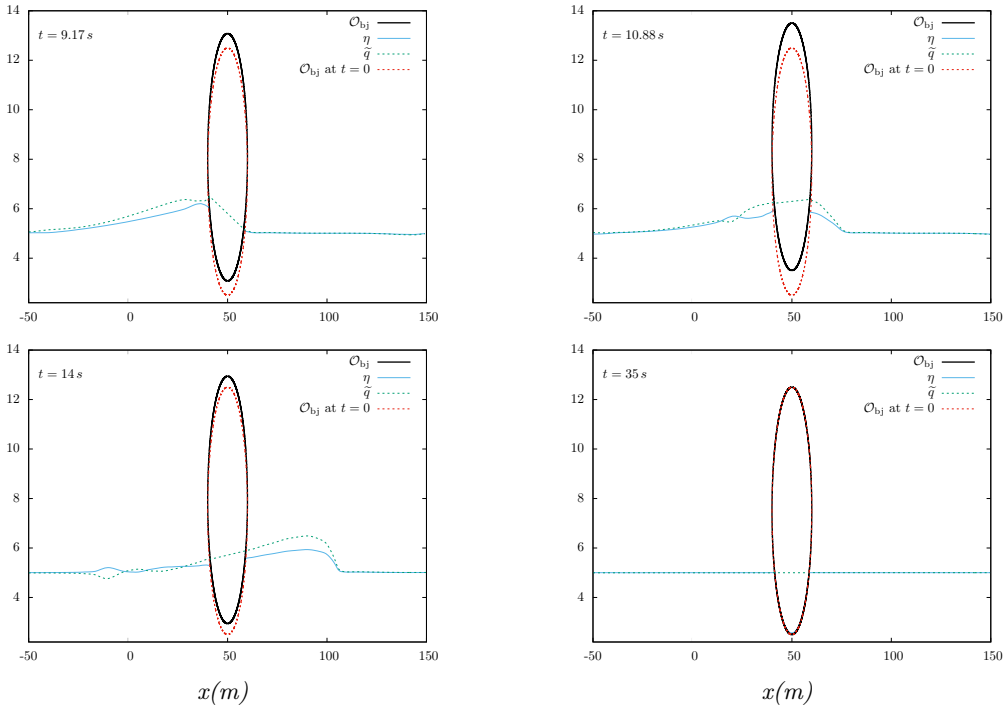

Figure 6.30: Test 36A - Pure heave motion - Free surface elevation and discharge computed for different values of time in the range $[0\,s, 35\,s]$ for $k = 3$, $n^e_{\rm el} = 50$ and $n^i_{\rm el} = 10$.

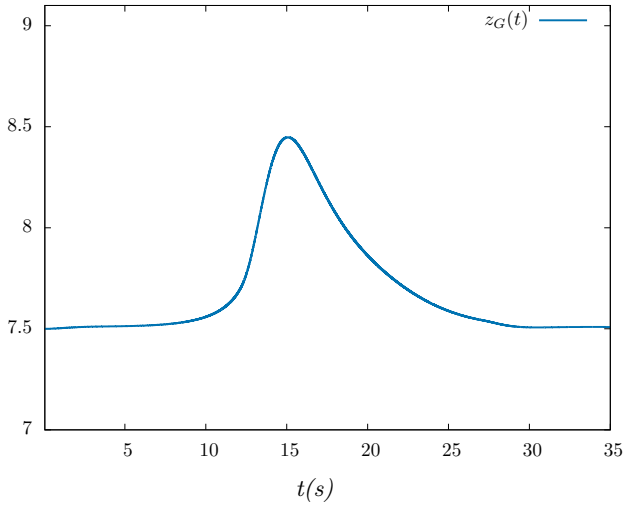

Figure 6.31: Test 36A - Pure heave motion - showing the variation of $z_G(t)$ in the range of time $[0\,s, 35\,s]$ for $k = 3$ and $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

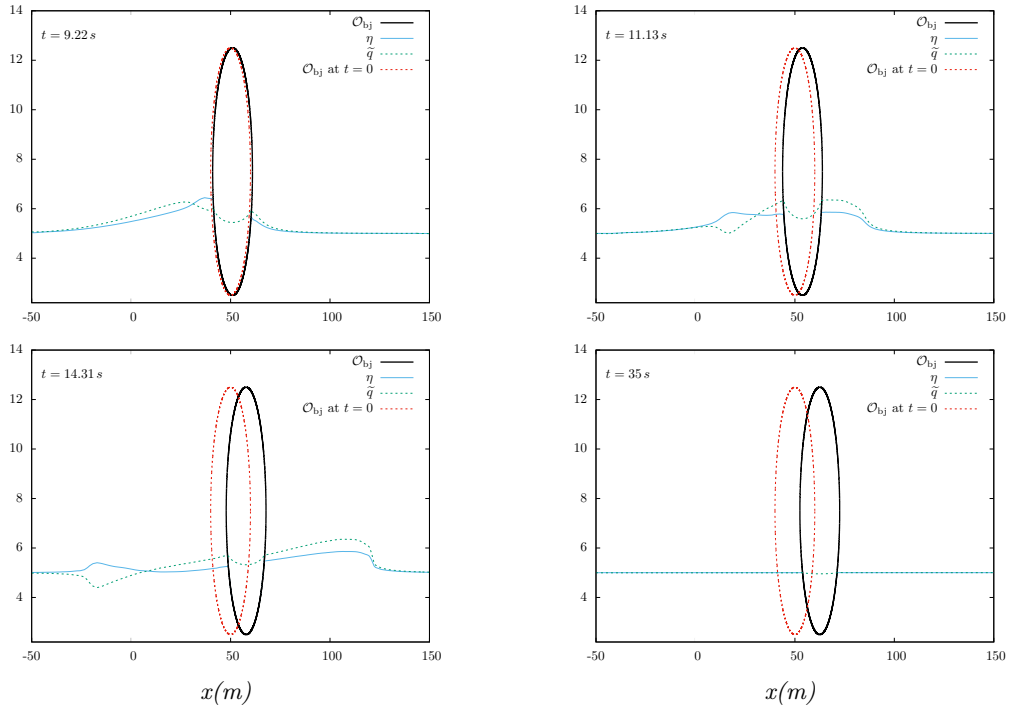

Figure 6.32: Test 36B - Pure surge motion - Free surface elevation and discharge computed for different values of time in the range $[0\,s, 35\,s]$ for $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

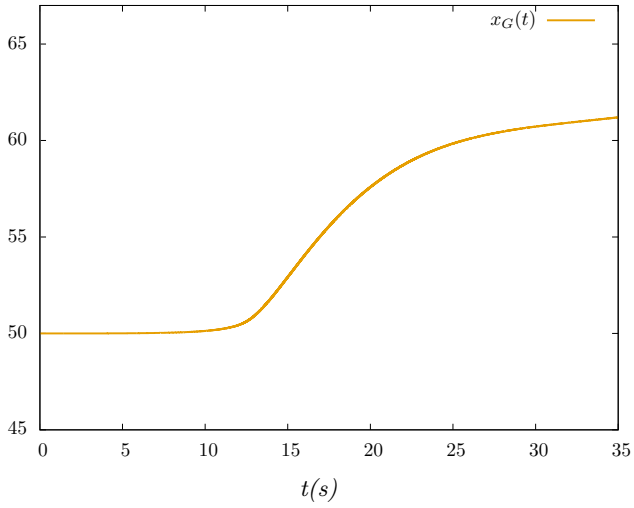

Figure 6.33: Test 36B - Pure surge motion - showing the variation of $x_G(t)$ in the range of time $[0\,s, 35\,s]$ for $k = 3$ and $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

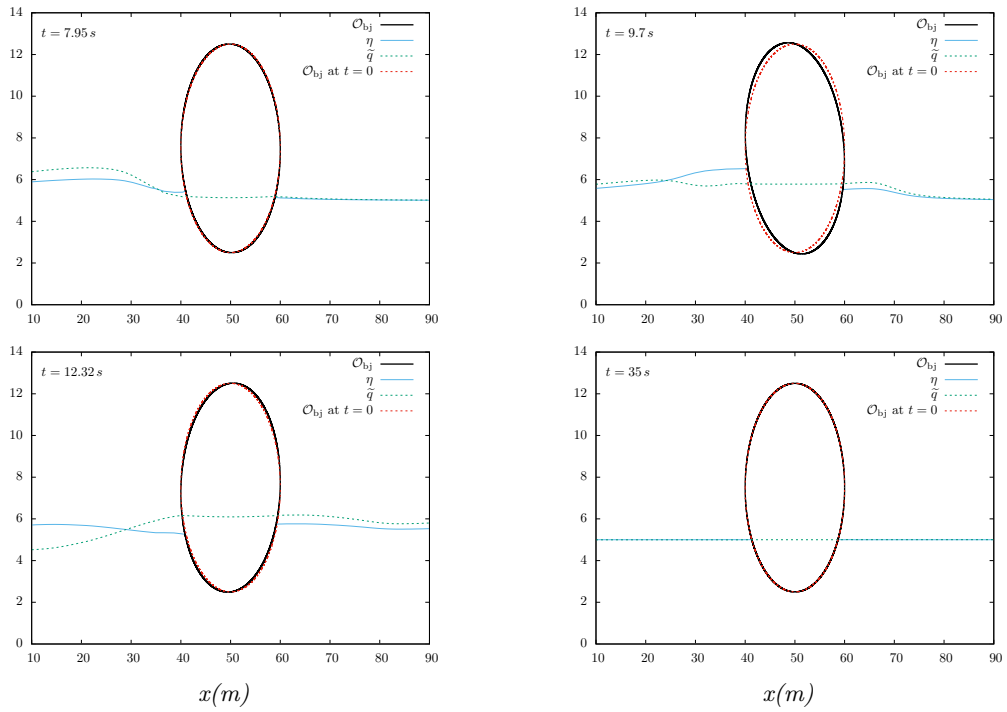

Figure 6.34: Test 36C - Pure pitch motion - Free surface elevation and discharge computed for different values of time in the range $[0\,s, 35\,s]$ for $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

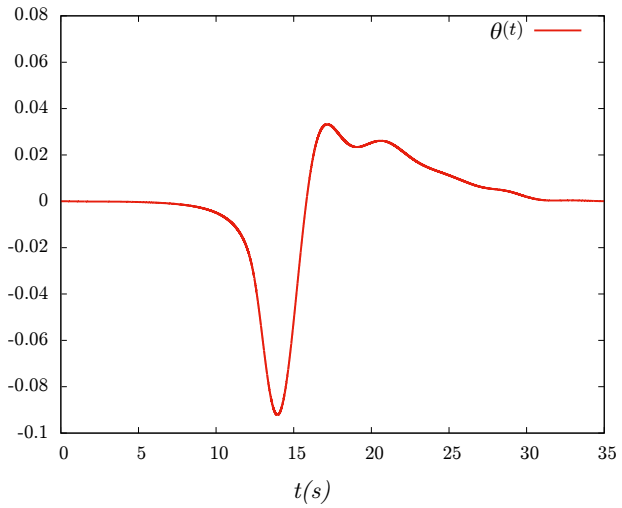

Figure 6.35: Test 36C - Pure pitch motion - showing the variation of $\theta(t)$ in the range of time $[0\,s, 35\,s]$ for $k = 3$ and $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

while the initial water elevation and discharge under the object are defined by:

$$\eta^i_0 = p^k_{\mathscr{T}^{i,0}_h}(\eta_{\text{lid}}) \qquad \text{and} \qquad q^i_0 = 0,$$

with $A_w = 0.55\,m$, $x_0 = -150\,m$, $c_{q_1} = 0.1$, $c_{q_2} = 0.05$ and $\gamma = \frac{c_{q_1}}{\sqrt{\frac{4H_0}{3A_w}}}$.

We set $n^e_{\text{el}} = 70$, $n^i_{\text{el}} = 10$, and $k = 3$. We show on Fig. 6.40 the free-surface obtained at several time values in the range $[16.78\,s,\ 150\,s]$. In Fig. 6.41, we show the trajectory of $\mathscr{M}_G$, under the form of time-series of its spatial coordinates (x_G, z_G) and the pitch angle θ. In Fig. 6.42, we show the free-surface at $t = 28.61\,s$, where corrected and uncorrected subcells are respectively plotted with blue squares and green dots (left), with a zoom on the shoreline (right). We observe that the body is pushed shoreward by the incoming wave, which is almost entirely transmitted to the shore, generating a run-up, before being reflected by the topography and hit the object again, pushing it seaward. The computations are performed in a very robust way, and we observe that as expected, the *a posteriori* LSC method only operate in the vicinity of the shoreline.

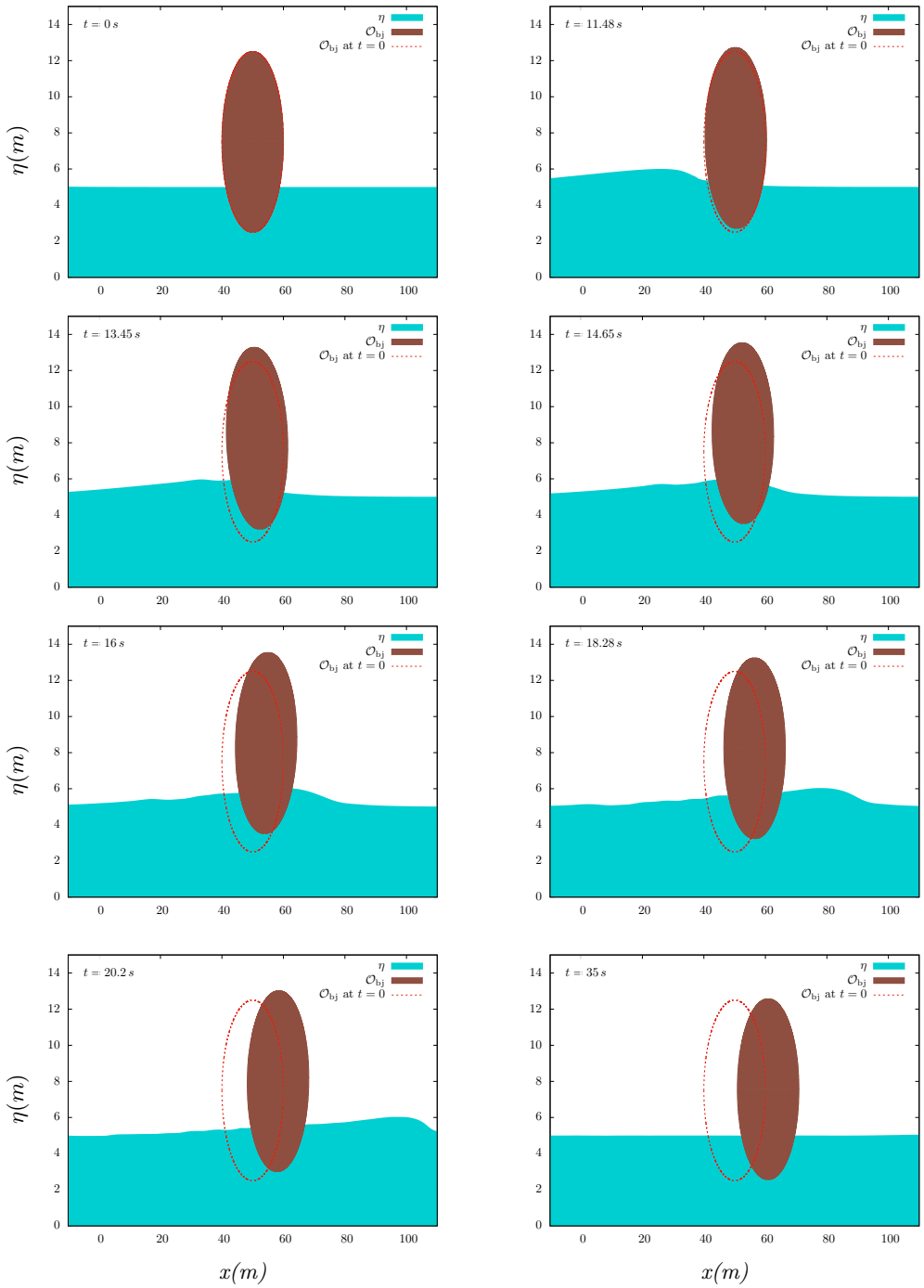

Figure 6.36: Test 36D - Heaving, surging and pitching are allowed - Free surface elevation computed for different values of time in the range $[0\,s, 35\,s]$ for $k = 3$, $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

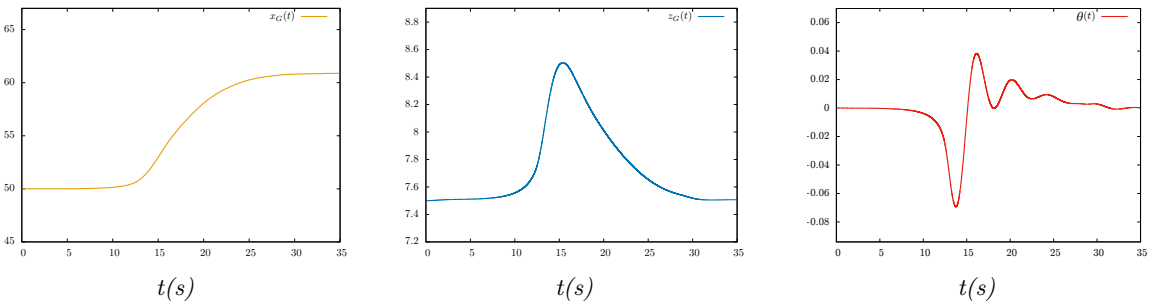

Figure 6.37: Test 36D - Heaving, surging and pitching are allowed - showing the variation of $x_G(t)$, $z_G(t)$ and $\theta(t)$ in the range of time $[0\,s, 35\,s]$ for $k = 3$ and $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

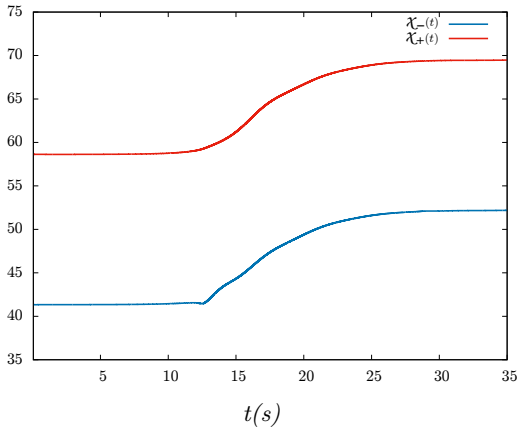

Figure 6.38: Test 36D - Heaving, surging and pitching are allowed - showing the variation of the horizontal coordinate of the contact points $\chi_-(t)$ and $\chi_+(t)$ in the range of time $[0\,s, 35\,s]$ for $k = 3$ and $n^e_{\text{el}} = 50$ and $n^i_{\text{el}} = 10$.

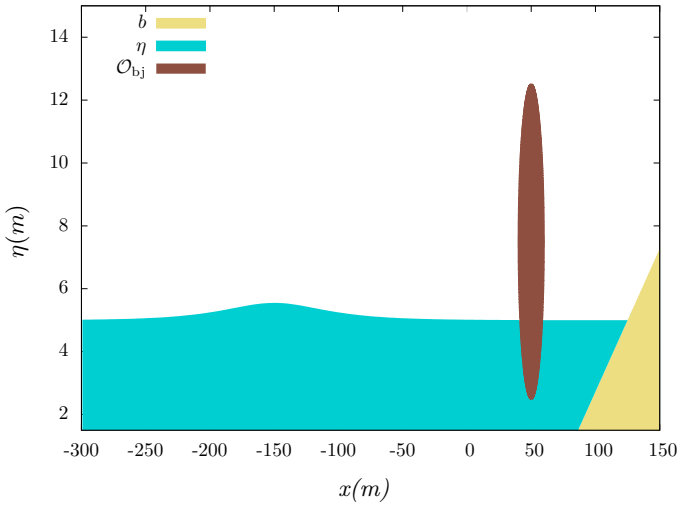

Figure 6.39: Test 37 - Free motion with a wet-dry transition - Free surface elevation at initial time for $k = 3$, $n^e_{\mathrm{el}} = 70$ and $n^i_{\mathrm{el}} = 10$.

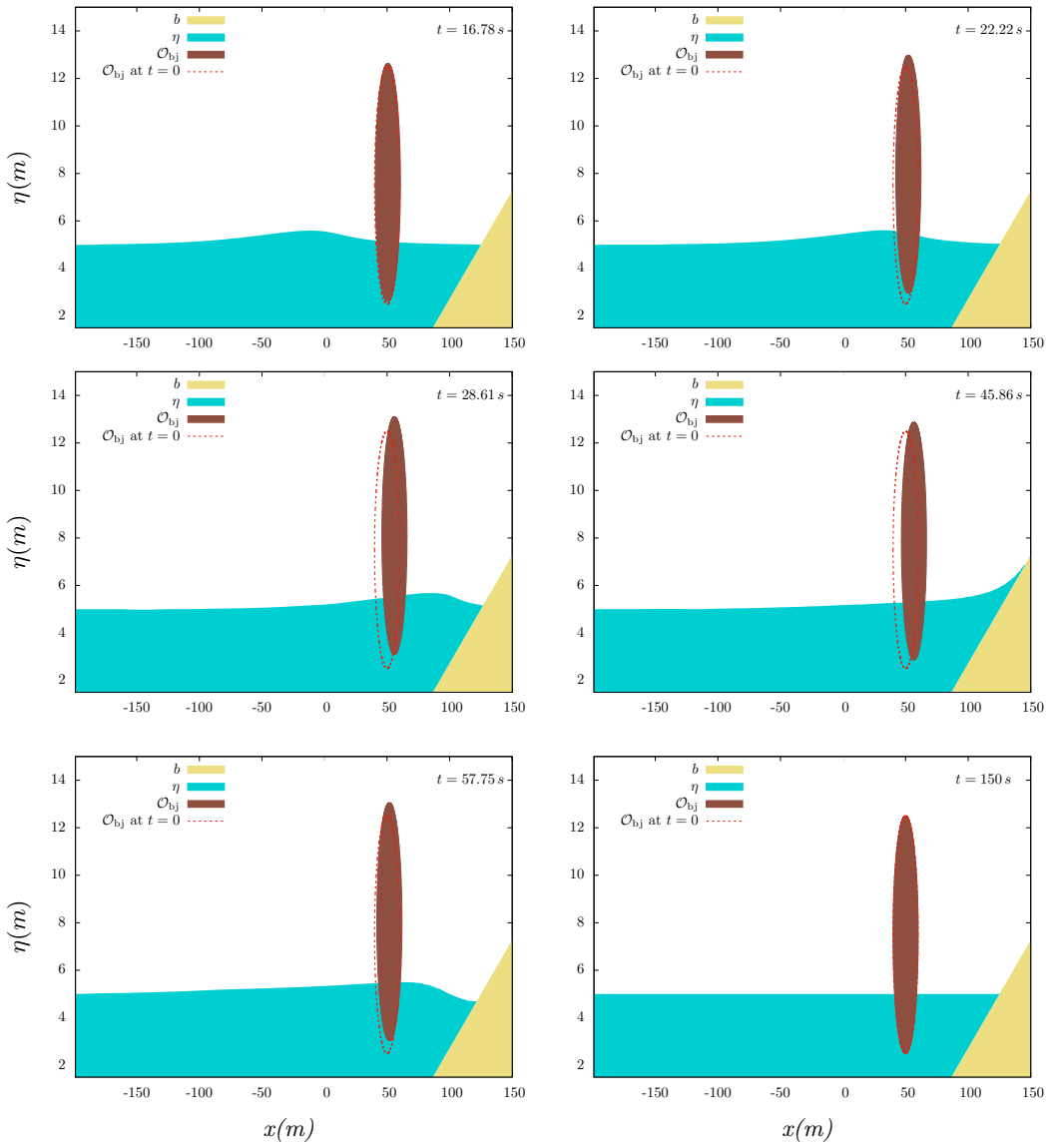

Figure 6.40: Test 37 - Free motion with a wet-dry transition - Free surface elevation for different values of time in the range $[16.78\,s,\ 150\,s]$, for $k = 3$, $n^e_{\text{el}} = 70$ and $n^i_{\text{el}} = 10$.

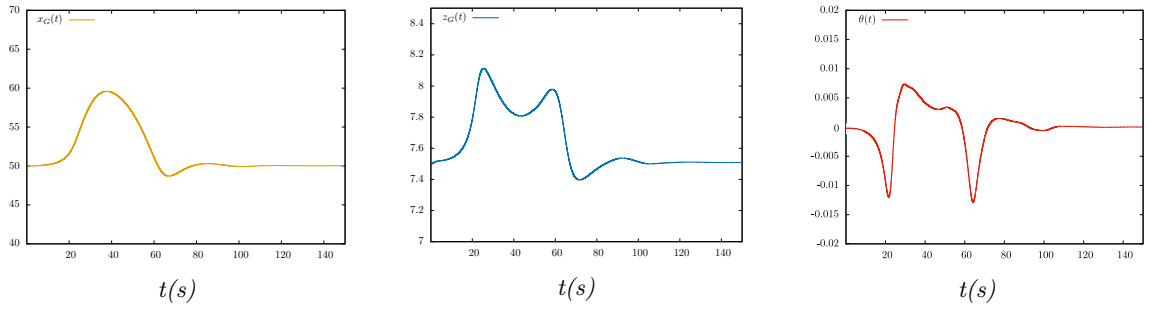

Figure 6.41: Test 37 - Free motion with a wet-dry transition - Time-series for $\boldsymbol{\mathcal{X}}_{\mathrm{G}} = (x_G, z_G, \theta)$ (from left to right) in the range of time $[0\,s, 150\,s]$, for $k = 3$, $n^e_{\mathrm{el}} = 70$ and $n^i_{\mathrm{el}} = 10$.

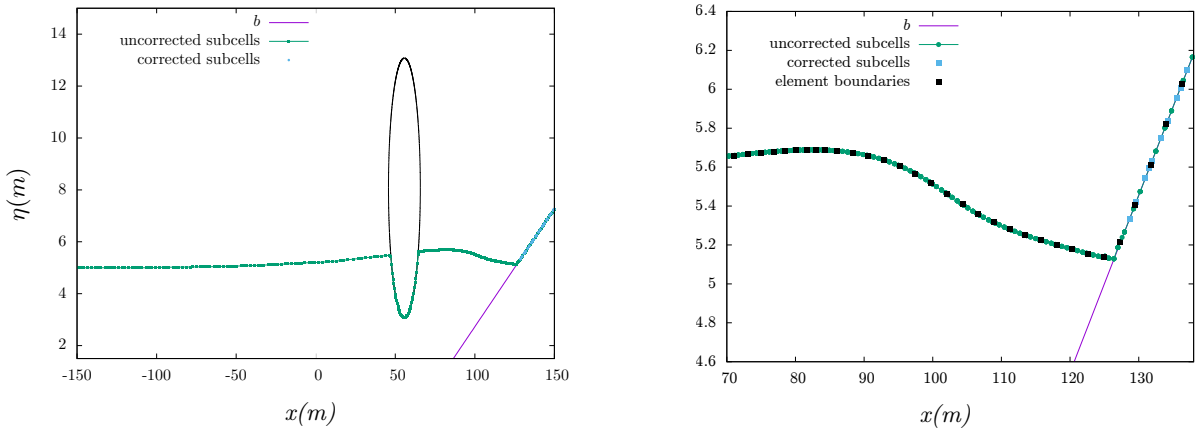

Figure 6.42: Test 37 - Free motion with a wet-dry transition - Free surface elevation computed at $t = 28.61\,s$ for $k = 3$, $n^e_{\mathrm{el}} = 70$ and $n^i_{\mathrm{el}} = 10$: corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the shoreline.

Chapter 7

Conclusions and perspectives

In this thesis, we begin with the introduction of a new well-balanced high-order DG discrete formulation with a FV-Subcell correction patch designed for the NSW equations. This method combines the very high accuracy of DG schemes along with a robust correction procedure ensuring the water-height positivity as well as addressing the issue of spurious oscillations in the vicinity of discontinuities. This robustness is enforced by means of an *a posteriori* LSC of the conservative variables. This procedure relies on an advantageous reformulation of DG schemes as a FV-like method on a sub-grid, which makes the correction strategy surgical and flexible, as well as conservative at the subcell level. Indeed, only the non-admissible subcells are marked and subject to correction, retaining as much as possible the very accurate subcell resolution of high-order DG formulations. The proposed strategy is investigated through an extensive set of benchmarks, including a brand new smooth solution for the computation of convergence rates, stabilization of flows with discontinuities, the preservation of motionless steady states, or moving shorelines over varying bottoms. We observe in particular that this approach provides a very accurate description of wet-dry interfaces even with the use of very high-order schemes on coarse meshes.
Regarding the potential advantages of this *a posteriori* limiting strategy compared to *a priori* limiters, because the troubled zone detection is performed *a posteriori*, the correction can be done only where it is absolutely necessary. Furthermore, the positivity preservation of the water-height is included without any additional effort, while it is generally not the case for *a priori* limitations of high-order schemes as highlighted in Chapter 2 for the $PL\ DG\text{-}FV_{subcell}$ limitation method. We further emphasize that our *a posteriori* LSC method scalability to any order of accuracy is also perfectly natural. Finally, it is important to note that this new correction procedure is totally parameter free. We also extend our *a posteriori* LSC method to the 2D NSW system and we briefly illustrate the potentiality of this approach with several numerical results.

In a second part, we introduce a novel numerical approximation algorithm allowing to compute fluid-structure interactions between a partially immersed floating object and shallow-water flows. This new discrete formulation is based on a DG-ALE global discretization for the flow model, coupled with a set of nonlinear ordinary differential equations for the resolution of the free-boundary problems associated with the time evolution of the air-fluid-structure interface, and the time evolution of the discharge beneath the object. The computation of the water free-surface beneath the floating object is reduced to a nonlinear algebraic equation to solve, essentially ruled by the object's position and underside's shape. In order to allow the computation of general waves interactions, possibly involving non-smooth surface waves, we extend the *a posteriori* LSC method of [80] to the current

DG-ALE description. In particular, we show that the resulting global flow discrete formulation preserves the DGCL, as well as the well-balancing property for motionless steady states, for any order of approximation in space. The resulting numerical strategy combines the high accuracy of DG approximations, with a robust stabilization process which ensures the positivity of the water-height at the subcell level, as well as preventing from the occurrence of spurious oscillations in the vicinity of discontinuities, discontinuities of the gradient and extrema. Indeed, we studied two types of object's motion: the object's motion may be either prescribed (where the case of a stationary body can be seen as a particular case of the prescribed body motion), or computed as a response to the hydrodynamic forcing associated with the flow motion. The floating body is allowed to move with heave, surge and pitch motions. These assets was numerically illustrated through an extensive set of manufactured benchmarks validating the water-body interaction model.

In future works, we plan to further investigate the 2D *a posteriori* LSC method and in particular the well-balancing and positivity-preserving) properties. Optimizing the code to reduce the computational time is also an important issue for practical applications. A more difficult goal is to generalize our numerical approximation algorithm for the wave-floating structure interaction problem in 2D.

Appendix A

Derivation of shallow-water asymptotics

Free surface flow: main notations and boundary conditions

For nearshore as well as in relatively deep waters, the coastal engineering community has used for a long time asymptotic depth averaged approximations. Following this approach, we aim in this work to use the NSW equations for our water-body interaction model. Historically, the physical model was first proposed by de Saint Venant in 1871 [50], obtained from asymptotic analysis and indeed depth-averaging the Navier-Stokes equations (see also [70]-[122]). Without including viscosity terms, a more simplified method to derive the non-linear shallow-Water system is by considering the NSW as an approximation of the incompressible Euler equations. These equations can be derived from the incompressible Euler's equations by combining a depth (height) averaging procedure with asymptotic expansions. These expansions are expressed in terms of two main dimensionless parameters. The first is the dispersion parameter μ

$$\mu = kh_r, \quad \text{with} \quad k\lambda = 2\pi, \tag{A.1}$$

where h_r is the reference water depth, λ the wavelength and k the wave number. Clearly, in shallow-waters, or for very long waves this parameter can be assumed to be small. The simplest models is obtained by considering a zeroth order approximation, which provides the well known hydrostatic Nonlinear shallow-water (NSW) model. This model gives a good representation of nonlinear waves as long as the dispersion parameter remains $\mu \leq \frac{\pi}{20}$. To account for the effects of shorter waves, higher order correction terms can be added. When doing it is customary to distinguish waves for which these effects are weakly or fully nonlinear. This is done introducing the nonlinearity parameter

$$\epsilon = \frac{A}{h_r}, \tag{A.2}$$

with A the wave elevation. For fully non-linear models $\epsilon = 1$, while weakly non-linear models are obtained under the hypothesis that $\epsilon \ll 1$. In this chapter, we follow this derivation procedure to arrive at the desired non-linear shallow-water approximation used in our numerical simulations. We repeat the formal derivation in one space dimension.

We start by introducing some notations that are going to be used in what follows. The model for the free surface problem is going to be presented and derived in two dimension (x, z) (one horizontal and the other vertical). We refer the reader to [106], where a Multidimensional expansion procedure is considered. As shown in Fig. A.1, we consider a cartesian coordinate system. We recall

that the characteristic scales for the flow are the wave amplitude A, wavelength λ and wave period T. The bathymetry is denoted by $b(x)$. The water-height (depth) $H(x,t)$ is the main unknown and this is defined as:

$$H(x,t) = H_0 + \zeta(x,t) - b(x), \tag{A.3}$$

where $\zeta(x,t)$ is the free surface elevation, relative to the reference still water depth H_0. The remaining unknowns are the vertical velocity $w(x,z,t)$, the horizontal velocity $u(x,z,t)$ and the pressure $P(x,z,t)$.

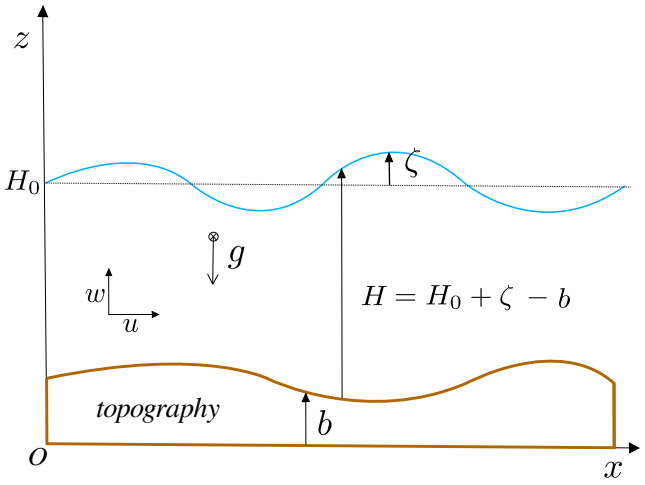

Figure A.1: Free surface flow: main notations

We note by ρ_ω the constant (and uniform in space) density of the fluid (water), neglecting the effects of the free surface tension and viscosity, the flow dynamics can be described by the incompressible Euler's equations (A.4)-(A.5). Starting by the incompressibility condition (volume conservation equation):

$$\partial_x u + \partial_z w = 0, \tag{A.4}$$

and the conservation of momentum equations (Newton's second law):

$$\partial_t u + u\partial_x u + w\partial_z u + \frac{1}{\rho_w}\partial_x P = 0, \tag{A.5a}$$

$$\partial_t w + u\partial_x w + w\partial_z w + \frac{1}{\rho_w}\partial_z P + g = 0. \tag{A.5b}$$

The boundary condition at the free surface elevation (kinetic condition) is:

$$w_f = \partial_t(H+b) + u_f\partial_x(H+b), \quad z = H+b, \tag{A.6}$$

or equivalently,

$$w_f = \partial_t\zeta + u_f\partial_x\zeta, \quad z = H+b, \tag{A.7}$$

where we have defined $u_f = u(x, H+b, t)$ and $w_f = w(x, H+b, t)$, the values of the velocities at the free surface. The boundary condition of the pressure at the free surface (dynamic condition) is:

$$P_f - P_{air} = 0, \quad z = H + b. \tag{A.8}$$

Since the atmospheric pressure is assumed to be constant, we can thus replace the pressure P by its relative value,

$$\Pi = P - P_{air}, \tag{A.9}$$

satisfying the boundary condition $\Pi_f = 0$. At seabed level ($z = b$), we have the impermeability condition, which is similar to the kinematic free surface condition, that is:

$$w_s = \partial_t b + u_s \partial_x b = u_s \partial_x b, \quad z = b, \tag{A.10}$$

where $u_s = u(x, b, t)$ and $w_s = w(x, b, t)$ are the values of the variables at the seabed, assuming that $\partial_t b = 0$. Finally, the models considered in this work are obtained under the hypothesis of irrotational flow which is (in 1d):

$$\partial_z u = \partial_x w. \tag{A.11}$$

Dimensional analysis

We introduce here the non-dimensional form of the problem. This is done as a first step toward the simplification of the Euler's equations. The nondimensional variables are evaluated dividing all the physical quantities by a set of selected reference scales for mass, time, length, flow speed and pressure. The parameters of dispersion μ (A.1) and nonlinearity ϵ (A.2) naturally appear in this process. We start by introducing the nonlinearity bathymetry parameter:

$$\beta = \frac{b_0}{H_0},$$

where b_0 is the characteristic variation of the bathymetry. And the non-dimensional bathymetry is defined by:

$$\widetilde{b}(\widetilde{x}) = \frac{b(x)}{\beta H_0}, \tag{A.12}$$

and the other non-dimensional variables are defined by:

$$\widetilde{t} = \mu \frac{\sqrt{\mathrm{g}H_0}}{H_0} t, \ \widetilde{x} = \frac{\mu}{H_0} x, \ \widetilde{z} = \frac{z}{H_0}, \ \widetilde{h_r}(\widetilde{x}) = \frac{h_r(x)}{H_0} = 1, \ \widetilde{\zeta}(\widetilde{x}, \widetilde{t}) = \frac{\zeta(x,t)}{\epsilon H_0}, \tag{A.13a}$$

$$\widetilde{H}(\widetilde{x}, \widetilde{t}) = \epsilon \widetilde{\zeta}(\widetilde{x}, \widetilde{t}) + \widetilde{h_r}(\widetilde{x}) - \beta b = \epsilon \widetilde{\zeta} + 1 - \beta b = \frac{H(x,t)}{H_0}, \ \widetilde{\Pi} = \frac{1}{\epsilon \mathrm{g} H_0} \frac{P - P_{air}}{\rho_\omega}, \tag{A.13b}$$

$$\widetilde{u} = \frac{1}{\epsilon \sqrt{\mathrm{g}H_0}} u, \ \widetilde{q} = \widetilde{H}\widetilde{u}, \ \widetilde{w} = \frac{1}{\epsilon \sqrt{\mathrm{g}H_0}} w, \ \widetilde{\mathrm{g}} = 1. \tag{A.13c}$$

Dropping the tilde notation, we write the incompressible Euler's equations in dimensionless form as:

$$\mu^2 \partial_x u + \partial_z w = 0, \tag{A.14a}$$

$$\partial_t u + \varepsilon \left(u \partial_x u + \frac{1}{\mu^2} w \partial_z u \right) + \partial_x \Pi = 0, \tag{A.14b}$$

$$\partial_t w + \varepsilon \left(u \partial_x w + \frac{1}{\mu^2} w \partial_z w \right) + \partial_z \Pi + 1 = 0. \tag{A.14c}$$

The nondimensional boundary conditions and irrotationality constraint can be written as

$$\mu^2 \left(\partial_t \zeta + \varepsilon u_f \partial_x \zeta \right) - w_f = 0, \qquad \text{at } z = 1 + \varepsilon\zeta, \tag{A.15a}$$

$$\Pi = 0, \qquad \text{at } z = 1 + \varepsilon\zeta, \tag{A.15b}$$

$$\beta \mu^2 u_s \partial_x b - w_s = 0, \qquad \text{at } z = \beta b, \tag{A.15c}$$

$$\partial_z u = \partial_x w. \tag{A.15d}$$

Depth averaging and asymptotic analysis

We derive the asymptotic approximations of the Euler equations in terms of depth averaged quantities:

$$\overline{u} = \frac{1}{1 + \varepsilon\zeta - \beta b} \int_{\beta b}^{1+\epsilon\zeta} u dz = \frac{1}{H} \int_{\beta b}^{1+\epsilon\zeta} u dz. \tag{A.16}$$

A variable playing a major role is also the volume flux (discharge) q:

$$q := \int_{\beta b}^{1+\varepsilon\zeta} u = (1 + \varepsilon\zeta - \beta b)\overline{u} = H\overline{u}.$$

One of the main tools used in the following analysis and reported here for completeness is the well known Leibnitz's integration rule:

$$\partial_x \left(\int_{a(x)}^{b(x)} f(x,z) dz \right) = \int_{a(x)}^{b(x)} \partial_x f(x,z) dz + \partial_x b(x) f(x, b(x)) - \partial_x a(x) f(x, a(x)), \tag{A.17}$$

where f, a and b are differentiable functions.

Mass equation

Integrating over the water depth equation (A.14a), we obtain:

$$\int_{\beta b}^{1+\varepsilon\zeta} \left(\mu^2 \partial_x u + \partial_z w \right) dz = 0. \tag{A.18}$$

Applying Leibnitz rule, equation (A.18) can be written as

$$w_f - w_s + \partial_x \left(\int_{\beta b}^{1+\varepsilon\zeta} \mu^2 u dz \right) + \beta \mu^2 \partial_x b u_s - \mu^2 \varepsilon \partial_x \zeta u_f = 0.$$

Substituting the kinetic (A.15a) and dynamic (A.15c) conditions to w_f and w_s and using the definition of the depth averaged velocity equation (A.16):

$$\partial_t \zeta + \partial_x (H\overline{u}) = 0,$$

or equivalently,

$$\partial_t H + \partial_x q = 0. \tag{A.19}$$

This is commonly called non-dimensional mass equation (or continuity equation) and it represents the conservation of volume (or equivalently mass) of water in the domain. The mass equation (A.19) is a direct consequence of combination of volume conservation with the boundary conditions of the Euler's equations.

Momentum equation

To obtain an evolution equation for the depth averaged horizontal velocity (multiplied by H) $H\overline{u}$, we integrate equation (A.14b) over the depth:

$$\int_{\beta b}^{1+\varepsilon\zeta} \left(\partial_t u + \varepsilon \left(u\partial_x u + \frac{1}{\mu^2} w \partial_z u \right) + \partial_x \Pi \right) dz = 0. \tag{A.20}$$

Evaluating each terms of (A.20) separately. Using Leibnit's integration rule (A.17), we get:

$$\int_{\beta b}^{1+\varepsilon\zeta} \partial_t u dz = \partial_t \left(\int_{\beta b}^{1+\varepsilon\zeta} u dz \right) - \varepsilon \partial_t \zeta u_f + \beta \partial_t b u_s = \left(\int_{\beta b}^{1+\varepsilon\zeta} u dz \right)_t - \varepsilon \partial_t \zeta u_f, \tag{A.21a}$$

$$\int_{\beta b}^{1+\varepsilon\zeta} \partial_x \left(\frac{u^2}{2} \right) dz = \partial_x \left(\int_{\beta b}^{1+\varepsilon\zeta} \frac{u^2}{2} dz \right) - \varepsilon \partial_x \zeta \frac{u_f^2}{2} + \beta \partial_x b \frac{u_s^2}{2}, \tag{A.21b}$$

$$\int_{\beta b}^{1+\varepsilon\zeta} \partial_x \Pi dz = \partial_x \left(\int_{\beta b}^{1+\varepsilon\zeta} \Pi dz \right) - \varepsilon \partial_x \zeta \Pi_f + \beta \partial_x b \Pi_s \overbrace{=}^{\Pi_f = 0} \partial_x \left(\int_{\beta b}^{1+\varepsilon\zeta} \Pi dz \right) + \beta \partial_x b \Pi_s \tag{A.21c}$$

using (A.14a) and then (A.15a)-(A.15c), one can write,

$$\int_{\beta b}^{1+\varepsilon\zeta} \partial_z u w dz = [\![uw]\!]_{\beta b}^{1+\varepsilon\zeta} - \int_{\beta b}^{1+\varepsilon\zeta} u \partial_z w dz = u_f w_f - u_s w_s + \mu^2 \int_{\beta b}^{1+\varepsilon\zeta} \partial_x \left(\frac{u^2}{2} \right) dz \tag{A.22a}$$

$$= \mu^2 \left(\partial_t \zeta u_f + \varepsilon \partial_x \zeta u_f \right) + \beta \mu^2 \partial_x b u_s + \mu^2 \int_{\beta b}^{1+\varepsilon\zeta} \partial_x \left(\frac{u^2}{2} \right) dz, \tag{A.22b}$$

where $\Pi_s = P(z = b)$ is the pressure at the seabed. Collecting all the terms and simplifying using again formula (A.17), we obtain the momentum equation:

$$\partial_t (H\overline{u}) + \varepsilon \partial_x \left(H\overline{u^2} \right) + \partial_x (H\overline{\Pi}) + \beta \partial_x b \Pi_s = 0, \tag{A.23}$$

having introduced the depth averaged pressure $\overline{\Pi}$:

$$\overline{\Pi} = \frac{1}{1 + \varepsilon\zeta - \beta b} \int_{\beta b}^{1+\varepsilon\zeta} \Pi dz.$$

Asymptotic velocites

Starting from equation (A.14a), we are going to express the expression for the vertical and horizontal velocity components in terms of the depth averaged horizontal velocity $\overline{u}$. We start by integrating eq. (A.14a) from the bottom to an arbitary depth z. Using (A.17) and then (A.15c), one can write:

$$w(z) = w_s - \mu^2 \left(\partial_x \left(\int_{\beta b}^{z} u dz \right) + \beta b_x u_s \right) = -\mu^2 \partial_x \left(\int_{\beta b}^{z} u dz \right). \tag{A.24}$$

Similarly, integrating the irrotationality condition (A.15d) and using equation (A.24) leads to

$$u(z) = u_s + \int_{\beta b}^{z} \partial_x w dz = u_s - \mu^2 \int_{\beta b}^{z} \partial xx \left(\int_{\beta b}^{z} u dz \right) dz. \tag{A.25}$$

We can now express the vertical velocity in terms of the bottom horizontal velocity by substituting equation (A.25) into equation (A.24), leading to the $\mathcal{O}\left(\mu^4\right)$ estimate:

$$w(z) = -\mu^2 \partial_x \left(\int_{\beta b}^{z} u_s dz \right) + \mathcal{O}\left(\mu^4\right) = -\mu^2 \partial_x \left(u_s (z - \beta b) \right) + \mathcal{O}\left(\mu^4\right). \tag{A.26}$$

Integrating again the irrotationality condition (A.15d), we can now, thanks to (A.26), improve the estimation of $u(z)$ as follows:

$$\begin{aligned} u(z) &= u_s - \mu^2 \int_{\beta b}^{z} \partial_{xx} \left(u_s (z - \beta b) \right) dz + \mathcal{O}\left(\mu^4\right), \\ &= u_s - \mu^2 \left(\partial_{xx} \left(u_s \right) \int_{\beta b}^{z} (z - \beta b) dz - u_s \int_{\beta b}^{z} \beta \partial_{xx} b dz \right) + \mathcal{O}\left(\mu^4\right), \\ &= u_s - \mu^2 \left(\partial_{xx} u_s \frac{(z - \beta b)^2}{2} - \beta u_s \partial_{xx} b (z - \beta b) \right) + \mathcal{O}\left(\mu^4\right). \end{aligned} \tag{A.27}$$

The depth averaged velocity can be now expressed as a function of the seabed velocity u_s using equation (A.27):

$$\overline{u} = \frac{1}{1 + \varepsilon\zeta - \beta b} \int_{\beta b}^{1+\varepsilon\zeta} u dz = u_s - \mu^2 \left(\partial_{xx} u_s \frac{H^2}{6} - \beta u_s \partial_{xx} b \frac{H}{2} \right) + \mathcal{O}\left(\mu^4\right). \tag{A.28}$$

We have thus a simple relation between the bottom velocity u_s and the depth averaged one $\overline{u}$:

$$\overline{u} = u_s + \mathcal{O}\left(\mu^2\right). \tag{A.29}$$

Using (A.29) we can invert (A.28) by writing:

$$u_s = \overline{u} + \mu^2 \left(\partial_{xx} \overline{u} \frac{H^2}{6} - \beta \overline{u} \partial_{xx} b \frac{H}{2} \right) + \mathcal{O}\left(\mu^4\right). \tag{A.30}$$

Using (A.30), we can, finally, express the horizontal and vertical velocity in terms of the depth averaged velocity:

$$u(z) = \overline{u} - \mu^2 \left(\partial_{xx} \overline{u} \left(\frac{(z - \beta b)^2}{2} - \frac{H^2}{6} \right) - \beta \overline{u} \partial_{xx} b \left(z - \beta b - \frac{H}{2} \right) \right) + \mathcal{O}\left(\mu^4\right), \tag{A.31a}$$

$$w(z) = -\mu^2 \partial_x \left(\overline{u} (z - \beta b) \right) + \mathcal{O}\left(\mu^4\right). \tag{A.31b}$$

Evaluation of nonlinear velocity term

As we are proceeding, we continue to express every term in terms of depth averaged velocity. The nonlinear term $\overline{u^2}$ at precision $\mathcal{O}\left(\mu^4\right)$ becomes:

$$\begin{aligned}\overline{u^2} &= \frac{1}{H}\int_{\beta b}^{1+\varepsilon\zeta} u^2 dz \\ &= \frac{1}{H}\int_{\beta b}^{1+\varepsilon\zeta}\left[\overline{u}-\mu^2\left(\overline{u}\left(\frac{(z-\beta b)^2}{2}-\frac{H^2}{6}\right)-\beta\overline{u}\partial_{xx}b\left(z-\beta b-\frac{H}{2}\right)\right)\right]^2 dz+\mathcal{O}\left(\mu^4\right) \\ &= \overline{u}^2+\mathcal{O}\left(\mu^4\right).\end{aligned} \tag{A.32}$$

Asymptotic pressure profile

The pressure is evaluated integrating from depth z to the free surface $1+\varepsilon\zeta$ equation (A.14c)

$$\Pi(z) = (1+\varepsilon\zeta-z)+\int_z^{1+\varepsilon\zeta}\partial_t w dz+\varepsilon\int_z^{1+\varepsilon\zeta} u\partial_x w dz+\frac{\varepsilon}{\mu^2}\int_z^{1+\varepsilon\zeta} w\partial_z w dz. \tag{A.33}$$

Evaluating each terms of (A.33) separately. Using equations (A.31a) and (A.31b) we get:

$$\begin{aligned}\int_z^{1+\varepsilon\zeta}\partial_t w dz &= -\mu^2\int_z^{1+\varepsilon\zeta}\partial_t\left(\partial_x(\overline{u}(z-\beta b))-\mathcal{O}\left(\mu^2\right)\right)dz \\ &= -\mu^2\int_z^{1+\varepsilon\zeta}\partial_t\left(\partial_x\overline{u}(z-\beta b)-\beta\overline{u}\partial_x b\right)dz+\mathcal{O}\left(\varepsilon\mu^4\right) \\ &= -\mu^2\left[\partial_{xt}\overline{u}\left(\frac{H^2}{2}-\frac{(z-\beta b)^2}{2}\right)-\beta\partial_t\overline{u}\partial_x b(1+\varepsilon\zeta-z)\right]+\mathcal{O}\left(\varepsilon\mu^4\right),\end{aligned}$$

$$\begin{aligned}\int_z^{1+\varepsilon\zeta} u\partial_x w dz &= -\mu^2\int_z^{1+\varepsilon\zeta}\left(\overline{u}+\mathcal{O}\left(\mu^2\right)\right)\partial_x\left(\partial_x\left(\overline{u}(z-\beta b)\right)-\mathcal{O}\left(\mu^2\right)\right)dz \\ &= -\mu^2\overline{u}\int_z^{1+\varepsilon\zeta}\partial_{xx}(\overline{u}(z-\beta b))dz+\mathcal{O}\left(\varepsilon\mu^4\right) \\ &= -\mu^2\overline{u}\partial_{xx}\overline{u}\left(\frac{H^2}{2}-\frac{(z-\beta b)^2}{2}\right)+\mu^2\overline{u}\left(2\beta\partial_x\overline{u}\partial_x b+\overline{u}\beta\partial_{xx}b\right)(1+\varepsilon\zeta-z)+\mathcal{O}\left(\varepsilon\mu^4\right),\end{aligned}$$

$$\begin{aligned}\int_z^{1+\varepsilon\zeta}\partial_z\left(\frac{w^2}{2}\right)dz &= \frac{\mu^4}{2}\int_z^{1+\varepsilon\zeta}\partial_z\left[\left(\partial_x(\overline{u}(z-\beta b))+\mathcal{O}\left(\mu^2\right)\right)^2\right]dz \\ &= \frac{\mu^4}{2}\int_z^{1+\varepsilon\zeta}\partial_z\left[\left(\partial_x\overline{u}(z-\beta b)-\beta\overline{u}\partial_x b\right)^2\right]dz+\mathcal{O}\left(\varepsilon\mu^4\right) \\ &= \frac{\mu^4}{2}\int_z^{1+\varepsilon\zeta}\partial_z\left[\partial_x\overline{u}\partial_x(\overline{u}(z-\beta b)^2)\right]dz+\mathcal{O}\left(\varepsilon\mu^4\right) = \mu^4\int_z^{1+\varepsilon\zeta}\partial_x\overline{u}\partial_x(\overline{u}(z-\beta b))dz+\mathcal{O}\left(\varepsilon\mu^4\right) \\ &= \mu^4\left[(\partial_x\overline{u})^2\left(\frac{H^2}{2}-\frac{(z-\beta b)^2}{2}\right)-\beta\overline{u}\partial_x\overline{u}\partial_x b(1+\varepsilon\zeta-z)\right]+\mathcal{O}\left(\varepsilon\mu^4\right).\end{aligned}$$

The pressure Π at depth z can thus be expressed at pressision $\mathcal{O}\left(\varepsilon\mu^4\right)$) as:

$$\begin{aligned}\Pi(z) = (1+\varepsilon\zeta - z) &- \mu^2\left[\partial_{xt}\overline{u}\left(\frac{H^2}{2} - \frac{(z-\beta b)^2}{2}\right) - \beta\partial_t\overline{u}\partial_x b(1+\varepsilon\zeta - z)\right] + \mathcal{O}\left(\varepsilon\mu^4\right)\\ &- \varepsilon\mu^2\overline{u}\left[\partial_{xx}\overline{u}\left(\frac{H^2}{2} - \frac{(z-\beta b)^2}{2}\right) - \left(2\beta\partial_x\overline{u}\partial_x b + \beta\overline{u}\partial_{xx}b\right)(1+\varepsilon\zeta - z)\right]\\ &+ \varepsilon\mu^2\left[(\partial_x\overline{u})^2\left(\frac{H^2}{2} - \frac{(z-\beta b)^2}{2}\right) - \beta\overline{u}\partial_x\overline{u}\partial_x b(1+\varepsilon\zeta - z)\right] + \mathcal{O}\left(\varepsilon\mu^4\right).\end{aligned} \tag{A.34}$$

We integrate the above pression equation to derive an explicit expression for the depth averaged pressure:

$$\begin{aligned}H\overline{\Pi} = \frac{H^2}{2} &- \mu^2\left(\partial_{xt}\overline{u}\frac{H^3}{3} - \beta\partial_t\overline{u}\partial_x b\frac{H^2}{2}\right)\\ &- \varepsilon\mu^2\overline{u}\left(\partial_{xx}\overline{u}\frac{H^3}{3} - \beta\left(2\partial_x\overline{u}\partial_x b + \overline{u}\partial_{xx}b\right)\frac{H^2}{2}\right)\\ &+ \varepsilon\mu^2\left((\partial_x\overline{u})^2\frac{H^3}{3} - \beta\overline{u}\partial_x\overline{u}\partial_x b\frac{H^2}{2}\right) + \mathcal{O}\left(\varepsilon\mu^4\right).\end{aligned} \tag{A.35}$$

Using (A.34) we can obtain the expression of the pressure at the seabed Π_s:

$$\Pi_s = \Pi(z=\beta b) = H - \mu^2\left(\partial_{xt}\overline{u}\frac{H^2}{2} - \beta\partial_t\overline{u}\partial_x bH\right) + \mathcal{O}\left(\mu^4, \varepsilon\mu^2\right). \tag{A.36}$$

NSW equations

Now we have all the tools to derive the nonlinear shallow-water model (NSW) approximation from the Euler's equations. The approximation hypothesis on which this model is based on is:

$$\varepsilon \approx 1 \quad \text{and} \quad \mu \ll 1,$$

such that all the terms of order $\mathcal{O}\left(\mu^2\right)$ or higher can be neglected. Substituting the equations (A.32), (A.35) and (A.36) into the momentum equation (A.23), the NSW nondimensional system of equations reads:

$$\begin{aligned}&\partial_t H + \partial_x(H\overline{u}) = 0\\ &(H\overline{u})_t + \varepsilon\left(H\overline{u}^2\right)_x + HH_x = -\beta H\partial_x b.\end{aligned} \tag{A.37}$$

Going back to dimensional variables and using the depth averaged flux (discharge) $q = H\overline{u}$, the dimensional NSW is:

$$\begin{aligned}&\partial_t H + \partial_x q = 0\\ &\partial_t q + \partial_x\left(\overline{u}q + \frac{1}{2}gH^2\right) = -gH\partial_x b.\end{aligned} \tag{A.38}$$

In this work, for sake of simplicity, the depth averaged horizontal velocity $\overline{u}$ notation are replaced by u.

Appendix B

New analytical smooth solutions for the NSW equations

This appendix aims at giving further details on the construction of a new smooth solution, of any arbitrary regularity, of the NSW equations. Following the methodology introduced in [163], we consider a smooth solution $\mathbf{v}$ in the context of flat bottom ($b = 0$), so that the NSW equations rewrite as:

$$\partial_t \mathbf{v} + \mathbf{A}(\mathbf{v})\partial_x \mathbf{v} = 0,$$

where the Jacobian matrix writes as:

$$\mathbf{A}(\mathbf{v}) = \nabla_{\mathbf{v}}\mathbf{F}(\mathbf{v}) = \begin{pmatrix} 0 & 1 \\ \mathrm{g}H - u^2 & 2u \end{pmatrix}.$$

The eigen-analysis of the matrix $\mathbf{A}(\mathbf{v})$ leads to the following pair of eigenvalues $\lambda^{\pm} = u \pm \sqrt{\mathrm{g}H}$ and eigenvectors:

$$E^{\pm} = \begin{pmatrix} 1 \\ u \pm \sqrt{\mathrm{g}H} \end{pmatrix}.$$

By diagonalizing the NSW system of equations, one finally gets the following Riemann invariants $\alpha^{\pm} = u \pm 2\sqrt{gH}$, governed by the following conservation laws:

$$\partial_t \alpha^{\pm} + \lambda^{\pm}\partial_x \alpha^{\pm} = 0. \tag{B.1}$$

In light of the definition of the Riemann invariants, the system eigenvalues can be reformulated in terms of $\alpha^{\pm}$ as follows:

$$\lambda^{\pm} = \frac{\alpha^{+}(2 \pm 1) + \alpha^{-}(2 \mp 1)}{4}.$$

To uncouple the two conservation laws (B.1), we consider a particular flow regime corresponding to the trans-critical particular situation where $\alpha^{-} = 0$, *i.e.* $u = 2\sqrt{gH}$. The NSW equations then finally reduce to the following very simple Burgers equation:

$$\partial_t u + \frac{3}{2} u\, \partial_x u = 0.$$

By definition, the characteristic curves $x(X, t)$ satisfy:

$$\begin{cases} \frac{\mathrm{d}(x(X,t)}{\mathrm{d}t} = \frac{3}{2}u(x(X,t),t), \\ x(X,0) = X, \end{cases}$$

u is constant all along the characteristic curve, i.e.,

$$\frac{\mathrm{d}u(x(x,t),t)}{\mathrm{dt}} = 0,$$

thus $u(x(X,t),t) = u_0(X)$. We obtain the characteristic curve equation

$$x(X,t) = \frac{3}{2}u_0(X)t + X.$$

To design a $\mathcal{C}^{N_s}$ smooth solution, we initialize the problem with the following initial data:

$$\eta_0 = \frac{u_0^2}{4\,g} \qquad \text{and} \qquad q_0 = \frac{u_0^3}{4\,g},$$

with the following $\mathcal{C}^{N_s}$ smooth initial velocity

$$u_0(x) = \begin{cases} 1 & \text{if } \; x \leq 0, \\ e^{-x^{\,N_s+1}} & \text{elsewhere.} \end{cases}$$

The method of characteristics provides us with the expression of the analytical solution, for any given $t \in [0, t_c[$:

$$u(x,t) = \begin{cases} 1 & \text{if } \; x \leq \frac{3}{2}\,t, \\ e^{-X^{N_s+1}} & \text{elsewhere,} \end{cases} \tag{B.2}$$

where the characteristic lines read $x(X,t) = \frac{3}{2}\,e^{-X^{N_s+1}}\,t + X$. For practical applications, to assess the position of the characteristic line origin point X given x and t, one may use an iterative root-finding process, as Newton's method, to solve the non-linear problem $g(X) = 0$, where for given x and t function $g(X) = \frac{3}{2}\,e^{-X^{N_s+1}}\,t + X - x$.

The analytical solution, (B.2), is defined $\forall\, t \in \; [0,\, t_c[$, where the critical time at which the characteristic lines cross is defined as follows:

$$t_c = \frac{2\;e^{\frac{N_s}{N_s+1}}}{3\;(N_s+1)^{\frac{1}{N_s+1}}\;N_s^{\frac{N_s}{N_s+1}}}.$$

Appendix C

Cut-off function

The cut-off function $\varphi \in \mathcal{D}(\mathbb{R})$ used in (6.23) is defined as follows:

$$\forall x \in \mathbb{R}, \quad \varphi(x) := e\,\psi_e(\varepsilon_0 x),$$

where

$$\forall x \in \mathbb{R}, \quad \psi_e(x) := \phi_e(1 - |x|^2),$$

and

$$\forall t \in \mathbb{R}, \quad \phi_e(t) := \begin{cases} e^{-t^{-1}} & \text{if} \quad t > 0 \\ \\ 0 & \text{elsewhere,} \end{cases}$$

Note that we have $\text{supp}(\psi_e) \subset \overline{B}(0,1)$, $\text{supp}(\varphi) \subset [-\frac{1}{\varepsilon_0}, \frac{1}{\varepsilon_0}]$ and ε_0 chosen such that we have $\varphi(x) = 1$, $\forall |x| \leq 1$.

Appendix D

Definition of the elliptic object

In this work, we consider a partially immersed object $\mathcal{O}_{\rm bj}$, which center of mass is located initially at $(x_{\rm G}(0), z_{\rm G}(0))$ and which boundary is denoted by $\partial\mathcal{O}_{\rm bj}$. Denoted respectively by a, b its major and minor radius, we define $\partial\mathcal{O}_{\rm bj}$ as an ellipse, so that we have:

$$(x, y) \in \partial\mathcal{O}_{\rm bj} \quad \Longleftrightarrow \quad \frac{(x - x_{\rm G}(0))^2}{a^2} + \frac{(z - z_{\rm G}(0))^2}{b^2} = 1.$$

The underside of the object may be locally parameterized as follows:

$$\forall x \in \mathcal{I}_{\rm lid} := [x_{\rm G}(0) - a, x_{\rm G}(0) + a], \quad \eta_{\rm lid}(x) := z_{\rm G}(0) - b\sqrt{1 - \frac{(x - x_{\rm G}(0))^2}{a^2}}.$$

We place the body at the initial time in a lake at rest of water elevation H_0, so that $z_G(0)$ is defined as, $z_G(0) = H_0 + e_0$, where e_0 is the height of the center of the gravity above the water surface at initial time. Proceeding this way one can easily determine the position of the contact points at initial time, that is:

$$X_0^{\pm} = \varkappa_{\pm}(0) = x_G(0) \pm \sqrt{a^2 - \frac{a^2 {e_0}^2}{b^2}}.$$

Appendix E

Mass and inertia coefficient of the elliptic object

When embedding a partly immersed object in a free-surface flow which is initially in a motionless steady-state, one may need to partly immerse it in such a way that the whole system "fluid-object" is also in equilibrium. In practice, for a given structure that comes with its (already defined) own mechanical properties (center of mass $\mathcal{M}_G$, boundary profile η_{lid}, mass m_o and inertia coefficient i_o), this boils down to accurately define the location $\boldsymbol{\mathcal{X}}_{\text{G}}(0)$ to ensure the balancing of Newton's law. That being said, for our numerical study, we choose to consider the following simpler "inverse" strategy: for a given object profile η_{lid}, and a given initial location $\boldsymbol{\mathcal{X}}_{\text{G}}(0)$, we define the mass and inertia coefficients m_o and i_o such that Newton's law are initially balanced.
Hence, our elliptic structure may be located with $(x_G(0), z_G(0)) = (x_G(0), H_0 + e_0)$ in a motionless steady-state flow configuration. For a given value of e_0 (and assuming $\theta(0) = 0$), this boils down to choosing m_o such that the initial acceleration of $\mathcal{M}_G$ is equal to zero:

$$-m_o \mathrm{g}\mathbf{e}_{\mathrm{z}} + \int_{\mathcal{I}(0)} \left(\underline{\mathrm{p}}^i - \mathrm{p}_{\mathrm{atm}}\right) \mathbf{n}^{\mathrm{i}} = 0. \tag{E.1}$$

Considering the second equation of (E.1), for a motionless steady-state, we get:

$$m_o = \int_{\mathcal{I}(0)} \rho(H_0 - \eta_{\text{lid}}(x))\, dx, \tag{E.2}$$

and the corresponding inertia coefficient is defined by:

$$i_o = \frac{m_o(a^2 + b^2)}{5}.$$

At the discrete level, considering the semi-discrete equations (6.48e) or the fully-discrete equation (6.75c), m_o is defined such that the initial discrete acceleration of $\mathcal{M}_G$ is equal to zero:

$$\begin{pmatrix} -m_o \mathrm{g}\mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} = \rho \mathfrak{S}^{h,n_{\mathrm{g}}}_{\mathcal{I}(0)} \left((f^{\star,0}_{1,h} + f^{\star,0}_{3,h}) \frac{\boldsymbol{\mathcal{T}}^{\star,0}_{\mathrm{G},h}}{H^{\mathrm{i},n}_h} \right).$$

Appendix F

Newton's second law and *added-mass* effect

The pressure $\underline{\mathrm{p}}^i$ satisfies the boundary value problem (5.31), whose solvability is guaranteed by (5.29). Then, $\underline{\mathrm{p}}^i$ satisfies:

$$\partial_x \underline{\mathrm{p}}^i = -\frac{\rho}{H_{\mathrm{i}}} \left(f_1^\star + f_2^\star + f_3^\star \right). \tag{F.1}$$

Noticing that,

$$\partial_x \boldsymbol{\mathcal{T}}_{\mathrm{G}} = \begin{pmatrix} -\mathbf{n}^{\mathrm{i}} \\ \mathbf{r}_{\mathrm{G}}^{\perp} \cdot \mathbf{n}^{\mathrm{i}} \end{pmatrix}, \tag{F.2}$$

and by using an integration by parts and the boundary condition $\underline{\mathrm{p}}^i = \mathrm{p}_{\mathrm{atm}}$ on $\chi_{\pm}$, we can rewrite (5.22) as:

$$\mathbb{M}_0 \boldsymbol{\vartheta}'_{\mathrm{G}} \;=\; -\begin{pmatrix} m_o\, \mathrm{g}\, \mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} + \int_{\mathcal{I}(t)} \left(\partial_x \underline{\mathrm{p}}^i \right) \boldsymbol{\mathcal{T}}_{\mathrm{G}}^\star, \tag{F.3}$$

by using (F.1), we have,

$$\mathbb{M}_0 \boldsymbol{\vartheta}'_{\mathrm{G}} \;=\; -\begin{pmatrix} m_o\, \mathrm{g}\, \mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho \int_{\mathcal{I}(t)} \left(f_1^* + f_2^* + f_3^* \right) \frac{\boldsymbol{\mathcal{T}}_{\mathrm{G}}^*}{H_{\mathrm{i}}}, \tag{F.4}$$

and by using the definition of f_2 in (5.30), (5.22) is finally reduced to the following ODE:

$$\left(\mathbb{M}_0 + \mathbb{M}_{\mathrm{a}} \left(H^i, \boldsymbol{\mathcal{T}}_{\mathrm{G}} \right) \right) \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\vartheta}_{\mathrm{G}} = \begin{pmatrix} -m_o \mathrm{g} \mathbf{e}_{\mathrm{z}} \\ 0 \end{pmatrix} - \rho \int_{\mathcal{I}(t)} \left(f_1^\star + f_3^\star \right) \frac{\boldsymbol{\mathcal{T}}_{\mathrm{G}}^\star}{H^i}. \tag{F.5}$$

We refer the reader to [87] for a complete description.

Bibliography

[1] R. Abgrall. Some remarks about conservation for residual distribution schemes. *Computational Methods in Applied Mathematics*, 18(3):327–351, 2018.

[2] V. Aizinger and C. Dawson. A discontinuous Galerkin method for two-dimensional flow and transport in shallow water. *Advances in Water Resources*, 25:67–84, 2002.

[3] F. Alcrudo and P. Garcia-Navarro. A high-resolution godunov-type scheme in finite volumes for the 2d shallow-water equations. *Internat. J. Numer. Methods Fluids*, 1993.

[4] Y. Allaneau and A. Jameson. Connections between the filtered discontinuous Galerkin method and the flux reconstruction approach to high order discretizations. *Comput. Meth. Appl. Mech. Engrg.*, 200:3628–3636, 2011.

[5] V.R. Ambati and O. Bokhove. Space–time discontinuous galerkin finite element method for shallow water flows. *Journal of Computational and Applied Mathematics*, 204:452–462, 2007.

[6] K Anastasiou and CT Chan. Solution of the 2d shallow water equations using the finite volume method on unstructured triangular meshes. *International Journal for Numerical Methods in Fluids*, 24(11):1225–1245, 1997.

[7] K. Anastasiou and C.T. Chan. Solution of the 2d shallow water equations using the finite volume method on unstructured triangular meshes. *Int J Numer Methods Fluids*, 24:1225–1245, 1997.

[8] L. Arpia and M. Ricchiuto. Well balanced residual distribution for the ALE spherical shallow water equations on moving adaptive meshes. *J. Comput. Phys.*, 405:109–173, 2019.

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

[10] D. Balsara, C. Altmann, C.D. Munz, and M. Dumbser. A sub-cell based indicator for troubled zones in RKDG schemes and a novel class of hybrid RKDG+HWENO schemes. *J. Comp. Phys.*, 226:586–620, 2007.

[11] S.R.M. Barros and J.W. Cardenas. A nonlinear galerkin method for the shallow-water equations on periodic domains. *J. Comput. Phys.*, 172:592–608, 2001.

[12] 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.

[13] C. Beels, P. Troch, K. De Visch, J.P. Kofoed, and G De Backer. Application of the time-dependent mild-slope equations for the simulation of wake effects in the lee of a farm of wave dragon wave energy converters. *Renewable Energ*, 35(8):1644–1661, 2010.

[14] A. Bermudez, A. Dervieux, J.-A. Desideri, and M.E. Vazquez. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Comput. Methods Appl. Mech. Engrg.*, 155:49–72, 1998.

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

[16] P.-E. Bernard, J.-F. Remacle, R. Comblen, V. Legat, and K. Hillewaert. High-order discontinuous galerkin schemes on general 2d manifolds applied to the shallow water equations. *J. Comput. Phys.*, 228:6514–6535, 2009.

[17] M. Berndt, J. Breil, S. Galera, M. Kucharik, P.-H. Maire, and M. Shashkov. Two-step hybrid conservative remapping for multimaterial arbitrary lagrangian–eulerian methods. *Journal of Computational Physics*, 230(17):6664–6687, 2011.

[18] 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(5):2587–2612, 2008.

[19] C. Berthon, M. M'Baye, M.H. Le, and D. Seck. A well-defined moving steady states capturing godunov-type scheme for shallow-water model. *Int. J. Finite Volume*, 15, 2020.

[20] R. Biswas, K.D. Devine, and J.E. Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14:255 – 283, 1994.

[21] O. Bokhove. Flooding and drying in finite-element discretizations of shallow-water equations. part 2: Two dimensions. 2003.

[22] O. Bokhove. Flooding and drying in discontinuous galerkin finite-element discretizations of shallow-water equations. part 1: One dimension. *J. Sci. Comput.*, 22-23:47–82, 2005.

[23] W. Boscheri, R. Loubere, and M. Dumbser. Direct arbitrary-lagrangian–eulerian ader-mood finite volume schemes for multidimensional hyperbolic conservation laws. *Journal of Computational Physics*, 292:56–87, 2015.

[24] U. Bosi, A.P. Engsig-Karup, C. Eskilsson, and M. Ricchiuto. A spectral/hp element depth-integrated model for nonlinear wave–body interaction. *Comp. Meth. App. Mech. Eng.*, 348:22–249, 2019.

[25] F. Bouchut, J. Le Sommer, and V. Zeitlin. Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. part 2. high-resolution numerical simulations. *Journal of Fluid Mechanics*, 514:35–63, 2004.

[26] J. Boussinesq. Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *Journal de Mathématiques Pures et Appliquées*, pages 55–108, 1872.

[27] K Budal. Theory for absorption of wave power by a system of interacting bodies. *Journal of Ship Research*, 21(4), 1977.

[28] S. Bunya, E. J. Kubatko, J. J. Westerink, and C. Dawson. A wetting and drying treatment for the runge-kutta discontinuous galerkin solution to the shallow water equations. *Comput. Methods Appl. Mech. Engrg*, 198:1548–1562, 2009.

[29] A. Burbeau, P. Sagaut, and C.-H. Bruneau. A problem-independent limiter for high-order Runge-Kutta discontinuous Galerkin methods. *J. Comput. Phys.*, 169(1):111 – 150, 2001.

[30] S. Busto, M. Dumbser, C. Escalante, N. Favrie, and S. Gavrilyuk. On high order ader discontinuous galerkin schemes for first order hyperbolic reformulations of nonlinear dispersive systems. *Journal of Scientific Computing*, 87(2):1–47, 2021.

[31] S. Busto, M. Dumbser, S. Gavrilyuk, and K. Ivanova. On thermodynamically compatible finite volume methods and path-conservative ader discontinuous galerkin schemes for turbulent shallow water flows. *Journal of Scientific Computing*, 88(1):1–45, 2021.

[32] A. Canestrelli, A. Siviglia, M. Dumbser, and E.F. Toro. Well-balanced high-order centred schemes for non-conservative hyperbolic systems. applications to shallow water equations with fixed and mobile bed. *Advances in Water Resources*, 32(6):634–644, 2009.

[33] G. Carrier and H. Greenspan. Water waves of finite amplitude on a sloping beach. *Journal of Fluid Mechanics*, 2:97—109, 1958.

[34] E. Casoni, J. Peraire, and A. Huerta. One-dimensional shock-capturing for high-order discontinuous Galerkin methods. *Int. J. Numer. Meth. Fluids*, 71:737–755, 2013.

[35] M.J. Castro Diaz, J.A. Lopez-Garcia, and C. Parès. High order exactly well-balanced numerical methods for shallow water systems. *J. Comput. Phys.*, 246:242–264, 2013.

[36] P. Caussignac and R. Touzani. Solution of three-dimensional boundary layer equations by a discontinuous finite element method, part i: Numerical analysis of a linear model problem. *Computer methods in applied mechanics and engineering*, 78(3):249–271, 1990.

[37] P. Caussignac and R. Touzani. Solution of three-dimensional boundary layer equations by a discontinuous finite element method, part ii: Implementation and numerical results. *Computer methods in applied mechanics and engineering*, 79(1):1–20, 1990.

[38] G. Chavent and B. Cockburn. The local projection-discontinuous-galerkin finite element method for scalar conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, 23(4):565–592, 1989.

[39] Q. Chen and I. Babuska. Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Comput. Methods Appl. Mech. Engrg*, 128:405–417, 1995.

[40] X.N. Chen and S.D. Sharma. A slender ship moving at a near-critical speed in a shallow channel. *J. Fluid Mech.*, 291:263–285, 1995.

[41] A. Chertock, S. Cui, A. Kurganov, and T. Wu. Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms. *International Journal for numerical methods in fluids*, 78(6):355–383, 2015.

[42] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws—multi-dimensional optimal order detection (mood). *J. Comput. Phys.*, 230:4028–4050, 2011.

[43] B. Cockburn, S. Hou, and C.-W. Shu. The runge-kutta local projection discontinuous galerkin finite element method for conservation laws. iv. the multidimensional case. *Mathematics of Computation*, 54(190):545–581, 1990.

[44] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta Discontinuous Galerkin Method for Conservation Laws V: Multidimensional Systems. *J. Comp. Phys.*, 141:199–224, 1998.

[45] B. Cockburn, S.Y. Lin, and C.-W. Shu. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws iii: One-dimensional systems. *J. Comput. Phys.*, 84(1):90 – 113, 1989.

[46] B. Cockburn and C.-W. Shu. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework. *Mathematics of computation*, 52(186):411–435, 1989.

[47] B. Cockburn and C.-W. Shu. The runge-kutta local projection-discontinuous-galerkin finite element method for scalar conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, 25(3):337–361, 1991.

[48] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems. *J. Comput. Phys.*, 141(2):199 – 224, 1998.

[49] J. N. de la Rosa and C. D. Munz. Hybrid DG/FV schemes for magnetohydrodynamics and relativistic hydrodynamics. *Comp. Phys. Commun.*, 222:113–135, 2018.

[50] A.J.-C. de Saint-Venant. Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit. *C.R. Acad. Sci. Paris, Section Mécanique*, 73:147–154, 1871.

[51] O Delestre. *Simulation du ruissellement d'eau de pluie sur des surfaces agricoles*. PhD thesis, Université d'Orléans; Université d'Orléans, 2010.

[52] O Delestre and F Marche. A numerical scheme for a viscous shallow water model with friction. *Journal of Scientific Computing*, 48(1):41–51, 2011.

[53] D. A. Di Pietro and A. Ern. Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations. *Math. Comp.*, 79(271):1303–1330, 2010.

[54] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69 of *Mathématiques and Applications*. Springer, 2012.

[55] S. Diot, S. Clain, and R. Loubère. Improved detection criteria for the multi-dimensional optimal order detection (MOOD) on unstructured meshes with very high-order polynomials. *Computers and Fluids*, 64:43–63, 2012.

[56] S. Diot, R. Loubère, and S. Clain. The MOOD method in the three-dimensional case: very-high-order finite volume method for hyperbolic systems. *Int. J. Numer. Meth. Fluids*, 73:362–392, 2013.

[57] D.Kuzmin. Slope limiting for discontinuous galerkin approximations with a possibly non-orthogonal taylor basis. *Int J Numer Methods Fluids*, 71(9):1178–1190, 2013.

[58] J. Donea, A. Huerta, J.-Ph. Ponthot, and A. Rodríguez-Ferran. *Arbitrary Lagrangian–Eulerian Methods, The Encyclopedia of Computational Mechanics*, pages 413–437. Wiley, 2004.

[59] M. Dumbser and R. Loubère. A simple robust and accurate a posteriorisub-cell finite volume limiter for the discontinuousGalerkin method on unstructured meshes. *J. Comp. Phys.*, 319:163–199, 2016.

[60] M. Dumbser, A. Uuriintsetseg, and O. Zanotti. On arbitrary-lagrangian-eulerian one-step weno schemes for stiff hyperbolic balance laws. *Communications in Computational Physics*, 14(2):301–327, 2013.

[61] M. Dumbser, O. Zanotti, R. Loubère, and S. Diot. A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *J. Comput. Phys.*, 278:47–75, 2014.

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

[63] K.S. Erduran, V. Kutija, and C.J.M. Hewett. Performance of finite volume solutions to the shallow water equations with shock-capturing schemes. *Int J Numer Methods Fluids*, 40:1237–1273, 2002.

[64] 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.

[65] C. Eskilsson and S.J.Sherwin. Discontinuous galerkin spectral/hp element modelling of dispersive shallow water systems. *J. Sci. Comput.*, 22:269–288, 2005.

[66] M. Feistauer, V. Dolejší, and V. Kučera. On the discontinuous Galerkin method for the simulation of compressible flow with wide range of Mach numbers. *Computing and Visualization in Science*, 10(1):17–27, 2007.

[67] L. Fraccarollo and E.F. Toro. Experimental and numerical assessment of the shallow water model for two-dimensional dam-break type problems. *J. Hydraulic Res.*, 33(6):843–863, 1995.

[68] J.M. Gallardo, C. Parés, and M Castro. On a well-balanced high-order finite volume scheme for shallow water equations with topography and dry areas. *J. Comput. Phys.*, 227(1):574–601, 2007.

[69] H. Gao and Z. J. Wang. A conservative correction procedure via reconstruction formulation with the Chain-Rule divergence evaluation. *J. Comp. Phys.*, 232:7–13, 2013.

[70] J.-F. Gerbeau and B. Perthame. *Derivation of viscous Saint-Venant system for laminar shallow water; numerical validation.* PhD thesis, INRIA, 2000.

[71] F.X. Giraldo, J.S. Hesthaven, and T. Warburton. Nodal high-order discontinuous galerkin methods for the spherical shallow water equations. *J. Comput. Phys.*, 181(2):499 – 525, 2002.

[72] E. Godlewski, M. Parisot, J. Sainte-Marie, and F. Wahl. Congested shallow water model: roof modelling in free surface flow. *ESAIM Math. Model. Numer. Anal.*, 52(5):1679 – 1707, 2018.

[73] E. Godlewski, M. Parisot, J. Sainte-Marie, and F. Wahl. Congested shallow water model: on floating body. *SMAI Journal of Computational Mathematics*, in press, 2022.

[74] S. Gottlieb, C.-W. Shu, and Tadmor E. Strong stability preserving high order time discretization methods. *SIAM Review*, 43:89–112, 2001.

[75] N Goutal. *Proceedings of the 2nd workshop on dam-break wave simulation.* Department Laboratoire National d'Hydraulique, Groupe Hydraulique Fluviale, 1997.

[76] A.E Green and P.M. Naghdi. A derivation of equations for wave propagation in water of variable depth. *Journal of Fluid Mechanics*, 78(2):237–246, 1976.

[77] J.-L. Guermond, R. Pasquetti, and B. Popov. Entropy viscosity method for nonlinear conservation laws. *J. Comput. Phys.*, 230:4248–4267, 2011.

[78] H. Guillard and C. Farhat. On the significance of the geometric conservation law for flow computations on moving meshes. *Comput. Methods Appl. Engrg*, 190:1467–1482, 2000.

[79] C. Hague and C. Swan. A multiple flux boundary element method applied to the description of surface water waves. *J. Comput. Phys.*, 228(14):5111–5128, 2009.

[80] A. Haidar, F. Marche, and F. Vilar. A posteriori finite-volume local subcell correction of high-order discontinuous galerkin schemes for the nonlinear shallow-water equations. *Journal of Computational Physics*, page 110902, 2021.

[81] R. Harris, Z.J. Wang, and Y. Liu. Efficient Quadrature-Free High-Order Spectral Volume Method on Unstructured Grids: Theory and 2D Implementation. *J. Comp. Phys.*, 227:1620–1642, 2008.

[82] H.S. Hassan, K.T. Ramadan, S.N. Hanna, et al. Numerical solution of the rotating shallow water flows with topography using the fractional steps method. *Applied Mathematics*, 1(02):104, 2010.

[83] C.W. Hirt, A.A. Amsden, and J.L. Cook. An arbitrary lagrangian–eulerian computing method for all flow speed. *J. Comput. Phys.*, 135:203–216, 1997.

[84] A. Huerta, E. Casoni, and J. Peraire. A simple shock-capturing technique for high-order discontinuous galerkin methods. *Int. J. Numer. Meth. Fluids*, 69:1614–1632, 2012.

[85] H. T. Huynh, Z. J. Wang, and P. E. Vincent. High-order methods for computational fluid dynamics: A brief review of compact differential formulations on unstructured grids. *J. Comp. Phys.*, 98:209–220, 2014.

[86] H.T. Huynh. A flux reconstruction approach to high-order schemes including discontinuous galerkin methods. In *18th AIAA computational fluid dynamics conference*, page 4079, 2007.

[87] T. Iguchi and D. Lannes. Hyperbolic free boundary problems and applications to wave-structure interactions. *Indiana Univ. Math. J.*, 70:353–464, 2021.

[88] M. Ioriatti and M. Dumbser. A posteriori sub-cell finite volume limiting of staggered semi-implicit discontinuous Galerkin schemes for the shallow water equations. *Applied Numerical Mathematics*, 135:443–480, 2019.

[89] M. Iskandarani, D. B. Haidvogel, and J. P. Boyd. A staggered spectral element model with application to the oceanic shallow water equations. *Int J Numer Methods Fluids*, 20(5):393–414, 1995.

[90] J. S. Park and C. Kim. Hierarchical multi-dimensional limiting strategy for correction procedure via reconstruction. *J. Comp. Phys.*, 308:57–80, 2016.

[91] J. S. Park and S.-H. Yoon and C. Kim. Multi-dimensional limiting process for hyperbolic conservation laws on unstructured grids. *J. Comp. Phys.*, 229:788–812, 2010.

[92] J. Jaffre, C. Johnson, and A. Szepessy. Convergence of the discontinuous galerkin finite element method for hyperbolic conservation laws. *Mathematical Models and Methods in Applied Sciences*, 5(03):367–386, 1995.

[93] F. James, P.-Y. Lagrée, M.H. Le, and M. Legrand. Towards a new friction model for shallow water equations through an interactive viscous layer. *ESAIM: Mathematical Modelling and Numerical Analysis*, 53(1):269–299, 2019.

[94] G. Jiang and C.-W. Shu. Efficient implementation of weighted eno schemes. *J. Comput. Phys.*, 126(1):202–228, 1996.

[95] T. Jiang. *Ship Waves in Shallow Water*. Verkehrstechnik, Fahrzeugtechnik. Fortschritt-Berichte VDIReihe, 2001.

[96] T. Jiang, R. Henn, and S.D. Sharma. Wash waves generated by ships moving on fairways of varying topography. In *24th Symposium on Naval Hydrodynamics Fukuoka, JAPAN*, 2002.

[97] F. John. On the motion of floating bodies i. *Communications on Pure and Applied Mathematics*, 2:13–57, 1949.

[98] C. Johnson and J. Pitkäranta. An analysis of the discontinuous galerkin method for a scalar hyperbolic equation. *Mathematics of computation*, 46(173):1–26, 1986.

[99] 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.

[100] G. Kesserwani and Q. Liang. Well-balanced RKDG2 solutions to the shallow water equations over irregular domains with wetting and drying. *Comput. Fluids*, 39(10):2040–2050, 2010.

[101] G. Kesserwani, Q. Liang, J. Vazquez, and R. Mose. Well-balancing issues related to the RKDG2 scheme for the shallow water equations. *Int. J. Numer. Meth. Fluids*, 62:428–448, 2010.

[102] R.M. Kirby and S.J. Sherwin. Stabilisation of spectral / hp element methods through spectral vanishing viscosity: Application to fluid mechanics. *Comput. Methods Appl. Mech. Engrg.*, 195:3128–3144, 2006.

[103] E.V. Koutandos, T.V. Karambas, and C.G. Koutitas. Floating breakwater response to waves action using a boussinesq model coupled with a 2dv elliptic solver. *J. Waterw. Port Coastal Ocean Eng.*, 130:243–255, 2004.

[104] E.E. Kriezi, T. Karambas, P. Prinos, and C. Koutitas. Interaction of floating breakwaters with waves in shallow waters. In *Proc., Int. Conf. IAHR, Beijing*, 2001.

[105] L. Krivodonova. Limiters for high-order discontinuous Galerkin methods. *J. Comp. Phys.*, 226:879–896, 2007.

[106] D. Lannes. *The water waves problem: mathematical analysis and asymptotics.* Number 188 in Mathematical Surveys and Monographs. American Mathematical Society, 2013.

[107] D. Lannes. On the dynamics of floating structures. *Annals of PDE*, 3(1):11, 2017.

[108] C. Lee and J.N. Newman. Computation of wave effects using the panel method. In *Numerical Models in Fluid-Structure Interaction*, 2005.

[109] 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.

[110] P. Lesaint. Finite element methods for symmetric hyperbolic equations. *Numerische Mathematik*, 21(3):244–255, 1973.

[111] P. Lesaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation. *Publications mathématiques et informatique de Rennes*, (S4):1–40, 1974.

[112] H. Li and R.-X. Liu. The discontinuous Galerkin finite element method for the 2d shallow water equations. *Mathematics and Computers in Simulation*, 56:223–233, 2001.

[113] L. Li and Q. Zhang. A new vertex-based imiting approach for nodal discontinuous galerkin methods on arbitrary unstructured meshes. *Comput. Fluids*, 159:316–326, 2017.

[114] C. Liang, F. Ham, and E. Johnsen. Discontinuous galerkin method with weno limiter for flows with discontinuity. *Center for Turbulence Research 335 Annual Research Briefs 2009*, 2009.

[115] Q. Liang and A. G. L. Borthwick. Adaptive quadtree simulation of shallow flows with wet-dry fronts over complex topography. *Comput. Fluids*, 38(2):221–234, 2009.

[116] 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.

[117] R. Liska, M. Shashkov, P. Váchal, and B. Wendroff. Synchronized flux corrected remapping for ale methods. *Computers & fluids*, 46(1):312–317, 2011.

[118] I. Lomtev, R.M. Kirby, and G.E. Karniadakis. A discontinuous Galerkin ALE method for compressible viscous flows in moving domains. *J. Comput. Phys.*, 155(1):128–159, 1999.

[119] M. Lukacova, S. Noelle, and M. Kraft. Well-balanced finite volume evolution galerkin methods for the shallow water equations. *J. Comput. Phys.*, 221(1):122–147, 2007.

[120] M. Lukacova-Medvidova, S. Noelle, and M. Kraft. Well-balanced finite volume evolution galerkin methods for the shallow water equations. *Journal of Computational Physics*, 221:122–147, 01 2007.

[121] H. Ma. A spectral element basin model for the shallow water equations. *J. Comput. Phys.*, 109(1):133 – 149, 1993.

[122] F. Marche. Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects. *European Journal of Mechanics-B/Fluids*, 26(1):49–63, 2007.

[123] A. Meister and S. Ortleb. On unconditionally positive implicit time integration for the DG scheme applied to shallow water flows. *Int. J. Numer. Meth. Fluids*, 76::69–94, 2014.

[124] A. Meister and S. Ortleb. A positivity preserving and well-balanced DG scheme using finite volume subcells in almost dry regions. *Appl. Math. Comp.*, 272:259–273, 2016.

[125] V. Michel-Dansac. A well-balanced scheme for the shallow-water equations with topography. *Computers and Mathematics with Applications*, 72(3):568–593, 2016.

[126] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography. *Comput. Math. Appl.*, 72:Comput. Math. Appl, 2016.

[127] H. Mirzaee, L. Ji, J.K. Ryan, and R.M Kirby. Smoothness-increasing accuracy-conserving (SIAC) post- processing for discontinuous Galerkin solutions over structured triangular meshes. *SIAM J. Numer. Anal.*, 49(5):1899–1920, 2011.

[128] R. D. Nair, S. J. Thomas, and R. D. Loft. A discontinuous galerkin global shallow water model. *Monthly Weather Review*, 133:876–888, 2004.

[129] I.M. Navon. Finite-element simulation of the shallow-water equations model on a limited-area domain. *Appl. Math. Modelling*, 3, 1979.

[130] S. Noelle, N. Pankratz, G. Puppo, and J.R. Natvig. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *J. Comput. Phys.*, 213(2):474–499, 2006.

[131] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced finite volume weno schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1):29–58, 2007.

[132] T.F. Ogilvie. Second-order hydrodynamic effects on ocean platforms. In *Proc. Intl. Workshop on Ship and Platform Motions, ed. R. W. Yeung, University of California, Berkeley*, pages 205–265, 1983.

[133] H. T. Ozkan-Haller and J.T.Kirby. A fourier-chebyshev collocation method for the shallow water equations including shoreline. *Applied Ocean Research*, 19:21–34, 1997.

[134] P.-O.Persson and J.Peraire. Sub-cell shock capturing for discontinuous galerkin methods. *AIAA Aerospace Sciences Meeting and Exhibit*, 112, 2006.

[135] K.T. Panourgiasa and J.A. Ekaterinaris. A nonlinear filter for high order discontinuous Galerkin discretizations with discontinuity resolution within the cell. *J. Comput. Phys.*, 326:234–257, 2016.

[136] J. Patera and V. Nassehi. A new two-dimensional finite element model for the shallow water equations using a lagrangian framework constructed along fluid particle trajectories. *Int. J. Numer. Meth. Fluids*, 39:4159–4182, 1996.

[137] P.-O. Persson, J. Peraire, and J. Bonet. Discontinuous Galerkin solution of the Navier–Stokes equations on deformable domains. *Comp. Meth. App. Mech. Eng.*, 198:1585–1595, 2009.

[138] B. Perthame and Y. Qiu. A variant of Van Leer's method for multidimensional systems of conservation laws. *J. Comput. Phys.*, 112(2):370–381, 1994.

[139] C.S. Peskin. The immersed boundary method the immersed boundary method. *Acta Numerica*, pages 479–517, 2002.

[140] J. Qiu and C.-W. Shu. A comparison of troubled-cell indicators for Runge-Kutta discontinuous Galerkin methods using weighted essentially nonoscillatory limiters. *SIAM J. Sci. Comput.*, 27:995–1013, 2005.

[141] J. Qiu and C.-W. Shu. Runge Kutta discontinuous Galerkin method using WENO limiters. *SIAM J. Sci. Comput.*, 26:907–929, 2005.

[142] W.H. Reed and T.R. Hill. Triangular mesh methods for the neutron transport equation. Technical report, Los Alamos Scientific Lab., N. Mex.(USA), 1973.

[143] M. Ricchiuto. An explicit residual based approach for shallow water flows. *J. Comput. Phys.*, 280:306–344, 2015.

[144] M. Ricchiuto and A. Bollermann. Stabilized residual distribution for shallow water simulations. *J. Comput. Phys.*, 228:1071–1115, 2009.

[145] 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.

[146] G. Russo and G. Puppo. High-order well-balanced schemes, in numerical methods for relaxation systems and balance equations. *eds., Quaderni di Matematica, Dipartimento di Matematica, Seconda Universita di Napoli, Italy*, 2009.

[147] D. Schwanenberg and M. Harms. Discontinuous galerkin finite-element method for transcritical two-dimensional shallow water flows. *J.Hydraul. Eng.*, 130(5):412–421, 2004.

[148] C.-W. Shu. Tvb uniformly high-order schemes for conservation laws. *Mathematics of Computation*, 49(179):105–121, 1987.

[149] C.-W. Shu and S. Osher. Efficient implementation of Essentially Non-Oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77:439–471, 1988.

[150] T. Song, A. Main, G. Scovazzi, and M. Ricchiuto. The shifted boundary method for hyperbolic systems: Embedded domain computations of linear waves and shallow water flows. Technical report, Inria, 2017.

[151] M. Sonntag and C.D. Munz. Shock capturing for discontinuous Galerkin methods using finite volume subcells. In *Finite Volumes for Complex Applications VII-Elliptic, Parabolic and Hyperbolic Problems*, pages 945–953, 2014.

[152] C.E. Synolakis, E.N. Bernard, V.V. Titov, U Kanoglu, and F.I. Gonzalez. Standards, criteria, and procedures for noaa evaluation of tsunami numerical models. *NOAA Tech. Memo.*, OAR PMEL-135, 2007.

[153] J. Tanner and E. Tadmor. Adaptive mollifiers - high resolution recover of piecewise smooth data from its spectral information. *Found. Comput. Math.*, 2:155–189, 2002.

[154] P.A. Tassi, O. Bokhove, and C.A. Vionnet. Space discontinuous galerkin method for shallow water flows—kinetic and hllc flux, and potential vorticity generation. *Advances in water resources*, 30(4):998–1015, 2007.

[155] P.D. Thomas and C.K. Lombard. Geometric conservation law and its applications to flow computations on moving grids. *AIAA Journal*, 17:1030–1037., 1979.

[156] E.F. Toro. *Shock-capturing methods for free-surface shallow flows*. Chichester: John Wiley and Sons, 2001.

[157] T. Utnes. A finite element solution of the shallow-water wave equations. *Appl. Math. Modelling*, 14:20–29, 1990.

[158] J.J.W. van de Vegt and Y. Xu. Space-time discontinuous galerkin method for nonlinear water waves. *J. Comput. Phys.*, 224:17–39, 2007.

[159] J.J.W. Van der Vegt and H. Van der Ven. Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flows. *J. Comput. Phys.*, 182:546–585, 2002.

[160] H. Vandeven. Family of spectral filters for discontinuous problems. *J. Sci. Comput.*, 8:159–192, 1991.

[161] S. Vater, N. Beisiegel, and J. Behrens. A limiter-based well-balanced discontinuous galerkin method for shallow-water flows with wetting and drying: One-dimensional case. *Advances in Water Resources*, 85:1–13, 2015.

[162] C.S. Ventakasubban. A new finite element formulation for ALE (arbitrary Lagrangian Eulerian) compressible fluid mechanics. *Int. J. Engrg. Sci.*, 33:1743–1762, 1995.

[163] F. Vilar. *A high-order discontinuous Galerkin discretization for solving two-dimensional Lagrangian hydrodynamics*. PhD thesis, Université Bordeaux I, 2012.

[164] F. Vilar. A posteriori correction of high-order discontinuous galerkin scheme through subcell finite volume formulation and flux reconstruction. *J. Comput. Phys.*, 387:245–279, 2019.

[165] F. Vilar and R. Abgrall. *A Posteriori* local subcell correction of high-order discontinuous galerkin scheme for conservation laws on two-dimensional unstructured grids. *SIAM J. Num. Anal.*, Under preparation.

[166] F. Vilar, P.-H. Maire, and R. Abgrall. A discontinuous galerkin discretization for solving the two-dimensional gas dynamics equations written under total lagrangian formulation on general unstructured grids. *Journal of Computational Physics*, 276:188–234, 2014.

[167] F. Vilar, P.H. Maire, and R. Abgrall. Cell-centered discontinuous Galerkin discretizations for two-dimensional scalar conservation laws on unstructured grids and for one-dimensional Lagrangian hydrodynamics. *Comput. Fluids*, 46:498–504, 2011.

[168] F. Vilar, C.-W. Shu, and P.-H. Maire. Positivity-preserving cell-centered lagrangian schemes for multi-material compressible flows: From first-order to high-orders. part i: The one-dimensional case. *Journal of Computational Physics*, 312:385–415, 2016.

[169] F. Vilar, C.-W. Shu, and P.-H. Maire. Positivity-preserving cell-centered lagrangian schemes for multi-material compressible flows: From first-order to high-orders. part ii: The two-dimensional case. *Journal of Computational Physics*, 312:416–442, 2016.

[170] P. E. Vincent, P. Castonguay, and A. Jameson. A New Class of High-Order Energy Stable Flux Reconstruction Schemes. *J. Sci. Comput.*, 47:50–72, 2011.

[171] S. Vukovic. Eno and weno schemes with the exact conservation property for one-dimensional shallow water equations. *J. Comput. Phys.*, 179(2):593–621, 2002.

[172] Z.J. Wang. High-Order Spectral Volume Method for Benchmark Aeroacoustic Problems. *AIAA Paper*, 2003-0880, 2003.

[173] A. N. Williams and W. G. McDougal. Flexible floating break-water. *J. Waterw. Port Coastal Ocean Eng.*, 117(5):429–450, 1991.

[174] D. Wirasaet, E.J. Kubatko, C.E. Michoski, S. Tanaka, and J.J. Westerink. Discontinuous Galerkin methods with nodal and hybrid modal/nodal triangular, quadrilateral, and polygonal elements for nonlinear shallow water flow. *Comput. Methods Appl. Mech. Engrg.*, 270:113–149, 2014.

[175] X. Xia and Q. Liang. A new efficient implicit scheme for discretising the stiff friction terms in the shallow water equations. *Advances in water resources*, 117:87–97, 2018.

[176] Y. Xing and C.-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.

[177] Y. Xing and C.-W. Shu. A new approach of high order well-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.

[178] 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.

[179] Y. Xing, X. Zhang, and C.-W. Shu. Positivity-preserving high order well-balanced discontinuous galerkin methods for the shallow water equations. *Advances in Water Resources*, 33(12):1476 – 1493, 2010.

[180] G.Q. Yang, O.M. Faltinsen, and R. Zhao. Wash of ships in finite water depth. In *Proceedings of the FAST 2001, Southhampton, UK*, 2001.

[181] M. Yang and Z.J. Wang. A parameter-free generalized moment limiter for high-order methods on unstrucured grids. *Adv. Appl. Math. Mech.*, 4:451–480, 2009.

[182] Y.H. Yu and L. Ye. Reynolds-averaged navier stokes simulation of the heave performance of a two-body floating-point absorber wave energy system. *Computers and Fluids*, 73:104–114, 2013.

[183] 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.

[184] X. Zhang and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. A*, 467:2752–2776, 2011.

[185] 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.

[186] J.G. Zhou, D.M. Causon, C.G. Mingham, and D.M. Ingram. The surface gradient method for the treatment of source terms in the shallow-water equations. *Journal of Computational physics*, 168(1):1–25, 2001.

[187] J. Zhu, J. Qiu, C.-W. Shu, and M. Dumbser. Runge–Kutta discontinuous Galerkin method using WENO limiters II: Unstructured meshes. *J. Comput. Phys.*, 227:4330–4353, 2008.

[188] J. Zhu, X. Zhong, C.-W. Shu, and J. Qiu. Runge Kutta discontinuous Galerkin method using a new type of WENO type limiters on unstructured meshes. *J. Comp. Phys.*, 248:200–220, 2013.