Upwind Residual discretization of enhanced Boussinesq equations for wave propagation over complex bathymetries

M. Ricchiuto*,+ and A.G. Filippini*

* Inria Bordeaux - Sud-Ouest,
200 Avenue de la Vieille Tour, 33405 Talence cedex - France

+ Institut de Mathématiques de Bordeaux,
351 cours de la Libération, F 33405 Talence cedex - France

Abstract

In this paper we consider the solution of the enhanced Boussinesq equations of Madsen and Sørensen (*Coast.Eng.* 18, 1992) by means of residual based discretizations. In particular, we investigate the applicability of upwind and stabilized variants of continuous Galerkin finite element and Residual Distribution schemes for the simulation of wave propagation and transformation over complex bathymetries. These techniques have been successfully applied to the solution of the nonlinear Shallow Water equations (see e.g. Hauke *CMAME* 163, 1998 and Ricchiuto and Bollerman *J.Comput.Phys* 228, 2009). In a first step toward the construction of a hybrid model coupling the enhanced Boussinesq equations with the Shallow Water equations in breaking regions, this paper shows that equal order and even low order (second) upwind/stabilized techniques can be used to model non-hydrostatic wave propagation over complex bathymetries. This result is supported by theoretical (truncation and dispersion) error analyses, and by thorough numerical validation.

Contents

1 Introduction

The accurate simulation of nonlinear and non-hydrostatic wave propagation and transformation on complex bathymetries in the near shore region, up to the shoreline, plays a major role in coastal engineering. Numerical models for the applications involved benefit on one hand from the development of mathematical models with improved dispersion and shoaling characteristics, and, on the other, from the availability of accurate and stable discretizations of these equations.

Significant effort has been put in the last 20 years in development of systems of depth averaged equations which correctly reproduce the dispersion characteristics of wave propagation in the near shore region. Starting from the Boussinesq equations of Peregrine [56], several improved and enhanced Boussinesq models have been proposed over the years, including, among those having the largest impact in literature and the most recent ones, the enhanced equations of Madsen and Sørensen [55], the extended formulation of Nwogu [53], genuinely nonlinear Serre-Green-Naghdi equations [49], and nonlinear and non-hydrostatic higher order Shallow-Water type models [37]. These models have been obtained by retaining asymptotic behavior of the order of $\mathscr{O}(\mu^2)$, μ being the ratio of water depth to wavelength. If h_0 is the value of a reference average depth, they give a correct description of the physics for values of the wave parameter $kh_0 \approx 3-5$. More accurate models, including effects up to the $\mathscr{O}(\mu^4)$ order have been proposed *e.g.* in [35].

Concerning the numerics used to solve these equations, the literature is full of promising schemes involving finite differences, finite volumes, or finite elements approaches. The major challenges that need to be dealt with are the approximation of the complex higher order derivative terms present in all non-hydrostatic depth-averaged models, and the accuracy requirements on the schemes in terms of low dispersion error. In addition, Boussinesq models can be coupled with the nonlinear Shallow-Water (NLSW) equations to model wave breaking [15, 68, 69, 67, 47]. While the mathematical character of the Boussinesq equations is (roughly) parabolic, the NLSW system is hyperbolic. As such it requires some degree of stabilization, e.g. in the form of some type of upwinding. A model coupling the Boussinesq system and the NLSW equations requires the underlying numerics used to robustly handle both the parabolic and purely hyperbolic limits.

The presence of higher order (third) partial derivatives has made the use of finite difference approximations appealing and quite popular (see *e.g.* [12, 33, 53, 34] to cite a few). The main drawback of the finite difference approach is the need of structured spatial meshes, even for irregular domains, and poor local mesh adaptivity potential (even tough hierarchical block structured multi-level approaches do exist, see *e.g.* [13]).

Fully unstructured solvers allowing for adaptive mesh refinement have been proposed, based either on the finite volume, or on the finite element approach. To the author's knowledge, genuinely multidimensional unstructured finite volume discretizations of enhanced Boussinesq equations have been actually proposed only in [46, 10], other works proposing some form of hybridization of finite volume/finite difference schemes on structured meshes or even in one space dimension (see *e.g.* [25, 16, 68] and references therein). The results presented in [46] are particularly encouraging, and the extension of the authors' model to wave breaking applications , presented at the Modeling the Earth system conference in Boulder [45], has shown the high potential of their approach. One criticism that can be made to the finite volume framework is that going beyond third order of accuracy might be quite hard, due to the necessity of introducing higher order multidimensional reconstructions for both the unknowns, and for the velocity divergence, to allow the discretization of the dispersive terms [46]. The advantage of the finite volume framework is of course an easy application of upwinding principles to properly handle the hyperbolic limit of the NLSW equations, and the use of well-established limiting techniques to avoid oscillations near bores and hydraulic jumps.

On the other hand, the finite element approximation gives a framework to naturally intro-

duce higher order polynomial representation of the unknowns and of their derivatives, simply by handling these as auxiliary variables. The work of [30, 27] on discontinuous Galerkin approximations of enhanced Boussinesq models shows the potential in terms of accuracy of the finite element approach. Continuous Galerkin discretizations of Bossinesq models have been proposed by several authors. For example, in [21] a Taylor-Galerkin formulation for the Peregrine equations is discussed. More recently, a model based on Taylor-Galerkin time integration and the enhanced Madsen and Sørensen equations has been proposed in [54], using mixed approximation space. Standard Galerkin approximations are also discussed in [50, 72] (see also the PhD [71]). These contributions show results at least as good as those obtained by means of finite difference schemes, with the additional flexibility of a natural unstructured mesh formulation.

In this paper we want to add to this panorama an additional element, by analyzing and testing continuous finite element and residual based schemes which include some form of upwind stabilization, and which have already been shown to accurately and robustly handle the Shallow Water equations. In particular, we consider the Residual Distribution (RD) schemes, developed e.g. in [61, 60, 22, 17], and the upwind stabilized Galerkin scheme known as Streamline Upwind Petrov Galerkin scheme (SUPG) of [39, 40, 43]. These schemes have shown very high potential in handling the NLSW, both in terms of preservation of physically relevant steady equilibria (well-balancedness), and in terms of a stable approximation of moving shorelines [61, 58, 39, 17]. For purely hyperbolic problems, it is known that, compared to finite differences, finite element schemes, and generally for residual based discretizations, have improved dispersion characteristics, due to the presence of a mass matrix.

While in the hyperbolic case this might seem like a drawback, in presence of mixed space and time derivatives, as in Boussinesq models, this gives an advantage, allowing to build discretizations that, on a reduced stencil, and even for low order interpolation (piecewise linear), yield dispersion properties similar to those of higher order finite difference schemes. Our aim is to analyze both theoretically and numerically second order upwind RD and SUPG discretizations for the enhanced Boussinesq equations of [55], and to asses their applicability to wave propagation. In the one-dimensional case, and for the linearized system, the paper presents a time-continuous error analysis based on a standard truncation error study of the finite difference form of the schemes, and a dispersion error analysis. The schemes are then thoroughly tested and compared to one another on one dimensional benchmarks taken from the literature. Both the analytical and numerical results lead to the conclusion that the Petrov-Galerkin approach might be the best suited for this application. We thus consider a two dimensional extension based on Petrov-Galerkin forms which, when considering the NLSW equations, give back the standard SUPG scheme, and the successful Multidimensional Upwind Residual Distribution scheme known as LDA scheme [22]. The results on well known two-dimensional benchmarks show that : on one hand the use of these schemes for wave propagation, on meshes with typical size comparable to that used by finite difference practitioners, is indeed feasible ; on the other hand, that, compared to the standard SUPG stabilization, Multidimensional Upwinding leads to slightly less pronounced shoaling, and a lower content in higher harmonics. This work has to be understood as a first step toward the construction of a model including coupling with the NLSW equations to handle wave breaking and moving shorelines.

The structure of the paper is the following. In section §2 we recall the basic form of the

enhanced Boussinesq model of [55], and in section §3 we present the schemes analyzed in one space dimension, including their implementation. Section §4 is devoted to a time continuous error analysis and comparison with second and higher order finite difference discretizations of the linearized equations. Sections §5 discusses the issue of the initial and boundary conditions. The one dimensional benchmarking is presented in sections §6 and §7, which discuss in detail the CPU cost of the schemes considered and their comparison on several well established numerical tests. In section §8, we discuss the extension of the schemes to two space dimension using a Petrov-Galerkin approach, and introducing two different generalizations of the upwind stabilization studied in 1d. Section §9 is devoted to the benchmarking of the Petrov-Gelerkin schemes in two space dimensions. The paper is ended by a summary of the results and an outlook on ongoing work.

2 Enhanced Boussinesq equations in one dimension

With reference to the notation of figure 1, the enhanced Boussinesq equations of Madsen and Sørensen [55] can be written as

$$\begin{cases} \partial_t \eta + \partial_x q = 0 \\ \partial_t q - Bh^2 \partial_{x^2 t} q - \dfrac{1}{3} h \partial_x h \partial_{xt} q + \partial_x (uq) + gH\partial_x \eta + \\ \qquad\qquad - \beta g h^3 \partial_{x^3} \eta - 2\beta g h^2 \partial_x h \partial_{x^2} \eta = 0 \end{cases} \tag{1}$$

where $\eta(x,t)$ and $h(x)$ denote the surface elevation and the depth at still water (cf. figure 1), while $H(x,t) = \eta(x,t) + h(x)$ and $q(x,t)$ are the total depth, and the discharge $q = Hu$, u denoting the depth averaged speed, as in the NSWE system. In addition, the brief notation ∂_{x^n} will be used within this work in order to indicate the recursive application of the partial derivative with respect to x for n times. These equations provide a description of $\mathcal{O}(\epsilon, \mu^2)$ of the wave propagation physics, recalling that the nonlinearity ϵ parameter represents the ratio of wave amplitude to depth, and the dispersion μ is the ratio of water depth to wavelength.

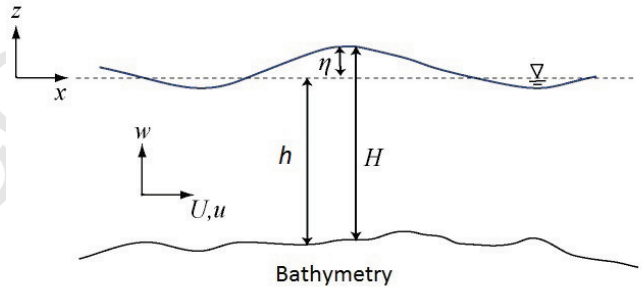

Figure 1: Sketch of the free surface flow problem, main parameters description.

This model is weakly nonlinear, preserving the same shallow water terms $\partial_x(uq)$ and $gH\partial_x\eta$ which are the only nonlinear in the system, being, thus, different from the fully nonlinear models, like the one proposed in [49]. It also contains additional dispersive terms in the

momentum equation, which are linear with respect to the unknowns η and q of the system. These are pre-multiplied by two numerical parameters B and β, whose values are obtained by optimizing the dispersion properties of the linearized model with respect to the Airy wave theory. In such a way, the two parameters assume the values $\beta = 1/15$ and $B = \beta + 1/3$ [55]. Note that the linearized Boussinesq equations of Peregrine are obtained by setting $\beta = 0$, while the NLSW equations are recovered by neglecting all the high order derivative terms.

For later use, we recall that, assuming that both η and u are very small perturbation of a still steady state, and flat bathymetry, the linearized version of system (1) reads

$$\left\{ \begin{array}{l} \partial_t \eta + h_0 \partial_x u = 0 \\ \partial_t u - B h_0^2 \partial_{x^2 t} u + g \partial_x \eta - \beta g h_0^2 \partial_{x^3} \eta = 0 \end{array} \right. \tag{2}$$

The dispersion characteristics of this linearized system are obtained by replacing η and u by a propagating Fourier mode

$$W = \left(\begin{array}{c} \eta \\ u \end{array} \right) = W_0 e^{\nu t + \mathrm{i} k x}$$

with i the imaginary unit, and with the wavenumber k related to the wavelength as $k = 2\pi/\lambda$, and with a complex amplification parameter $\nu = \xi + \mathrm{i}\omega$, ξ representing the dissipation parameter, and ω the phase. The so-called Airy theory for water wave propagation (see *e.g.* [53] and references therein) gives for these parameters the analytical values

$$\left\{ \begin{array}{l} \xi_{\mathrm{Airy}} = 0 \\ \\ \omega_{\mathrm{Airy}}^2 = C_0^2 k^2 \dfrac{\tanh(k h_0)}{k h_0} \end{array} \right. \tag{3}$$

with $C_0^2 = g h_0$, the linearized Shallow Water celerity. Substitution of the Fourier mode in (2) quickly provides the approximated value given by the Madsen and Sørensen model

$$\left\{ \begin{array}{l} \xi_{\mathrm{MS}} = 0 \\ \\ \omega_{\mathrm{MS}}^2 = C_0^2 k^2 \dfrac{1 + \beta (k h_0)^2}{1 + B (k h_0)^2} \end{array} \right. \tag{4}$$

Relation (4) is known to be a significantly improved approximation of (3) w.r.t. the relations obtained with the linearized Shallow Water equations or the Peregrine Boussinesq equations [55, 65].

3 Space discretization

3.1 Generalities and notation

Let Ω denote the spatial domain. We consider a tesselation Ω_{h} composed by a set of non-overlapping elements, the subscript h denoting the reference mesh size. The generic element K is defined by a set of nodes, *e.g.* in one space dimension $\mathrm{K} \equiv [x_j,\, x_{j+1}]$, with $\mathrm{h}_K = x_{j+1} - x_j$. Unknowns are stored at nodes as time dependent values $\{\eta_i(t)\}_{i \geq 1}$ and $\{q_i(t)\}_{i \geq 1}$. For a generic node i we will also denote by K_i the set of elements containing i as a node. As in

the standard P^1 finite element method, nodal values are interpolated by means of piecewise linear continuous shape functions $\varphi_i(x)$, the interpolated values being denoted by η_h, and q_h with (cf. figure 2)

$$\begin{aligned}\eta_{\rm h}(t,x) &= \sum_{i\geq 1}\eta_i(t)\varphi_i(x) = \sum_K\sum_{j\in K}\eta_j(t)\varphi_j(x)\\ q_{\rm h}(t,x) &= \sum_{i\geq 1} q_i(t)\varphi_i(x) = \sum_K\sum_{j\in K} q_j(t)\varphi_j(x)\end{aligned} \tag{5}$$

with the $\varphi_i(x)$ the standard continuous piecewise linear finite element basis functions assuming value 1 in node i and zero in all the other nodes. As discussed in the introduction, in this paper we focus on piecewise linear interpolation in order to show the feasibility of the use of compact low order discretizations for wave propagation. However all the developments presented, including the general form of the schemes, extend naturally to higher order polynomial approximation.

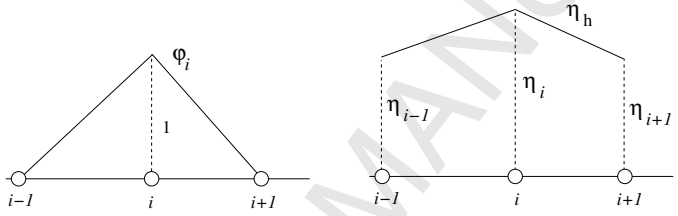

Figure 2: P^1 finite element interpolation

In the following subsections we present the schemes studied in the paper in the hypothesis that periodic boundary conditions are used. More details concerning boundary conditions and wave generation are given in sections §5 and §6.

3.2 Continuous Galerkin approximation

Following [72], we write the continuous Galerkin approximation of system (1) as, seek $\forall, i \in \Omega_h$, the solution of

$$\begin{aligned}
&\int_{\Omega_{\rm h}} \varphi_i\partial_t\eta_{\rm h} - \int_{\Omega_{\rm h}} q_{\rm h}\,\partial_x\varphi_i = 0\\
&\int_{\Omega_{\rm h}} \varphi_i\partial_t q_{\rm h} + \int_{\Omega_{\rm h}} B\partial_{xt}q_{\rm h}\partial_x(h^2\varphi_i) - \int_{\Omega_{\rm h}}\frac{1}{3}\varphi_i h\partial_x h\partial_{xt}q_{\rm h} - \int_{\Omega_{\rm h}}(uq)_{\rm h}\partial_x\varphi_i - \int_{\Omega_{\rm h}} g\frac{H_{\rm h}^2}{2}\partial_x\varphi_i\\
&\qquad - \int_{\Omega_{\rm h}}\varphi_i g H_{\rm h}\partial_x h\int_{\Omega_{\rm h}}\varphi_i\beta g h^3\partial_x w_{\rm h}^\eta - \int_{\Omega_{\rm h}}\varphi_i 2\beta g h^2\partial_x h\, w_{\rm h}^\eta = 0\\
&\int_{\Omega_{\rm h}}\varphi_i w_{\rm h}^\eta + \int_{\Omega_{\rm h}}\partial_x\eta_{\rm h}\partial_x\varphi_i = 0
\end{aligned} \tag{6}$$

where the auxiliary variable w^η is an approximation to the second order spatial derivative of the free surface level η. Note that in (6) we have separated and integrated by parts the terms corresponding to the NLSW flux terms $\partial_x(uq)$ and $g\partial_x H^2/2$. However, due to the piecewise continuous nature of the approximation, and to the assumed periodic boundary conditions, the terms corresponding to these fluxes can be re-integrated by parts giving

$$-\int_{\Omega_\mathrm{h}} (uq)_\mathrm{h}\partial_x\varphi_i = \sum_{\mathrm{K}\in\Omega_\mathrm{h}} \int_{\mathrm{K}} \varphi_i \partial_x (uq)_\mathrm{h} - \sum_{\mathrm{K}\in\Omega_\mathrm{h}} [\varphi_i (uq)_\mathrm{h}]_\mathrm{K} = \sum_{\mathrm{K}\in\Omega_\mathrm{h}} \int_{\mathrm{K}} \varphi_i \partial_x (uq)_\mathrm{h} = \int_{\Omega_\mathrm{h}} \varphi_i \partial_x (uq)_\mathrm{h}$$

the sum of jump terms on element boundaries $[\varphi_i (uq)_\mathrm{h}]_\mathrm{K}$ being zero due to the continuity of φ_i and $(uq)_\mathrm{h}$, and to the compact support of the basis functions. Proceeding similarly for the $g\partial_x H^2/2$ term, we can recast (6) as

$$\begin{aligned}
&\int_{\Omega_\mathrm{h}} \varphi_i \partial_t \eta_\mathrm{h} - \int_{\Omega_\mathrm{h}} q_\mathrm{h}\, \partial_x \varphi_i = 0 \\
&\int_{\Omega_\mathrm{h}} \varphi_i \partial_t q_\mathrm{h} + \int_{\Omega_\mathrm{h}} B\partial_{xt} q_\mathrm{h} \partial_x (h^2 \varphi_i) \\
&\qquad + \int_{\Omega_\mathrm{h}} \varphi_i \left[-\frac{1}{3} h\partial_x h \partial_{xt} q_\mathrm{h} + \partial_x (uq)_\mathrm{h} + gH_\mathrm{h}\partial_x \eta_\mathrm{h} - \beta g h^3 \partial_x w^\eta_\mathrm{h} - 2\beta g h^2 \partial_x h\, w^\eta_\mathrm{h} \right] = 0 \\
&\int_{\Omega_\mathrm{h}} \varphi_i w^\eta_\mathrm{h} + \int_{\Omega_\mathrm{h}} \partial_x \eta_\mathrm{h} \partial_x \varphi_i = 0
\end{aligned} \tag{7}$$

The actual discretization is obtained by evaluating all the integrals by numerical quadrature over each element $\mathrm{K} \in \mathrm{K}_i$, with the assumption of piecewise linear variation of all the quantities involved. Note that the introduction of the auxiliary variable w^η is made necessary by the presence of the higher (third) order derivatives of the free surface level.

3.3 Centered Residual Distribution

Even though sharing a similar cell-vertex philosophy, the Residual Distribution (RD) discretization is obtained with a different strategy. Discrete equations are obtained by first computing integrated values of the elemental residuals, denoted by Φ^K, and then by distributing fractions of these elemental residuals to the nodes forming the element (cf. figure 3 and refer to [22, 61, 60] for details).

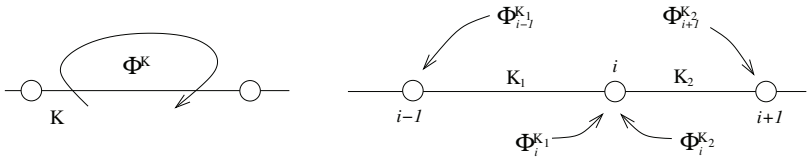

Figure 3: One dimensional Residual Distribution

In particular, with the same underlying continuous spatial approximation used for the Galerkin scheme, we define the element residual $\Phi^K = \left[\Phi^{\mathrm{K}}_\eta \quad \Phi^{\mathrm{K}}_q\right]^T$ as

$$\begin{aligned}
\Phi^{\mathrm{K}}_\eta &= \int_{\mathrm{K}} \left(\partial_t \eta_{\mathrm{h}} + \partial_x q_{\mathrm{h}}\right) \\
\Phi^{\mathrm{K}}_q &= \int_{\mathrm{K}} \Big(\partial_t q_{\mathrm{h}} - Bh^2 \partial_{x^2 t} q_{\mathrm{h}} - \frac{1}{3} h \partial_x h \partial_{xt} q_{\mathrm{h}} + \partial_x (uq)_{\mathrm{h}} \\
&\qquad + gH_{\mathrm{h}} \partial_x \eta_{\mathrm{h}} - \beta g h^3 \partial_{x^3} \eta_{\mathrm{h}} - 2\beta g h^2 \partial_x h \partial_{x^2} \eta_{\mathrm{h}} \Big)
\end{aligned} \tag{8}$$

Discrete equations for the nodal values are obtained by numerically integrating the relations defining the elemental residual (8), and then by distributing the resulting quantity to the nodes of element K by means of a distribution matrix β^{K}_i, so that nodal discrete equations are obtained as

$$\sum_{\mathrm{K} \in \mathrm{K}_i} \beta^{\mathrm{K}}_i \Phi^{\mathrm{K}} = 0 \tag{9}$$

The centered RD scheme (cRD) is obtained (in one space dimension) with the simple choice $\beta^{\mathrm{K}}_i = \mathrm{I}_2/2 \; \forall \, i$ and $\forall \, \mathrm{K}$, with I_2 denoting the 2×2 identity matrix.

As in the case of the Galerkin scheme, the presence of high order derivative terms requires the definition of nodal values of the derivatives of the variables, which are otherwise only defined at the elemental level. As in finite element schemes, this is achieved in the RD community by treating these nodal derivatives as a set of auxiliary variables, for which reconstruction strategies or ad-hoc discrete equations are developed (see *e.g.* [2, 5, 57, 52] and references therein). Following the work done in [2, 5, 57], and using the auxiliary variable already introduced for the Galerkin scheme (cf. equation (7)), we have in practice computed the second in (8) as

$$\begin{aligned}
\Phi^{\mathrm{K}}_q = \int_{\mathrm{K}} \Big(&\partial_t q_{\mathrm{h}} - Bh^2 \partial_{xt} w^q_{\mathrm{h}} - \frac{1}{3} h \partial_x h \partial_{xt} q_{\mathrm{h}} + \partial_x (uq)_{\mathrm{h}} \\
&+ gH_{\mathrm{h}} \partial_x \eta_{\mathrm{h}} - \beta g h^3 \partial_x w^\eta_{\mathrm{h}} - 2\beta g h^2 \partial_x h \, w^\eta_{\mathrm{h}} \Big)
\end{aligned} \tag{10}$$

with auxiliary variables given by

$$\begin{aligned}
\int_{\Omega_{\mathrm{h}}} \varphi_i w^q_{\mathrm{h}} + \int_{\Omega_{\mathrm{h}}} q_{\mathrm{h}} \partial_x \varphi_i = 0 \\
\int_{\Omega_{\mathrm{h}}} \varphi_i w^\eta_{\mathrm{h}} + \int_{\Omega_{\mathrm{h}}} \partial_x \eta_{\mathrm{h}} \partial_x \varphi_i = 0
\end{aligned} \tag{11}$$

where now the auxiliary variable w^q is an approximation of the nodal derivative of q. More details on the treatment of these variables will be given in section §3.5.

3.4 Upwind discretizations

Schemes (7) and (8)-(10) are centered approximations of the equations and are not well suited for the discretization of the Shallow Water limit for which some form of upwinding is necessary to stabilize the system, or to provide positivity corrections in correspondence of moving

shorelines and discontinuities [60, 61]. In view of the coupling of the Boussinesq equations with the NLSW system to handle wave breaking, we will analyze here two upwind schemes obtained by adding to schemes (7) and(8)-(10) an upwind bias *based on the characteristic decomposition of the NLSW equations.*

We start by rewriting scheme (7) with the short-notation

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) = 0 \tag{12}$$

with $\mathcal{R}_i^{\mathrm{cG}}(\eta, q)$ the array whose components are the left hand sides of the first two equations in (7). In order to construct an upwinding operator, we consider now the quasi-linear form of the NLSW equations which can be recast as

$$\partial_t \begin{pmatrix} \eta \\ q \end{pmatrix} + A\partial_x \begin{pmatrix} \eta \\ q \end{pmatrix} = 0\,, \quad A = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix}$$

where $c^2 = gH$ is the Shallow Water celerity. We recall that matrix A admits a full set of real linearly independent eigenvectors, associated to the two eigenvalues $u \pm c$. Following the SUPG stabilization technique [40, 43, 61, 1], we define the *stabilized* variant of the continuous Galerkin (7) as

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_h} \int_{\mathrm{K}} A\partial_x\varphi_i\, \tau_{\mathrm{K}}\, r_{\mathrm{h}}^{\mathrm{MS}} = 0 \tag{13}$$

where the matrix τ_{K} is the so-called SUPG stabilization parameter, and having denoted by $r_{\mathrm{h}}^{\mathrm{MS}}$ the local residual value of the Madsen and Sørensen equations :

$$r_{\mathrm{h}}^{\mathrm{MS}} = \begin{pmatrix} \partial_t\eta_{\mathrm{h}} + \partial_x q_{\mathrm{h}} \\ \partial_t q_{\mathrm{h}} - Bh^2\partial_{x^2t} q_{\mathrm{h}} - \dfrac{1}{3}h\partial_x h\partial_{xt} q_{\mathrm{h}} + \partial_x(uq)_{\mathrm{h}} + gH_{\mathrm{h}}\partial_x\eta_{\mathrm{h}} - \beta g h^3\partial_{x^3}\eta_{\mathrm{h}} - 2\beta g h^2\partial_x h\partial_{x^2}\eta_{\mathrm{h}} \end{pmatrix}$$

As done previously, to evaluate the integral of the SUPG stabilization, we introduce the auxiliary variables $w^\eta \approx \partial_{x^2}\eta$ and $w^q \approx \partial_x q$, and rewrite $r_{\mathrm{h}}^{\mathrm{MS}}$ as

$$r_{\mathrm{h}}^{\mathrm{MS}} = \begin{pmatrix} \partial_t\eta_{\mathrm{h}} + \partial_x q_{\mathrm{h}} \\ \partial_t q_{\mathrm{h}} - Bh^2\partial_{xt} w_{\mathrm{h}}^q - \dfrac{1}{3}h\partial_x h\partial_{xt} q_{\mathrm{h}} + \partial_x(uq)_{\mathrm{h}} + gH_{\mathrm{h}}\partial_x\eta_{\mathrm{h}} \\ \qquad\qquad - \beta g h^3\partial_x w_{\mathrm{h}}^\eta - 2\beta g h^2\partial_x h\, w_{\mathrm{h}}^\eta \end{pmatrix} \tag{14}$$

where now all the quantities involved have a piecewise linear variation. In addition, for a P^1 approximation, the basis functions derivative $\partial_x\varphi_i$ are constant within each element, thus if a one point linearization of the NLSW Jacobian is used in evaluating the stabilization integral, we are left with

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_h} A^{\mathrm{K}}\partial_x\varphi_i^{\mathrm{K}}\tau_{\mathrm{K}} \int_{\mathrm{K}} r_{\mathrm{h}}^{\mathrm{MS}} = 0 \tag{15}$$

with A^{K} denoting hte local linearization of the NLSW flux Jacobian. Comparing (15) with (14) and (10), we see that, *in the P^1 case*, scheme (15) can be recast as

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_h} A^{\mathrm{K}}\partial_x\varphi_i^{\mathrm{K}}\tau_{\mathrm{K}}\Phi^{\mathrm{K}} = 0 \tag{16}$$

At last, we employ here the definition of the SUPG stabilization parameter allowing to recover the upwind discretization of a first order hyperbolic operator (see *e.g.* [11, 23] and references therein), namely

$$\tau_{\mathrm{K}} = \frac{1}{\sum\limits_{j\in\mathrm{K}} |\partial_x \varphi_j^K|} |A^{\mathrm{K}}|^{-1} \tag{17}$$

with the absolute value $|A^{\mathrm{K}}|$ computed by means of standard eigenvalue decomposition. In one space dimension this leads to the SUPG scheme

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_h} \mathrm{sign}(\partial_x \varphi_i^{\mathrm{K}}) \frac{\mathrm{sign}(A^{\mathrm{K}})}{2} \Phi^{\mathrm{K}} = 0 \tag{18}$$

where again the sign matrix $\mathrm{sign}(A^{\mathrm{K}})$ is computed by means of standard eigenvalue decomposition. In one space dimension, the stabilization operator for node i only acts in the cells $\mathrm{K}^{i-1/2} \equiv [x_{i-1},\, x_i]$ and $\mathrm{K}^{i+1/2} \equiv [x_i,\, x_{i+1}]$ providing the additional terms $\pm\mathrm{sign}(A^{\mathrm{K}})\Phi^{\mathrm{K}^{i\mp1/2}}/2$.

In the RD case, a more intuitive procedure is used to project the elemental residual Φ^{K} onto the NLSW characteristics, and split each term according to the sign of the corresponding eigenvalue. For a node i this boils down to the following discrete equations :

$$\frac{\mathrm{I}_2 + \mathrm{sign}(A^{\mathrm{K}^{i-1/2}})}{2} \Phi^{\mathrm{K}^{i-1/2}} + \frac{\mathrm{I}_2 - \mathrm{sign}(A^{\mathrm{K}^{i+1/2}})}{2} \Phi^{\mathrm{K}^{i+1/2}} = 0 \tag{19}$$

which can be readily recast as

$$\mathcal{R}_i^{\mathrm{cRD}}(\eta_{\mathrm{h}}, q_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_h} \mathrm{sign}(\partial_x \varphi_i^{\mathrm{K}}) \frac{\mathrm{sign}(A^{\mathrm{K}})}{2} \Phi^{\mathrm{K}} = 0 \tag{20}$$

having denoted by $\mathcal{R}_i^{\mathrm{cRD}}$ the algebraic component associated to the centered RD scheme ($\beta_i^{\mathrm{K}} = \mathrm{I}_2/2$ in equation (9) plus (8) and (10)).

3.5 Auxiliary variables, conservation and well balancedness

As defined by equations (7) and (11), the L^2 projections defining the nodal values of the auxiliary variables w_{h}^q and w_{h}^η require the solution of a linear system whose matrix is the Galerkin mass matrix

$$m_{ij}^{\mathrm{cG}} = \int_{\Omega_{\mathrm{h}}} \varphi_i \varphi_j$$

This matrix being symmetric, positive definite, and constant, this system can be solved very efficiently, and its LU decomposition can be actually stored, reducing the reconstruction of the nodal values $\{w_i^q\}_{i\geq1}$ and $\{w_i^\eta\}_{i\geq1}$ to a matrix-vector product.

However, as remarked in [71], for P^1 elements, in practice it makes no difference whether the left hand side of the projection operator is evaluated exactly, or if a mass lumping procedure is employed, yielding, in one space dimension

$$(x_{i+1} - x_{i-1}) w_i^q = q_{i+1} - q_{i-1} \tag{21}$$

and

$$\frac{x_{i+1}-x_{i-1}}{2}w_i^{\eta}=\frac{\eta_{i+1}-\eta_i}{x_{i+1}-x_i}-\frac{\eta_i-\eta_{i-1}}{x_i-x_{i-1}} \tag{22}$$

Both formulas reduce to standard second order central differencing on an equally spaced mesh. Note that, while giving a simpler and more efficient formulation, and allowing for a straightforward inclusion of these terms in the evaluation of a numerical Jacobian if using implicit time integration with Newton iterations, this approach limits the type of finite elements that can be used to those for which mass lumping works properly. This rules out a certain number of elements, including standard P^2 Lagrange elements (in 2D), or low-order spectral elements (see e.g. [29]). It however allows the use of other type of elements, such as e.g. those based on Bezier polynomials [63, 6]. In all the other cases, the full Galerkin matrix must be retained in the L^2 projection.

In all the following analysis, and in all the numerical results, we have made explicit use of the simpler finite difference reconstruction formulae (21) and (22).

Additional precisions are in order concerning the form of the equations in (7), (8)-(9), and (13)-(14). Firs of all, we remark that for exact integration w.r.t. the assumed polynomial variation of H_h and h_h, and due to the use of the same polynomial expansion for H_h and h_h, the direct use of the free surface level in the hydrostatic terms in the q equations, or the separate use of the conservative form of the flux $\partial_x(gH_\mathrm{h}^2/2)$ plus the bathymetry term $-gH_\mathrm{h}\partial_x h_\mathrm{h}$ is absolutely equivalent, as already shown in section §3.2 (cf. also [61] for the RD case). This actually means that our discretization of the Shallow Water part of the model is perfectly conservative. As a direct consequence of being able to write the hydrostatic term in function of the gradient of the free surface level, we obtain that in correspondence of the so-called lake at rest state $\eta=0$ and $q=0$, the schemes proposed reduce to a system of EDOs of the type

$$M\frac{dW_\mathrm{h}}{dt}=0$$

with W_h is the array containing all the nodal values of η_h and q_h (cf. next section). In other words, our approach preserves indefinitely the lake at rest state (cf. [61] and reference therein for more).

3.6 Time integration and implementation in 1d

Time integration has been performed with the Crank-Nicholson (CN) scheme for all the benchmarks shown in the paper. In particular, the discretization steps described in the previous sections have been applied to the semi-discrete version of (1) that reads

$$\left\{\begin{array}{l}\Delta^{n+1}\eta+\Delta t\partial_x q^{n+1/2}=0\\ \Delta^{n+1}q-Bh^2\partial_{x^2}(\Delta^{n+1}q)-\dfrac{1}{3}h\partial_x h\partial_x(\Delta^{n+1}q)+\Delta t\partial_x(uq)^{n+1/2}\\ \quad+\Delta tgH^{n+1/2}\partial_x\eta^{n+1/2}-\Delta t\beta gh^3\partial_{x^3}\eta^{n+1/2}-2\Delta t\beta gh^2\partial_x h\partial_{x^2}\eta^{n+1/2}=0\end{array}\right. \tag{23}$$

where $\Delta^{n+1}(\cdot)=(\cdot)^{n+1}-(\cdot)^n$, and $(\cdot)^{n+1/2}=\big((\cdot)^{n+1}+(\cdot)^n\big)/2$. The choice of the CN integrator is related to its simplicity, and stability. Its non-dissipative nature allows to freely set the time step, the (neutral) linear stability of the schemes in being guaranteed by the results of

the dispersion analysis of section §4.3. Clearly, the CN scheme introduces a degree of artificial dispersion which indeed will influence the results. As a consequence, whenever a quantitative measure of the accuracy was needed we have properly scaled the time step to make sure that the time integration error should be below the spatial truncation error. Note however, that the choice of a second order time integration should not, a priori, provide worse results than schemes based on a splitting approach, already proposed in published literature [16]. The interaction of spatial and temporal discretization will have to be studied, but it has been left out of the present paper, which focuses on the influence of residual based upwind dissipation in the discretization of enhanced Boussinesq equations.

When applied to (23), all the schemes studied in this paper reduce to a *nonlinear algebraic system* for the unknowns η_{h}^{n+1} and q_{h}^{n+1} which can be generally written as

$$F(W_{\mathrm{h}}|\eta_{\mathrm{h}}^{n},\, q_{\mathrm{h}}^{n}) = 0 \tag{24}$$

where W_{h} is the array containing all the nodal values of η_{h}^{n+1} and q_{h}^{n+1}, and with the dependence of the system on the variables at the known time step kept explicitly. System (24) has to be solved by some iterative method. In one space dimension, all the schemes have been coded with Matlab®, and share the same computational skeleton. In particular, the algebraic equations are solved with a Newton loop which can be summarized as follows

1. Set $W_0 = (\eta_0^{n+1},\, q_0^{n+1})^T = (\eta^n\, q^n)^T$
2. Evaluate the *frozen* Jacobian matrix

$$\mathcal{M} = \frac{\partial F}{\partial W_{\mathrm{h}}}(W_{\mathrm{h}} = W_0|\eta_{\mathrm{h}}^{n},\, q_{\mathrm{h}}^{n})$$

3. for $\mathrm{k} = 1,\, \mathrm{k}_{\max}$ do
   (a) Evaluate $F(W_{k-1}|\eta_{\mathrm{h}}^{n},\, q_{\mathrm{h}}^{n})$
   (b) If $\|F\| \leq \epsilon$ set $\mathrm{k} = \mathrm{k}_{\max}$ and exit, else evaluate $W_k = W_{k-1} - \mathcal{M}^{-1}F(W_{k-1}|\eta_{\mathrm{h}}^{n},\, q_{\mathrm{h}}^{n})$
4. Set $W_{\mathrm{h}} = W_{\mathrm{k}_{\max}}$

In all the benchmarks presented in the following sections, step 2. is the only one in which the Jacobian $\mathcal{M}$ of the nonlinear equations is assembled. The matrix $\mathcal{M}$ is stored using Matlab®'s built-in sparse format, and the solve at steb 3.(b) is performed using Gaussian elimination (Matlab®'s built-in "mldivide" operator). Note that the structure of $\mathcal{M}$ is quite different for the schemes analyzed so far, as it an be easily deduced from the finite difference equations reported in appendix A. The SUPG and URD schemes are those with the most important fill-in, the Jacobian having a full penta-diagonal block structure with full blocks coupling all the equations for η and q. The other schemes also have a penta-diagonal structure, but they have a much sparser fill-in pattern. In particular, the FD4 scheme has no off diagonal entries of the type $\eta - \eta$ coupling nodal values of the free surface, while the FD2 and cG scheme fill in is almost tri-diagonal, with the only exception of the penta-diagonal $q - \eta$ coupling terms in the q equations. The cG scheme also has $\eta - \eta$ terms in the first equation, arising from the mass matrix, absent in the FD schemes.

These differences influence the CPU cost of the schemes. A quantitative analysis is discussed later in the paper.

4 Time continuous error analysis

In this paragraph we present a truncation and dispersion error analysis of the schemes introduced in section §3. While it is natural to consider the linearized system (2) for the dispersion analysis, for simplicity we will consider equations (2) for the truncation error as well. The analysis is time-continuous, meaning that we do not take into account time integration, thus allowing to better underline differences between the space discretizations considered. In particular, we will study six different schemes : the four described in section §3, plus the second order centered finite difference scheme, denoted by FD2, and the centered finite difference schemes designed by Wei and Kirby [34] for Nwogu's extended Boussinesq equations, denoted by FDWK. This last scheme is obtained simply using centered finite difference approximation of the fourth order accuracy in first order terms of (2), while higher order dispersive terms are discretized by means of centered finite difference approximation of the second order. This approach is very common among the scientific community and is used as a reference. Note that the standard centered fourth order finite difference schemes, initially included in the comparison, yields dispersion errors almost identical to those of the FDWK scheme, and for this reason it is not shown in the results of the analysis.

Following the work done in [71, 72, 29], we consider the equivalent finite difference form of all the schemes, and analyze it w.r.t. its truncation and dispersion error. Clearly, this approach is not the standard one for finite element and Residual Distribution (cf. [18, 38, 4, 22]), and it is only suited in one space dimension, and its results cannot be generalized to the multidimensional case, especially on irregular meshes. It does however provide a very good indication of the potential of the schemes, especially when regular, or even structured, grids can be used in two dimensions.

Note that the truncation error of the C^0 continuous Galerkin scheme was already considered in [71, 72, 29]. However, in the references the authors limit themselves to use the form of the truncation error to argue what the dispersion properties of the scheme will be. Here, we explicitly compute the dispersion coefficient of the schemes, and compare the dispersion error of different schemes as a function of the number of points per wavelength.

4.1 Equivalent finite difference equations

The expressions obtained when discretizing (2) are quite long and reporting them in the body of the paper might lead to useless excessive length. The full discrete equations are thus reported in appendix A (equations (50), (51), (52), (53), (54), and (55) for the FD2, FDWK, cG, cRD, SUPG, and URD schemes respectively). We limit ourselves to a few preliminary observations. The first is that, by comparing (50) and (52), we see that the only difference between the FD2 and cG schemes is the treatment of the first order time derivatives $\partial_t \eta$ and $\partial_t u$. The FD2 scheme provides a pointwise approximations

$$\frac{d\eta_i}{dt}, \quad \frac{du_i}{dt}$$

while the cG scheme yielding a coupling of neighboring nodes via the mass matrix :

$$\frac{1}{6}\frac{d\eta_{i-1}}{dt} + \frac{2}{3}\frac{d\eta_i}{dt} + \frac{1}{6}\frac{d\eta_{i+1}}{dt}, \quad \frac{1}{6}\frac{du_{i-1}}{dt} + \frac{2}{3}\frac{du_i}{dt} + \frac{1}{6}\frac{du_{i+1}}{dt}$$

The two schemes are otherwise identical.

The cRD scheme (53) differs from both FD2 and cG not only due to its mass matrix, but also for the approximation of the second order derivative in the mixed term $\partial_{x^2t}u$ which, due to the introduction of the auxiliary variable w^u, is based on a larger stencil :

$$\partial_{xx}(\cdot) \approx \frac{(\cdot)_{i+2} - 2(\cdot)_i + (\cdot)_{i-2}}{2\Delta x^2}$$

Otherwise the three schemes provide the same approximation of the third order term, and have the same overall stencil.

Lastly, looking at finite difference form of the SUPG and URD schemes, equations (54) and (55) respectively, we can easily identify the terms associated to the streamline upwind integral :

$$\eta_i \text{ equation : } -\frac{C_0}{2g}\left\{\frac{\Delta x}{2}\left(\frac{du_{i+1}}{dt} - \frac{du_{i-1}}{dt}\right) - \frac{Bh_0^2}{\Delta x}\left(\frac{du_{i+2}}{dt} - 2\frac{du_{i+1}}{dt} + 2\frac{du_{i-1}}{dt} - \frac{du_{i-2}}{dt}\right)\right.$$
$$\left. \underline{-g(\eta_{i+1} - 2\eta_i + \eta_{i-1})} - \underline{\underline{\frac{\beta g h_0^2}{\Delta x^2}(\eta_{i+2} - 4\eta_{i+1} + 6\eta_i - 4\eta_{i-1} + \eta_{i-2})}}\right\}$$

$$u_i \text{ equation : } -\frac{g}{2C_0}\left\{\frac{\Delta x}{2}\left(\frac{d\eta_{i+1}}{dt} - \frac{d\eta_{i-1}}{dt}\right) + \underline{h_0(u_{i+1} - 2u_i + u_{i-1})}\right\}$$

Dissipative terms can be clearly identified. For example, the underlined expressions clearly represent approximations of second order derivatives in space, while the doubly underlined terms are nothing else than a fourth order dissipation. The remaining difference expressions, which do not have any apparent property, arise from the coupling of all the terms of the equation (and of the two equations) introduced by the residual based upwinding.

4.2 Truncation error

After lengthy calculations, the leading order terms of the truncation errors of the finite difference form of the schemes reported in appendix A (cf. equations (50), (51), (52), (53), (54), and (55)) can be shown to be the following :

FD2 scheme.

$$\begin{aligned} \mathrm{TE}^{\eta}_{\mathrm{FD2}} &= \frac{h_0\Delta x^2}{6}\partial_{x^3}u_i + \mathscr{O}(\Delta x^4) \\ \mathrm{TE}^{u}_{\mathrm{FD2}} &= \frac{\Delta x^2}{6}\partial_{x^2}\left(-\frac{Bh_0^2}{2}\partial_{x^2t}u_i + g\partial_x\eta_i - \frac{3}{2}\beta g h_0^2\partial_{x^3}\eta_i\right) + \mathscr{O}(\Delta x^4) \end{aligned} \tag{25}$$

FDWK scheme.

$$\begin{aligned} \mathrm{TE}^{\eta}_{\mathrm{FDWK}} &= \frac{h_0\Delta x^4}{30}\partial_{x^5}u_i + \mathscr{O}(\Delta x^6) \\ \mathrm{TE}^{u}_{\mathrm{FDWK}} &= \frac{\Delta x^2}{4}\partial_{x^4}\left(\frac{1}{3}Bh_0^2\partial_t u_i + \beta g h_0^2\partial_x\eta_i\right) + \mathscr{O}(\Delta x^4) \end{aligned} \tag{26}$$

cG scheme.

$$\begin{aligned} \mathrm{TE}^{\eta}_{\mathrm{cG}} &= \frac{\Delta x^4}{24}\partial_{x^4}\left(\frac{1}{3}\partial_t\eta_i + \frac{h_0}{5}\partial_x u_i\right) + \mathscr{O}(\Delta x^6) \\ \mathrm{TE}^{u}_{\mathrm{cG}} &= \frac{\Delta x^2}{12}\partial_{x^4}\left(Bh_0^2\partial_t u_i - \beta g h_0^2\partial_x\eta_i\right) + \mathscr{O}(\Delta x^4) \end{aligned} \tag{27}$$

cRD scheme.

$$\begin{aligned}\mathrm{TE}^{\eta}_{\mathrm{cRD}} &= \frac{\Delta x^2}{2}\partial_{x^2}\left(\frac{1}{2}\partial_t\eta_i + \frac{h_0}{3}\partial_x u_i\right) + \mathscr{O}(\Delta x^4)\\ \mathrm{TE}^{u}_{\mathrm{cRD}} &= \Delta x^2\partial_{x^2}\left(\frac{1}{4}\partial_t u_i - \frac{1}{3}Bh_0^2\partial_{x^2t}u_i + \frac{1}{6}g\partial_x\eta_i - \frac{1}{4}\beta g h_0^2\partial_{x^3}\eta_i\right) + \mathscr{O}(\Delta x^4)\end{aligned} \tag{28}$$

SUPG scheme.

$$\begin{aligned}\mathrm{TE}^{\eta}_{\mathrm{SUPG}} &= \frac{C_0\Delta x^3}{2g}\partial_{x^3}\left(\frac{1}{3}\partial_t u_i - \frac{1}{2}Bh_0^2\partial_{x^2t}u_i + \frac{1}{6}g\partial_x\eta_i - \frac{1}{3}\beta g h_0^2\partial_{x^3}\eta_i\right) + \mathscr{O}(\Delta x^4)\\ \mathrm{TE}^{u}_{\mathrm{SUPG}} &= \frac{\Delta x^2}{12}\partial_{x^4}\left(Bh_0^2\partial_t u_i - \beta g h_0^2\partial_x\eta_i\right) + \mathscr{O}(\Delta x^3)\end{aligned} \tag{29}$$

URD scheme.

$$\begin{aligned}\mathrm{TE}^{\eta}_{\mathrm{URD}} &= \frac{\Delta x^2}{2}\partial_{x^2}\left(\frac{1}{2}\partial_t\eta_i + \frac{h_0}{3}\partial_x u_i\right) + \mathscr{O}(\Delta x^3)\\ \mathrm{TE}^{u}_{\mathrm{URD}} &= \Delta x^2\partial_{x^2}\left(\frac{1}{4}\partial_t u_i - \frac{1}{3}Bh_0^2\partial_{x^2t}u_i + \frac{1}{6}g\partial_x\eta_i - \frac{1}{4}\beta g h_0^2\partial_{x^3}\eta_i\right) + \mathscr{O}(\Delta x^3)\end{aligned} \tag{30}$$

Comparing the above expression we can make the following remarks. The FDWK, cG, and SUPG schemes provide higher consistency w.r.t. the η equation. In particular, both FDWK and cG are consistent up to $\mathcal{O}(\Delta x^4)$, while SUPG provides an $\mathcal{O}(\Delta x^3)$ approximation of the first equation. This is well known for the FDWK and cG scheme, albeit for different equations, and it is an interesting result for the SUPG. In particular, looking at the form of the truncation error, we can immediately guess what the dispersion error of the schemes might be. The FD2, cRD and URD schemes all contain in the leading $\mathcal{O}(\Delta x^2)$ truncation error terms third order derivatives of the same form of those contained in the equations. These terms will affect the dispersion relation significantly. As a side remark, note that the leading term of the URD scheme is exactly equal to the one of the cRD one, the effects of the streamline dissipation being of order $\mathcal{O}(\Delta x^3)$, as evident by the error of the SUPG scheme. On the contrary, also for the enhanced equations of [55] the FDWK and cG schemes show a truncation error containing only fifth order derivatives, which allows to guess a very good discrete dispersion relation. Similarly, the TE of the SUPG scheme does not contain any third order derivative similar to those present in the equations, but only fifth and fourth order terms related to higher order dispersion and dissipation. Once more this allows to anticipate a surprisingly good dispersion relation. These qualitative observations will be quantitatively verified by means of the explicit study of the dispersion error of the schemes, which is the subject of the following section.

4.3 Dispersion error

To obtain the discrete dispersion relations of the schemes, we proceed as follows. First, we replace the nodal values of η and u in each of the finite difference equations reported in the appendix (cf. equations (50), (51), (52), (53), (54), and (55) for the FD2, FDWK, cG, cRD, SUPG, and URD schemes respectively) by a propagating Fourier mode :

$$W_j = (\eta_j(t), u_j(t))^T = W_0 e^{\nu_{\mathrm{h}}t + \mathrm{i}kx_j}$$

with i the imaginary unit and k the wavenumber, and where $\nu_h = \xi_h + i\omega_h$ is the dispersion coefficient, with real part ξ_h representing the amplification rate, and imaginary part ω_h being the phase speed. The algebraic expressions obtained can be easily recast in terms of the nodal value W_j using relations of the type $W_{j+1} = e^{ik\Delta x}$, and $dW_j/dt = \nu W_j$. The resulting equation constitutes a complex eigenvalue problem whose solution is the dispersion factor ν_h of the scheme. The formulas obtained , reported in appendix A, are quite complex, and involve combinations of functions of the parameters $\mu = kh_0$ and $\mu_{\Delta x} = k\Delta x$.

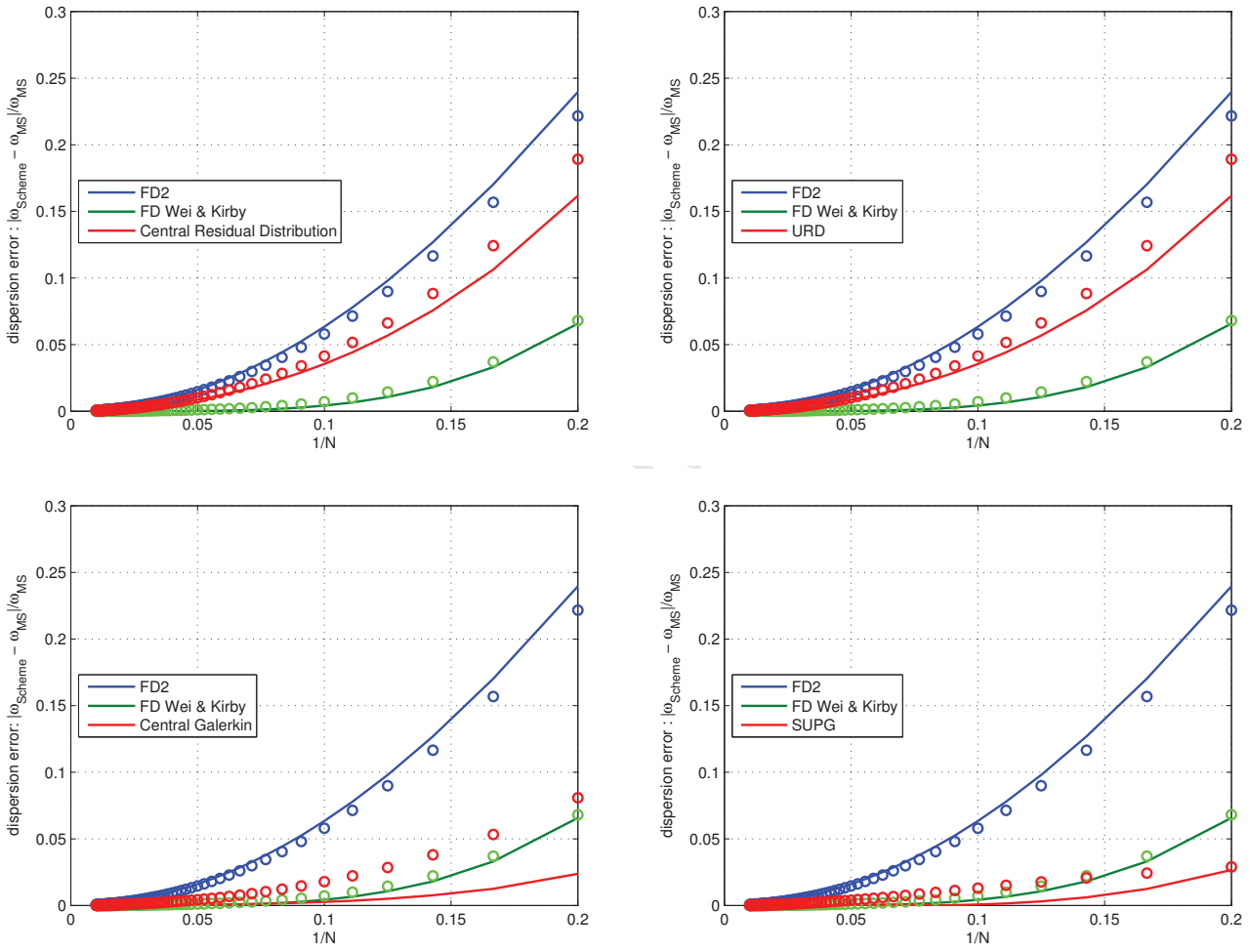

Figure 4: Dispersion error of the schemes as a function of the number of nodes per wavelength N, for $kh_0 = 0.5$ (solid lines) and $kh_0 = 2.6$ (circles). Top : cRD (left) and URD (right). Bottom : cG (left) and SUPG (right)

Some remarks can be however made. The first is that, as it might have been expected, we obtain $\xi_h = 0$ for all the centered schemes. What might appear as a surprising result is that we also have $\xi_{URD} = 0$ for the URD scheme. This result is however consistent with what can be obtained quite easily for the scalar advection equation

$$\partial_t u + a\partial_x u = 0, \quad a > 0$$

For this simple problem indeed one easily checks that the URD scheme has no amplification or dumping. A heuristic explanation of this fact is that in this case the URD scheme gives the nodal semi-discrete equation

$$\frac{\Delta x}{2}(\frac{du_{i-1}}{dt} + \frac{du_i}{dt}) + a(u_i - u_{i-1}) = 0$$

which actually is a centered approximation w.r.t. the $i - 1/2$ average cell value $(u_i + u_{i-1})/2$. For the enhanced Boussinesq equations analyzed here things are of course more complicated, however the behavior observed is the same.

Lastly, the SUPG provides the expected result of having a non-zero damping factor. In particular, we obtain $\xi_{\text{SUPG}} \leq 0$, as it should be expected from the truncation error of the scheme, and from the form of the fourth order dissipation terms in the finite difference equation (cf. sections §4.1 and §4.2).

Concerning the dispersion errors, the formulas are harder to interpret, so we have chosen to present the results in the form of comparison plots. The results are summarized in figure 4. In the plots, on the x axis we have the inverse of the number of nodes per wavelength, and on the y axis the relative error $|\omega_{\text{h}} - \omega_{MS}|/\omega_{MS}$ (cf. equation (4)) obtained for $kh_0 = 0.5$ and $kh_0 = 2.6$. In all the figures we have reported for comparison the ratios $|\omega_{FD2} - \omega_{MS}|/\omega_{MS}$ and $|\omega_{FDWK} - \omega_{MS}|/\omega_{MS}$. We can see from the plots that the RD schemes, while providing better dispersion relations than the FD2 schemes provide dispersion errors significantly above those of the FDWK scheme. On the other hand, both the cG and SUPG provide dispersion errors comparable if not lower than those of the FDWK method. In particular, both schemes provide dispersion errors lower than those of the FDWK scheme for longer waves ($kh_0 = 0.5$). For shorter waves ($kh_0 = 2.6$) the FDWK has better dispersion propertied than the Galerkin scheme, however, the SUPG scheme gives better accuracy if more than 7 points per wavelength are employed. This is a positively surprising result showing very good potential for the SUPG scheme.

5 Boundary conditions and wave generation

Before presenting the numerical validation of the schemes, we discuss some important issues related to the solution of (1), namely the techniques used to impose the boundary conditions and, more importantly, the way in which proper initial conditions, including solitary and periodic waves, have been generated.

5.1 Boundary conditions

Two types of Boundary Conditions (BC) have been considered here : periodic and outflow conditions. Periodic BCs have been applied quite naturally by adding both the right hand sides and linear system matrix lines (cf. paragraph on Newton solver in section §6) of two coupled periodic nodes.

Outflow BCs represent a bigger challenge since wave reflection at the boundaries might pollute the inner domain solution. As in many other works (see *e.g.* [71, 72, 26] and references therein), this condition is mimicked here with a viscous *sponge layer* added to the spatial

domain whose sole function is to dump completely all the waves passing through. In these layers, (1) is replaced by

$$\begin{cases} \partial_t \eta + \partial_x q = \nu \partial_{x^2} \eta \\ \partial_t q - Bh^2 \partial_{x^2 t} q - \dfrac{1}{3} h \partial_x h \partial_{xt} q + \partial_x (uq) + gH \partial_x \eta + \\ \qquad\qquad - \beta g h^3 \partial_{x^3} \eta - 2\beta g h^2 \partial_x h \partial_{x^2} \eta = \nu \partial_{x^2} q \end{cases}$$

with the additional terms discretized with a Galerkin scheme, independently on the choice of the method used to approximate the rest of the system. Following [71], the viscosity ν is set to zero everywhere except in the layer $x \in [X_{s1}, X_{s2}]$ where it is defined as:

$$\nu = n_1 \frac{e^{n_2 \left(\dfrac{x - X_{s1}}{X_{s2} - X_{s1}} \right)} - 1}{e - 1} \tag{31}$$

Here n_1 and n_2 are constants used to fine tune the amount of viscosity. The values of these coefficients, as well as the length of the layer, are problem dependent. For the problems considered here, we found that for the problems considered here a layer length of approximately $5-10$ meters, and values of the coefficients of $n_1 \approx 10^{-3}$ and $n_2 \approx 10$ are a good start, but fine tuning is always preferable to make sure that the amount of reflected information is negligible.

5.2 Solitary wave generation

Several test problems involve the propagation of solitary waves over complex bathymetries. An analytical exact soliton for (1) is not available, however, a numerical approximation of exact solitary waves for (1) can be obtained. In particular, following [44], we seek a solution of (1), on a flat bathymetry with reference depth h_0, having the self similar behavior

$$W = (\eta,\ q)^T = W(\xi) = W(x - Ct)$$

with a certain celerity C. We also require this solution and all its derivatives with respect to ξ to go to zero at infinity. Plugging the expression $W = W(\xi)$ in (1) with $d = c^t = h_0$, integrating once between $-\infty$ and ξ, we obtain the relations

$$\begin{cases} -C\eta + q = 0 \\ -Cq + CBh_0^2 q'' + \dfrac{q^2}{h_0 + \eta} + g h_0 \eta + \dfrac{g}{2}\eta^2 - \beta g h_0^3 \eta'' = 0 \end{cases} \tag{32}$$

The first relation can be used to derive a second order ODE for q, namely :

$$Ch_0^2 \left(B - \beta \frac{C_0^2}{C^2} \right) q'' + C \left(\frac{C_0^2}{C^2} - 1 \right) q + \frac{g}{2C^2} q^2 + C \frac{q}{q + Ch_0} = 0 \tag{33}$$

with $C_0^2 = gh_0$. Note that this equation can be pre-multiplied by q' and integrated once more between $-\infty$ and ζ to give

$$\frac{Ch_0^2}{2} \left(B - \beta \frac{C_0^2}{C^2} \right) (q')^2 + \frac{C}{2} \left(\frac{C_0^2}{C^2} - 1 \right) q^2 + \frac{g}{6C^2} q^3 + Cq\left(\frac{q}{2} - Ch_0\right) + C^3 h_0^2 \ln \frac{q + Ch_0}{Ch_0} = 0 \tag{34}$$

Under the hypothesis that a solitary wave solution does exist, and that in correspondence of its maximum we have $q' = 0$, and $q = q_{\max} = CA$, with A the amplitude of the wave, we obtain a relation between the wave celerity and the wave amplitude :

$$\left(\frac{C}{C_0}\right)^2 = \frac{1}{2}\left(\frac{A}{h_0}\right)^2 \frac{1 + \frac{1}{3}\frac{A}{h_0}}{\frac{A}{h_0} - \ln\left(1 + \frac{A}{h_0}\right)} \tag{35}$$

Once we set the value of the ratio A/h_0, we can compute the celerity from (35). With these data, we have integrated (33) as a first order system of ODEs from $\xi = 0$ to $\xi = \infty$ with initial conditions $q_0 = CA$, and $q_0' = 0$. This has been done with the standard third order Runge-Kutta SSP scheme [36]. As an example, the η profile obtained for $h_0 = 1\,[m]$ and $A/h_0 = 0.2$ is given in figure 5.

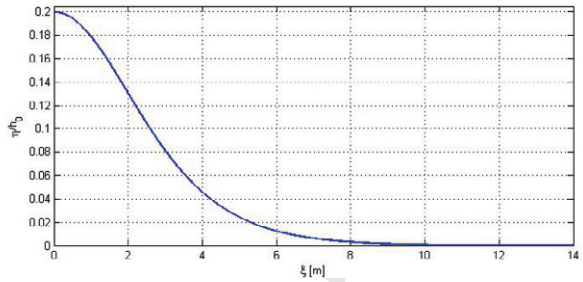

Figure 5: Soliton profile for $h_0 = 1\,[m]$ and $A/h_0 = 0.2$ obtained by numerically integrating (34)

To confirm the validity of the solution, preliminary calculations have been performed with the Galerkin scheme on a very fine mesh. The results for $A/h_0 = 0.2$ are reported in figure 6 showing the perfect math between the fine mesh solution of the Galerkin scheme and the approximate exact solution computed integrating (34).

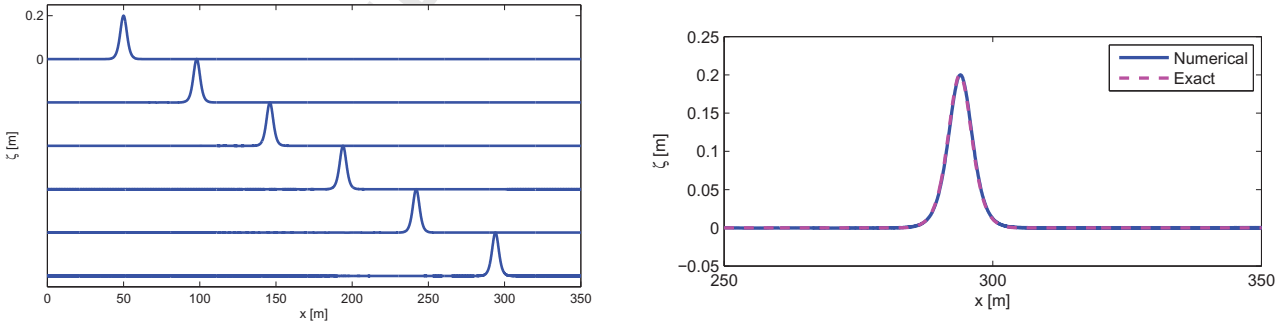

Figure 6: Left : computed evolution of the exact $A/h_0 = 0.2$ soliton. Right : comparison between the exact and the numerically computed solitary wave at the final time of the simulation.

5.3 Internal wave generation

A large number of tests involve the interaction of monochromatic periodic waves of small amplitude with a given bathymetry. In absence of an exact solution, the generation of such

periodic waves is a bit tricky. For very small amplitude waves, some authors suggest the use of Dirichlet type conditions with imposition at the *inlet* boundary of the condition (see *e.g.* [68] and references therein)

$$(\eta_b(t),\, q_b(t))^T = (A\sin(\omega t),\, C_0 A\sin(\omega t))^T$$

with $C_0^2 = gh_0$. We have found numerically the use of this approach relatively inefficient, first due to the fact that the signal obtained presents a transient phase after which it stabilizes to a periodic wave of amplitude generally larger than A, and more importantly due to a poor iterative convergence in the newton loop, requiring from two to three more iterations w.r.t. the convergence usually observed.

We have thus chosen a different approach, which is quite successfully used for the Boussinesq model of Nwogu [53]. This approach consists in adding to the η equation an internal source term of periodic variation in time. Following [73, 71], the first equation of (1) is modified as follows :

$$\partial_t(\eta + h_{\text{iwg}}) + \partial_x q = 0 \tag{36}$$

where the form of the internal wave generation term is taken to be:

$$h_{\text{iwg}}(x,t) = f_{\text{iwg}}(x)\, A_{\text{iwg}} \sin(\omega t) \tag{37}$$

with $\omega = 2\pi/T$, with T the period of the required signal. The spatial dumping function $f_{\text{iwg}}(x)$ is set to

$$f_{\text{iwg}}(x) = \Gamma_{\text{iwg}} e^{-(x-x_{\text{iwg}})^2/d_{\text{iwg}}^2} \tag{38}$$

These expressions simulate an undulating Gaussian hill centered at the position $x = x_{\text{iwg}}$. The constant A_{iwg} is always set equal to the amplitude of the signal we want to obtain, so the main trick is the choice of d_{iwg} and Γ_{iwg}. The choice of these parameters depends on the type of Boussinesq model considered and should be performed using Green's function method to the linearized equations including the source, as in [73]. Such a study, which is beyond the scope of this paper and a subject of investigation in itself, has never been performed for the model considered here. We can only mention the work of [48], which however does not consider a Gaussian source as in (38), but a localized delta function (cf. also the study presented in [66]).

To obtain some guidelines to fix the constants involved, at least in the range of amplitudes and wave periods considered in the benchmarks presented later, we have proceeded as follows. As in [73], we assume that the width the generation region $[x_{\text{iwg}} - a,\, x_{\text{iwg}} + a]$ should be proportional to the wavelength λ of the signal to be generated :

$$2a = \alpha_{\text{iwg}}\lambda$$

Of course we must make sure that on the boundaries of the generation zone $f_{\text{iwg}}(\pm a)$ is very small. So we set $f_{\text{iwg}}(\pm a) = 10^{-9}$ and we obtain

$$\ln 10^{-9} = -\frac{\alpha_{\text{iwg}}^2 \lambda^2}{4 d_{\text{iwg}}^2}$$

from which we deduce

$$d_{\text{iwg}} = \frac{1}{2}\frac{\alpha_{\text{iwg}}\lambda}{\sqrt{-\ln 10^{-9}}} \approx \frac{1}{2}\frac{\alpha_{\text{iwg}}\lambda}{\sqrt{20}}$$

which is consistent with the estimate for the β constant in the source function of [73] (equation (47) on page 279 in the reference). Next we assume the scaling constant Γ_{iwg} to depend on the dispersion parameter λ/h_0. In particular, we numerically found this parameter to be more important for longer waves. So we have set $\Gamma_{\text{iwg}} = \beta_{\text{iwg}}\lambda/h_0$ leading to

$$f_{\text{iwg}}(x) = \beta_{\text{iwg}}\frac{\lambda}{h_0}e^{-(x-x_{\text{iwg}})^2/d_{\text{iwg}}^2}, \quad d_{\text{iwg}} = \frac{1}{2}\frac{\alpha_{\text{iwg}}\lambda}{\sqrt{20}}$$

In practice, one might either use the value of the wavelength of the signal to be generated, or simplify further by using the assumption $\lambda \approx C_0 T = \sqrt{gh_0}T$, with T the period of the signal sought, leading to

$$f_{\text{iwg}}(x) = \beta_{\text{iwg}}\frac{\sqrt{gh_0}T}{h_0}e^{-(x-x_{\text{iwg}})^2/d_{\text{iwg}}^2}, \quad d_{\text{iwg}} = \frac{1}{2}\frac{\alpha_{\text{iwg}}\sqrt{gh_0}T}{\sqrt{20}} \tag{39}$$

In our tests we have found this choice of the source to work quite well for $\alpha_{\text{iwg}} \in [1, 4]$ (the higher the longer the wave), and for $\beta_{\text{iwg}} \in [0.15, 0.21]$. Fine tuning, always allowing to obtain the desired amplitude within few percent of error.

Concerning the numerical treatment of (36), the term $h_{\text{iwg}}(x,t)$ is carried along in *all the discretization steps* together with η. Thus all the time increments $\Delta^{n+1}\eta$ (cf. (23)) are replaced everywhere by $\Delta^{n+1}(\eta + h_{\text{iwg}}) = \eta^{n+1} + h_{\text{iwg}}^{n+1} - \eta^n - h_{\text{iwg}}^n$. Typically, the use of this wave generation technique produces a transient signal of a few periods which must be discarded in the simulation.

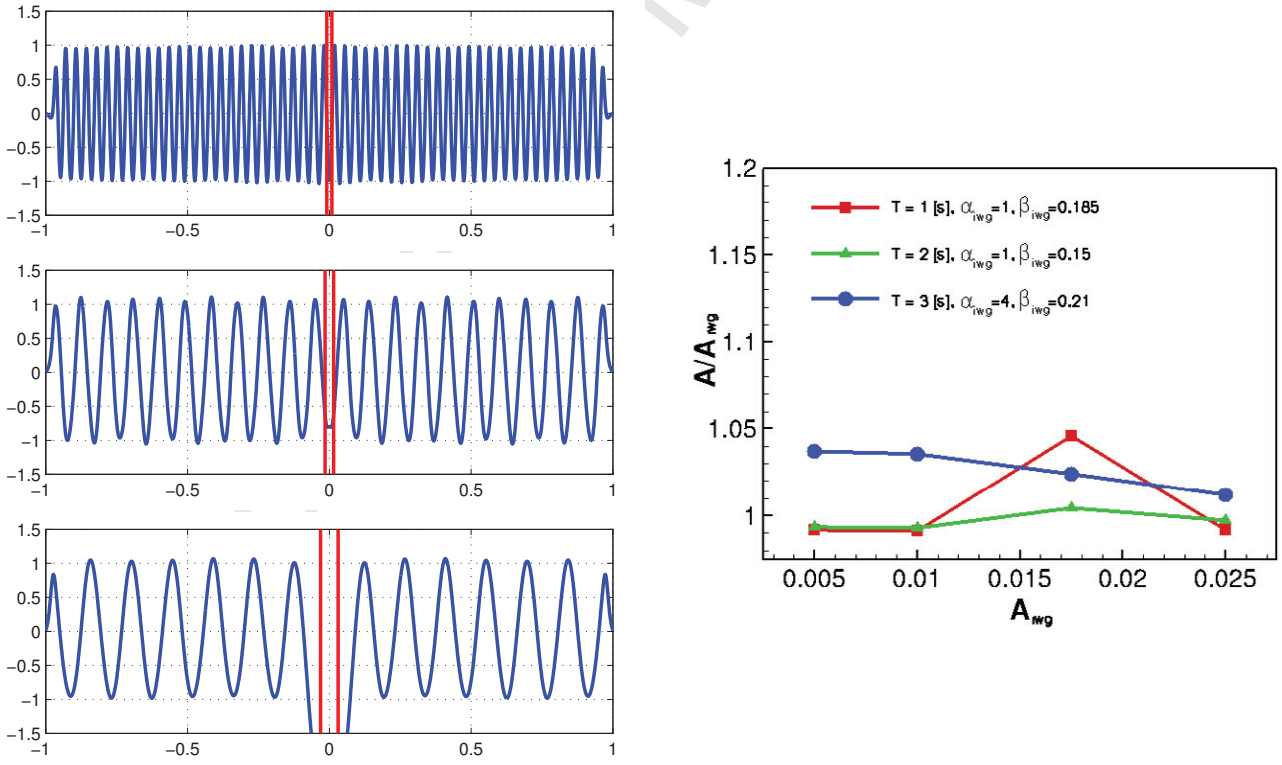

Figure 7: Left : waves obtained for the values $(T,\, A) = (1\,[s],\, 0.025\,[m])$ (top), $(T,\, A) = (1\,[s],\, 0.01\,[m])$ (middle), and $(T,\, A) = (3\,[s],\, 0.005\,[m])$ (bottom), both the x and η axes are normalized. Right : ratio between amplitude of signals obtained and A_{iwg} as function of A_{iwg}.

To give an example of the behavior obtained, we consider the generation of waves of periods $T = \{1\,[s],\, 2\,[s],\, 3\,[s]\}$, with different amplitudes. To generate the signals we have set, independently on the amplitude sought, $(\alpha_{\text{iwg}},\, \beta_{\text{iwg}}) = (1,\, 0.185)$, for the case $T = 1,[s]$, $(\alpha_{\text{iwg}},\, \beta_{\text{iwg}}) = (1,\, 0.15)$, for the case $T = 2,[s]$, and $(\alpha_{\text{iwg}},\, \beta_{\text{iwg}}) = (4,\, 0.21)$, for the case $T = 3,[s]$. On the right on figure 7 we show the ration between the value of the amplitude obtained over the input amplitude A_{iwg}, as a function of A_{iwg} itself, for the values of A_{iwg} close to those used in the benchmarks discussed later in the paper. The picture shows that, even without fine tuning of the parameters for each different amplitude, the amplitude of the signal obtained is within 5% of A_{iwg}. The left pictures on the same figure show the scaled values of $\eta(x)$ obtained for the cases $(T,\, A) = (1\,[s],\, 0.025\,[m])$ (top), $(T,\, A) = (1\,[s],\, 0.01\,[m])$ (middle), and $(T,\, A) = (3\,[s],\, 0.005\,[m])$ (bottom). The red vertical lines enclose the generation region, while the left and right boundaries are sponge layers.

6 CPU cost estimation in one dimension

To have an indication of the computational cost of the schemes and of the savings obtained by freezing the Newton Jacobian, we have compared the schemes' unit cost, defined as the CPU time per time step, node, and Newton iteration, and the cost per node, and time step. These two quantities, denoted by CPU and $\text{CPU}_{\Delta t}$, are obtained by computing approximations of the solitary wave solution of section §6.2 on different meshes, and computing the average times and Newton iterations needed to converge at each time step and for each scheme.

The results are summarized in table 1, in which we report: the unit CPU time per time step, node, and Newton iteration in the frozen Jacobian (column 2) and variable Jacobian (column 3) cases, and the ratio of these two (column 4), giving an indication of the relative cost of the Jacobian assembly for each scheme; the average (per time step) number of Newton iterations needed to converge at each time step (column 5) ; the unit CPU time per time step and per node in the frozen Jacobian (column 6) and variable Jacobian (column 7) cases, and the ratio of the two (last column), giving an indication of the relative cost of the Jacobian assembly per time step. The Newton convergence threshold for all the cases is 10^{-14}, and all the computations have been run on a portable 2.66 Ghz Intel Dual Core PC with 4 GB of RAM memory

	CPU^0	CPU	CPU^0/CPU	$\bar{\text{k}}_{\max}$	$\text{CPU}^0_{\Delta t}$	$\text{CPU}_{\Delta t}$	$\text{CPU}^0_{\Delta t}/\ \text{CPU}_{\Delta t}$
FD2	1.9965	2.5532	0.78196	4	7.89725	10.2737	0.7687
FDWK	2.3129	2.93118	0.78907	3.5	8.0469	10.2529	0.7848
cG	3.4338	3.83428	0.89556	2.5	8.513	9.7083	0.87688
cRD	3.19209	3.16557	1.0084	2	6.3842	6.33113	1.008
SUPG	5.56998	6.12684	0.90911	2.5	13.84496	16.49361	0.8394
URD	4.56813	4.79167	0.95335	2.5	11.17793	11.83333	0.9446

Table 1: Performance comparison of the 1d implementation of the schemes. CPU time units : $10^{-5}\,[s]$ (see text for a complete description).

The table shows that the FD2 scheme is the least costly in terms of operations per node and time step and Newton iteration, while the SUPG is the most expensive one. The relative

cost of the Jacobian assembly is measured by the fourth column which shows that for the finite difference schemes freezing the Jacobian leads to average savings per Newton step of over 20%. This is also a consequence of the relatively large number of Newton steps to converge, which is of about 4. On the other hand the cRD scheme, which only requires 2 Newton iterations to converge, shows absolutely no gains in Freezing the Jacobian $\mathcal{M}$. For the cG, SUPG, and URD schemes the gain the third column shows that the Jacobian assembly in average takes between 5 and 15% of the CPU time.

The results reported in the last three column are more significant from the point of view of obtaining the final result, giving the average CPU times to advance of one time step. These figures bring the Jacobian assembly costs at around 30-35% for the FD schemes, and confirm the rough 5-15% estimate for the cG, SUPG, and URD schemes. On the other hand, we see a slight reduction of the advantage of the FD schemes over the others. In particular, the FD2, FDWK and cG schemes have now comparable CPU times per time step, while the SUPG remains the most expensive scheme, requiring roughly 60% more CPU than the cG scheme, and 70% more when compared with the FDWK in the frozen Jacobian case. The URD figures are slightly better than the SUPG ones (roughly 40% times more expensive than the FDWK scheme), while the cRD scheme, due to its rapid Newton convergence, is the fastest of all.

Given our unoptimized implementation of the schemes one must take these figures very carefully. However, to be fair to the centered schemes, one must recognize that using an upwind stabilization brings a considerable additional computational cost, here perhaps overestimated at roughly 60%-70% times the CPU time of an "unstabilized" scheme. These figures can be compared to those, perhaps more reliable, obtained with our 2d implementation and discussed later in section §8.4. From the results of table 1 we also see that just by freezing the system Jacobian on obtains gains of about 10-15% for the residual based schemes, and of 20%-25% for the FD schemes. In section §8.4 we will see how this can be considerably improved.

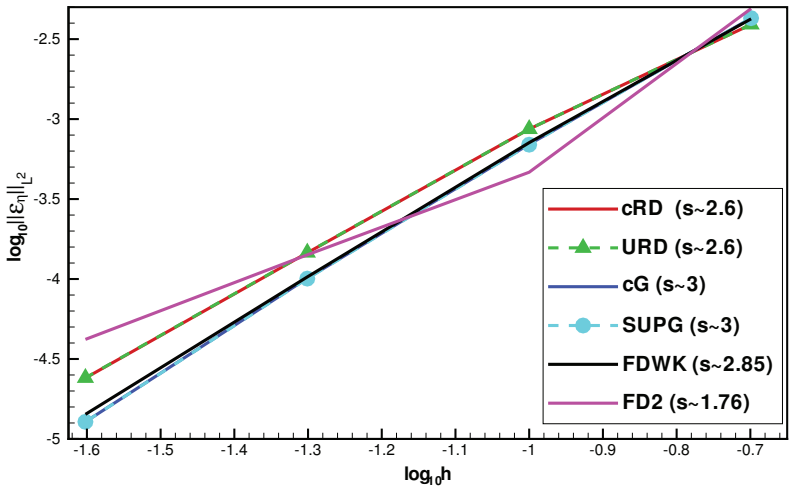

Figure 8: Grid convergence for the cG, cRD, SUPG, URD, FD2, and FDWK schemes.

7 Numerical tests in one dimension

7.1 Solitary wave propagation : grid convergence

We have verified the accuracy of the schemes by performing a grid convergence study on the solitary wave solutions of section §6.2. The solution used is a soliton propagating on a depth of $h_0 = 1$m, and with a ratio $A/h_0 = 0.2$, giving a celerity $C \approx 3.44$m/s (a value of $g = 9.8066\text{m/s}^2$ has been used everywhere). The error is computed after the solitary wave has moved of 100m from its initial position, on meshes containing 1000, 2000, 4000, and 8000 cells.

The results are reported on figure 8 in terms of L^2 norm of the error in η, and confirm, in part, the theoretical expectations. For the FD2 scheme we get the expected second order rate (the slope obtained is actually 1.76), with very small errors on the lower resolutions. For both the RD schemes we obtain slopes well above 2 (close to 2.6), with errors below the FD2 scheme only on the finest resolution used. Finally, the FDWK, cG, and SUPG schemes provide a slope very close to 3 (close 2.98 for cG and SUPG, and closer to 2.85 for FDWK). These results confirm numerically our previous observation concerning the accuracy of the SUPG scheme. We did not observe the fourth order of convergence neither for the FDWK nor for the cG, which really show errors of the same magnitude of the SUPG. Quite surprisingly, the presence of the upwind term affects the error to the order of the second or third decimal (in $\log_{10}$). Note that, in order to isolate the effects of the spatial discretization, to obtain these results the time step has been set to

$$\Delta t = 100\frac{\Delta x^3}{C}$$

The number of time steps thus grows significantly on the finer meshes, leading to a number of time iterations and CPU times which are not representative of the cost of the spatial discretization. Clearly, high order time stepping would have been preferable for a full comparison including computational times.

In conclusion, to be completely fair, if on one hand we stress once more that the SUPG scheme provides error magnitudes and convergence rates very close to the cG and FDWK schemes, the performance comparison of section §6 leads to the final result that retaining this level of accuracy with an upwind scheme requires a considerably higher CPU time.

7.2 Head-on collision of two solitary waves

A common test for the Boussinesq-type and non-hydrostatic models is the simulation of the interaction of two identical solitary waves propagating in opposite directions. After the interactions, one should ideally recover the initial profiles. The collision of the two waves presents additional challenges to the model by a sudden change of the nonlinear and frequency dispersion characteristics. The numerical model must handle the equilibrium between amplitude and frequency dispersion to propagate the wave profile at constant shape and speed.

We present here the interaction of two solitons propagating on a depth of $h_0 = 1$m with amplitude $A/h_0 = 0.2$. The spatial domain is [0 200] m, and the the initial solution is represented in figure 9. The computation uses a grid size $\Delta x = 0.08$ m, and a time step $\Delta t = 0.015$ s.

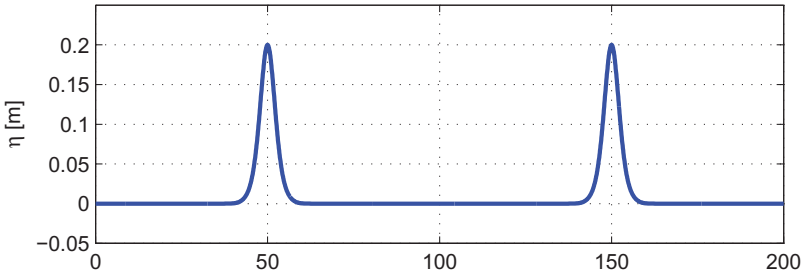

Figure 9: Head on collision of two solitary waves : initial solution.

The results obtained are summarized on figures 10 and 11. In particular, the pictures on figure 10 show the superposition of the solutions obtained with the four schemes studied in the paper at the time during the interaction when maximum amplitude is reached (left picture) and after the interaction (right picture). No differences can be observed at this scale. On the left on figure 11 the soliton profiles recovered at the final state with the SUPG scheme are plotted against the exact profiles, showing an excellent agreement, while on the right a zoom the the soliton peak on the left is presented. This close up view shows once more the higher resolution of the cG and SUPG schemes w.r.t. the RD schemes, and also the very similar accuracy of the cG and SUPG.

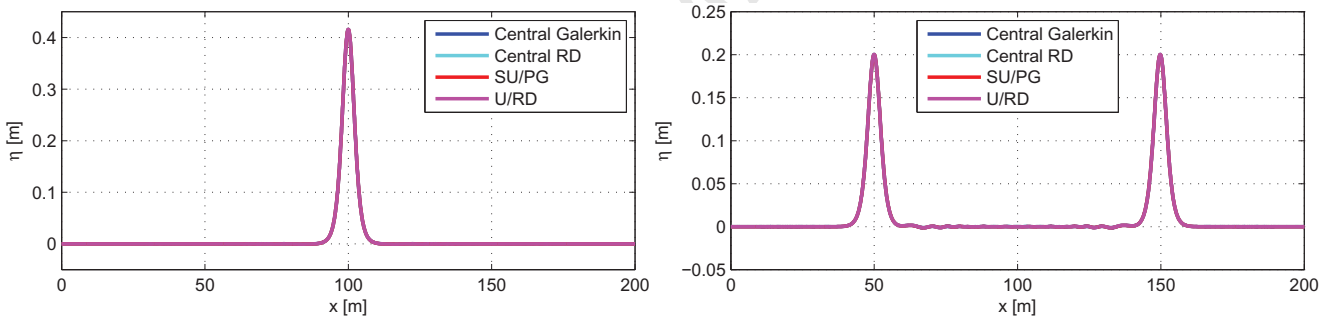

Figure 10: Head on collision of two solitary waves. Left : solution corresponding to maximum amplitude. Right : final state after the interaction. The solutions of all the schemes are superposed.

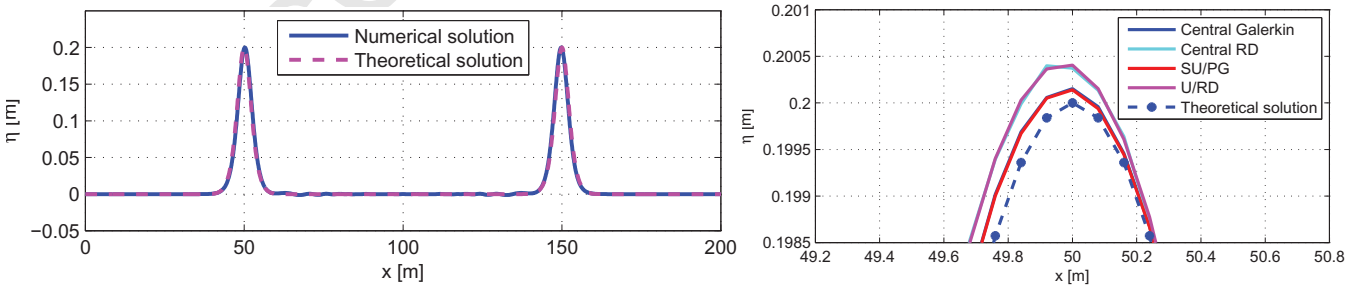

Figure 11: Head on collision of two solitary waves. Left : final state after the interaction, comparison of the SUPG solution with the exact soliton profiles. Right : final state after the interaction, close up view of the schemes resolution of the soliton peak.

7.3 Wave propagation over a shelf

This test is proposed in [44] : a solitary wave of amplitude $A = 0.2$ m propagates over a still water level of depth $h_0 = 1$ m. At $t = 0$ the soliton is placed at $x = 80$ m and it propagates on a shelf of slope $1 : 20$ (figure 12), over which the water depth is reduced to $h = 0.5$. The computational domain for this test is [0 280]m. As in [44] we have set the grid spacing to $\Delta x = 0.1$m, while the time step has been set to $\Delta t = 0.029$ s.

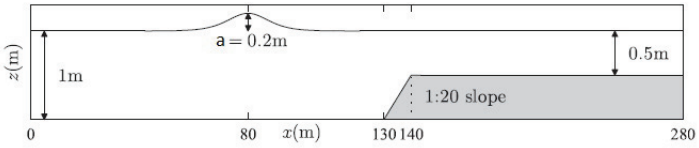

Figure 12: Sketch of the submerged shelf test.

Due to the interaction with the shelf, the solitary splits into several waves. The smallest in amplitude is a reflected wave characterized by a very long wavelength, while three forward waves with considerably higher amplitude are observed. A visualization of the wave transformation in time computed by the SUPG scheme is reported on figure 13, showing the nice capturing of the wave shoaling and splitting due to the interaction with shelf. The result compares very well with that reported in [44].

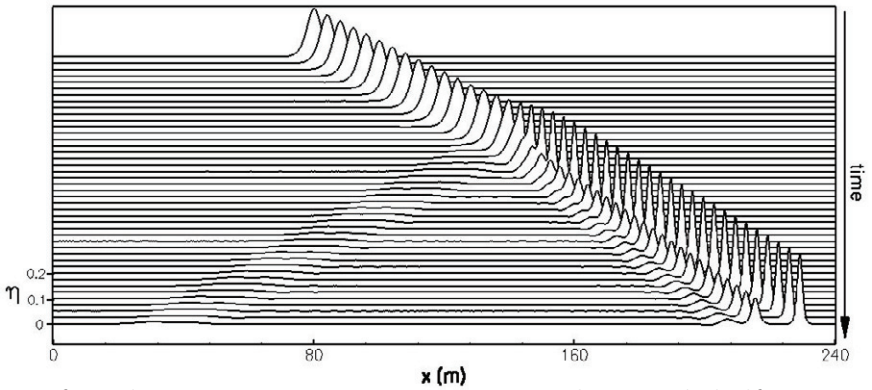

Figure 13: Splitting of a solitary wave propagating over a submerged shelf : numerical computation using SUPG scheme.

On figure 14 we compare the the results computed by the different schemes, having added the FDWK result for comparison. Note that the cRD and URD curves are right on top of each other, thus the curve relative to the cRD result (green) is not visible. The interaction is relatively fast, so it is hard to see any relevant differences between the prediction of the different schemes on the overall shape of the solution, as it is clear from the top picture on figure 14. The close ups of the three forward moving waves (bottom pictures in the same figure), only show a small phase difference, particularly visible on the highest wave produced for which the RD and FDWK scheme provided a larger celerity. In the bottom-right figure, we report a reference solution obtained with the cG scheme on a refined mesh (dashed line). With respect to this reference solution, the SUPG is the one providing the smallest phase error.

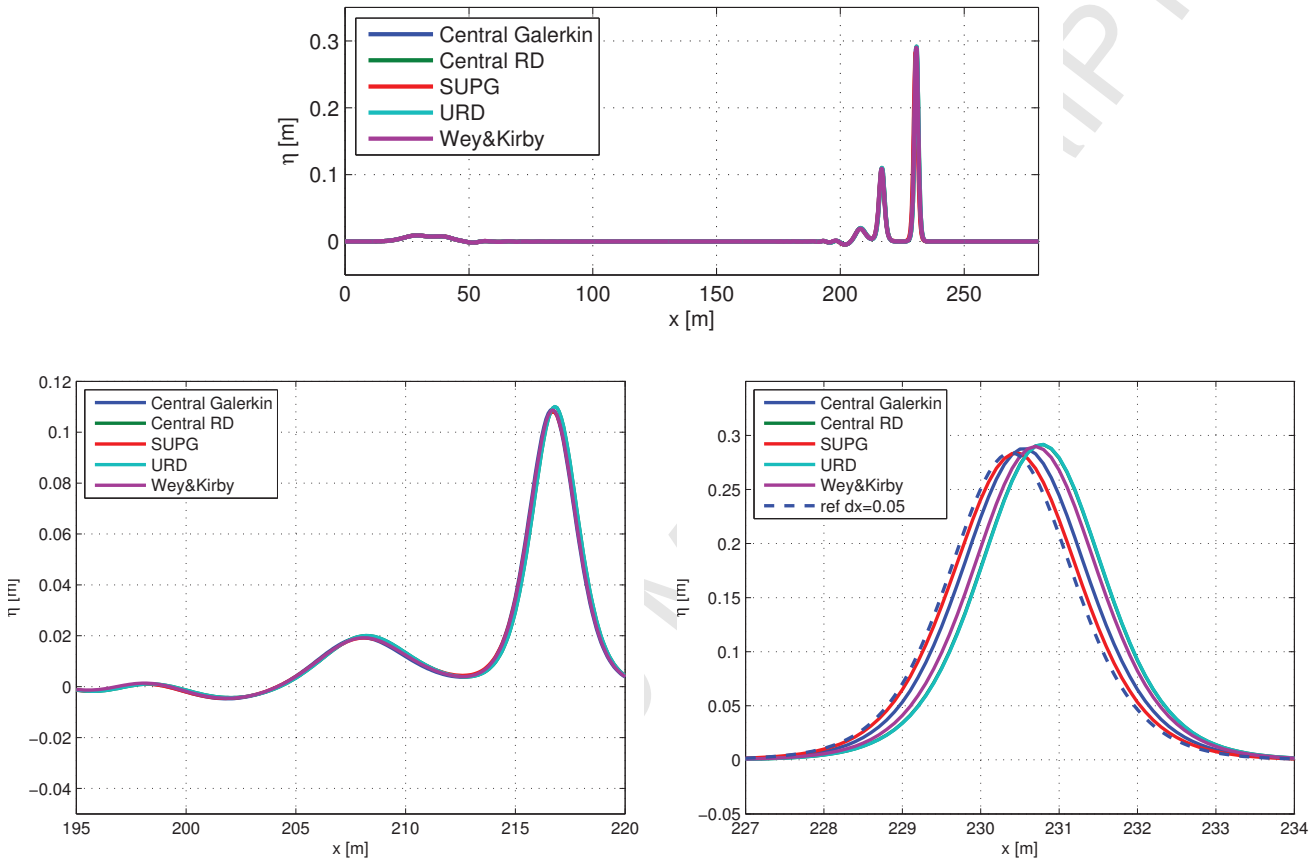

Figure 14: Solitary wave interaction with a shelf. Superposition of the solutions obtained with the cRD, URD, cG, SUPG, and FWDK schemes. Top: overall view. Bottom left: close-up of the shorter peaks. Bottom right: close-up of the highest peak.

7.4 Periodic wave propagation over a submerged bar

In [64] a series of laboratory experiments investigating wave propagation and dispersion over a submerged bar are described. The experimental data collected are often used to validate non-hydrostatic models for wave propagation. In particular, we consider here the experiment with the configuration shown in figure 15. The test consists in a periodic wave of period $T = 2.02$ s and amplitude $A = 0.01$ m which propagates into an initially undisturbed region of depth $h_0 = 0.4$m before reaching a bar of the shape and proportions given in the figure.

The numerical set for this test case needs the use of the periodic internal wave generator centered at the coordinate $x = 10$ m of the domain $x \in [0\ 35]$ m. Two sponge layers are then used at the two boundaries of the domain, extended on the regions $x \in [0\ 3]$ m on the left and $x \in [32\ 35]$ m on the right, in order to absorb any wave reaching the boundaries.

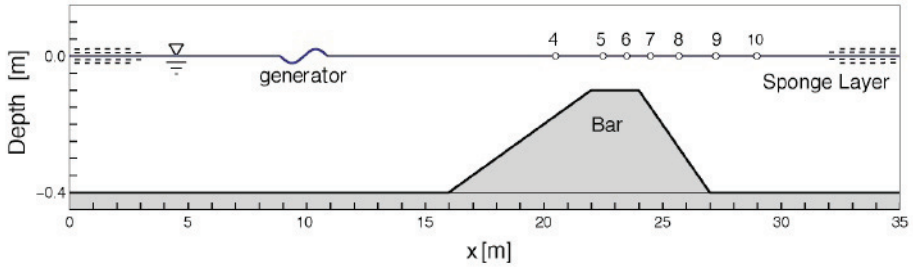

Figure 15: Sketch of the computational configuration of the numerical test of the propagation over a submerged bar.

The values of grid size and time step have been set to $\Delta x = 0.04$m and $\Delta t = 0.0323$, similar to those used by other authors (see *e.g.* [20] and references therein). For this choice of parameters, the cRD scheme has shown a strongly unstable behavior, so no results are shown with this scheme. For the cG, URD, SUPG, and FDWK schemes, the numerical results are compared with experimental data in gauges placed at $x_4 = 20.5$m, $x_5 = 22.5$ m, $x_6 = 23.5$ m, $x_7 = 24.5$ m, $x_8 = 25.7$ m, $x_9 = 27.3$m, and $x_10 = 29$m, denoted as gauge g4, g5, g6, g7, g8, g9, g10. Numerical simulations have been run for a over 40s, and the computed signals in the gauges have been compared to the experimental ones after a steady periodic flow has been obtained. Due to the periodic character of the solution, as in [20, 71], a phase calibration has been necessary to be able to compare computed and measured signals. This calibration has been performed here by introducing a small time shift in the computed results, such that the phase error w.r.t. the data in g4 is minimized, as shown on figure 16. Once this time shift is obtained, the same value is applied to all the other gauges.

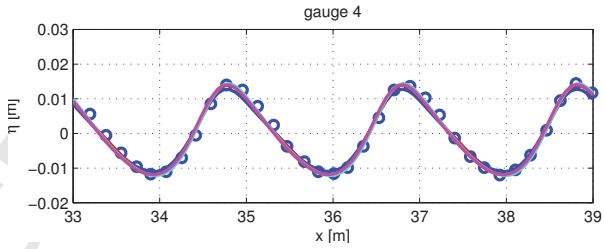

Figure 16: Data in gauge g4 for the submerged bar experiment of [64] (amplitude $A = 0.01$ m and period $T = 2.02$ s): phase calibration. Experimental data (o), FDWK scheme (—), cG scheme (—), SUPG scheme (—), and URD scheme (—).

This problem provides a severe test of the model, as nonlinearity initially steepens the waves on the up-slope, and then the increasing depth behind the bar decomposes the waves into short wave components. This results in a rapidly varying profile behind the bar with the exact form depending crucially on the dispersive characteristics of the numerical model. Due to the theoretical limitations of the model used, some discrepancies may appear show up as higher harmonics are released behind the bar. The comparison between computed and experimental signals is reported on figure 17. All the schemes perform very well up to the

position of gauge g7 at the end of the plateau. Starting from g8, we can see major differences appearing. In particular, already in g8 the URD scheme shows a significant phase shift w.r.t. all the other schemes, while in g9 and g10 we can clearly see the appearance of spurious higher harmonics.

The cG, SUPG, and FDWK provide a satisfactory prediction of the experimental data, the FDWK giving a slight underestimation of the amplitudes in both g9 and g10. Nevertheless the agreement is still very reasonable, and the discretization schemes proposed represent quite accurately the nonlinear and dispersive properties of the original continuous model. Moreover the similar resolution properties of the SUPG, cG, and FDWK schemes are once more confirmed.

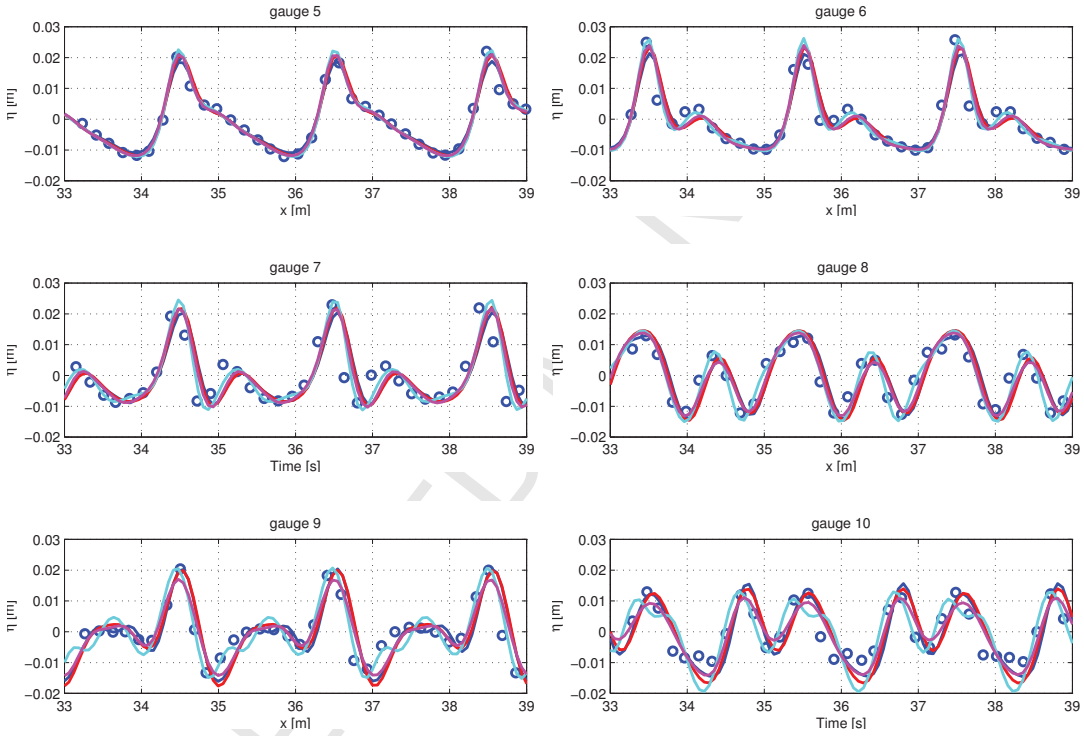

Figure 17: Data in gauges g5 to g10 for the submerged bar experiment of [64] (amplitude $A = 0.01$ m and period $T = 2.02$ s). Experimental data (o), FDWK scheme (—), cG scheme (—), SUPG scheme (—), and URD scheme (—).

8 Extension to two space dimensions

The objective of this section is to show that also in two space dimensions, as in 1d, the upwind Petrov-Galerkin schemes studied in this paper provide results comparable to those used in published literature, on meshes comparable to those used to obtain these results.

In particular we consider the discretization of a two-dimensional version of (1) on unstructured triangulations by means of two upwind stabilized Petrov-Galerkin schemes, both of which reduce in 1d reduce to the SUPG scheme studied in the first half of the paper. These schemes can both be written as the C^0 Galerkin scheme plus a local Petrov-Galerkin stabiliza-

tion depending on an upwind biased perturbation of the test function. In the Shallow Water limit the schemes reduce to the standard SUPG scheme [43], and to the P^1 LDA Residual Distribution scheme with Petrov-Galerkin mass matrix [22, 61, 59].

8.1 Enhanced Boussinesq equations in two space dimensions

We consider the Boussinesq equations in the improved form proposed in [65, 54]. Denoting by $\vec{q}$ the mass flux vector, $\vec{q} \equiv (q_x, q_y)$, with $\vec{u} \equiv (u_x, u_y)$ the velocity vector, and with the same notation used in one dimension for the other variables, the model equations read :

$$\begin{cases} \partial_t \eta + \nabla \cdot \vec{q} = 0 \\ \\ \partial_t q + \nabla \cdot (\vec{u} \otimes \vec{q}) + gH\nabla\eta + \vec{\psi} = 0 \end{cases} \tag{40}$$

where $\vec{\psi} \equiv (\psi_x, \psi_y)$ are the dispersive terms of the model which can be written as [65, 54]

$$\begin{cases} \psi_x = -Bh^2\partial_{tx}\nabla\cdot\vec{q} - \dfrac{1}{6}h\partial_x h\partial_t\left(\nabla\cdot\vec{q} + \partial_x q_x\right) - \dfrac{1}{6}h\partial_y h\partial_{tx} q_y - \beta g h^2 \partial_x w^\eta \\ \\ \psi_y = -Bh^2\partial_{ty}\nabla\cdot\vec{q} - \dfrac{1}{6}h\partial_y h\partial_t\left(\nabla\cdot\vec{q} + \partial_y q_y\right) - \dfrac{1}{6}h\partial_x h\partial_{ty} q_x - \beta g h^2 \partial_y w^\eta \\ \\ w^\eta = \nabla\cdot(h\nabla\eta) \end{cases} \tag{41}$$

This form of the enhanced model reduces exactly to (1) for constant bathymetry in the one dimensional case. For variable bathymetry (40)-(41) correspond to a variant of the model with improved shoaling coefficient proposed in [65]. The main advantage of using this form of the equations, as compared to *e.g.* the ones used in [68], is that all the higher order derivatives of η, including the mixed ones, appear through the sole auxiliary variable w^η.This allows to reduce the number of auxiliary variables whose reconstruction is needed in the numerical implementation of the model (cf. next section). The values of the dispersion coefficients β and B are set, as in one dimension and as in [65, 54], to $\beta = 1/15$ and $B = \beta + 1/3$.

8.2 C^0 continuous Petrov-Galerkin discretization

Let Ω_h denote an unstructured triangulation of the spatial domain Ω, with h the reference element size, and with K denoting the generic element. For a node $i \in \Omega_\mathrm{h}$, let K_i denote the set of elements containing i. On Ω_h we consider piecewise linear continuous approximations η_h and $\vec{q}_\mathrm{h}$ of the variables of type (5), with standard piecewise linear continuous Lagrange bases (cf. figure 18).

The discrete version of (40) for a node $i \in \Omega_\mathrm{h}$ reads

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_\mathrm{h}, \vec{q}_\mathrm{h}) + \sum_{\mathrm{K}\in\Omega_\mathrm{h}} \int_\mathrm{K} \Gamma_i^\mathrm{K}\, r_\mathrm{h}^\mathrm{MS} = 0 \tag{42}$$

where, as in one dimension, $\mathcal{R}_i^{\mathrm{cG}}$ represents the nodal equation obtained with the Galerkin scheme, while Γ_i^K is a local stabilization matrix introducing some upwind bias in the scheme.

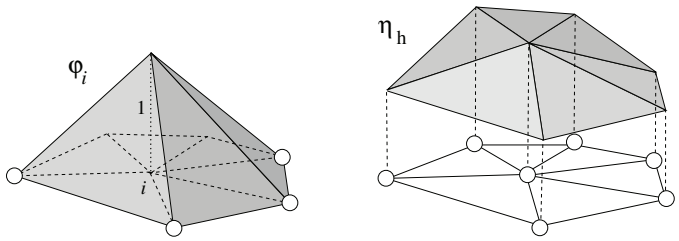

Figure 18: P^1 finite element interpolation

As done in one space dimension, in the P^1 case we can safely assume Γ_i^{K} to be constant within each element. This leads to the simpler form of our prototype scheme

$$\mathcal{R}_i^{\mathrm{cG}}(\eta_{\mathrm{h}},\, \vec{q}_{\mathrm{h}}) + \sum_{\mathrm{K}\in\Omega_{\mathrm{h}}} \Gamma_i^{\mathrm{K}} \Phi^{\mathrm{K}} = 0 \tag{43}$$

where, as in one space dimension, $\Phi^{\mathrm{K}} = \begin{bmatrix}\Phi_\eta^{\mathrm{K}} & \Phi_{\vec{q}}^{\mathrm{K}}\end{bmatrix}^T$ is the element residual defined as (cf. section §3.2, equation (8) and (10))

$$\begin{aligned}
\Phi_\eta^{\mathrm{K}} &= \int_K (\partial_t \eta_{\mathrm{h}} + \nabla\cdot\vec{q}_{\mathrm{h}}) \\
\Phi_{\vec{q}}^{\mathrm{K}} &= \int_K (\partial_t \vec{q}_{\mathrm{h}} + \nabla\cdot(\vec{u}\otimes\vec{q})_{\mathrm{h}} + gH_{\mathrm{h}}\nabla\eta_{\mathrm{h}}) + \Phi_{\vec{\psi}}^{\mathrm{K}} \\
\Phi_{\vec{\psi}}^{\mathrm{K}} &= \left(\Phi_{\psi_x}^{\mathrm{K}},\, \Phi_{\psi_y}^{\mathrm{K}}\right)^T = \int_{\mathrm{K}} \vec{\psi}
\end{aligned} \tag{44}$$

In practice, the cG term $\mathcal{R}_i^{\mathrm{cG}}$ has been implemented as

$$\mathcal{R}_i^{\mathrm{cG}} = \mathcal{R}_i^{\text{cG-NLSW}} + \mathcal{R}_i^{\text{cG-}\psi} \tag{45}$$

where, as in the one dimensional case, the first term can be written after re-integrating by-parts the Galerkin statement as

$$\mathcal{R}_i^{\text{cG-NLSW}} = \begin{bmatrix} \int_{\Omega_{\mathrm{h}}} \varphi_i\,(\partial_t\eta_{\mathrm{h}} + \nabla\cdot\vec{q}_{\mathrm{h}}) \\ \int_{\Omega_{\mathrm{h}}} \varphi_i\,(\partial_t\vec{q}_{\mathrm{h}} + \nabla\cdot(\vec{u}\otimes\vec{q})_{\mathrm{h}} + gH_{\mathrm{h}}\nabla\eta_{\mathrm{h}}) \end{bmatrix} \tag{46}$$

Fully discrete terms are obtained by evaluating the last expressions exactly w.r.t. a linear variation of η_{h}, $\vec{q}_{\mathrm{h}}$, H_{h}, and of the non-linear component of the NLSW advective flux $(\vec{u}\otimes\vec{q})_{\mathrm{h}}$. For the dispersive component we have $\mathcal{R}_i^{\text{cG-}\psi} = (0,\, \mathcal{R}_i^{\text{cG-}\psi_x},\, \mathcal{R}_i^{\text{cG-}\psi_y})^T$, where the momentum terms have been implemented as

$$\begin{aligned}
\mathcal{R}_i^{\text{cG-}\psi_x} = &\int_{\Omega_h} Bh^2 \left(\partial_{tx} q_{x\mathrm{h}}\partial_x\varphi_i + \frac{1}{2}(\partial_x q_{y\mathrm{h}}\partial_y\varphi_i + \partial_y q_{y\mathrm{h}}\partial_x\varphi_i)\right) \\
&-\frac{1}{6}\int_{\Omega_{\mathrm{h}}} \varphi_i h\,(\partial_x h\partial_t(\nabla\cdot\vec{q}_{\mathrm{h}} + \partial_x q_{x\mathrm{h}}) + \partial_y h\partial_{tx} q_{y\mathrm{h}}) - \int_{\Omega_{\mathrm{h}}} \beta g h^2 \partial_x w_{\mathrm{h}}^{\eta}
\end{aligned} \tag{47}$$

Fully discrete expressions are obtained after evaluation of (47) with approximate quadrature over each mesh element, with the hypothesis of piecewise linear variation of all the quantities involved. Note that, for symmetry reasons, the mixed derivative term in the first line of (47) has been expressed as an average of two terms for which the partial derivative in each space direction is passed onto the basis function. The form of the $\mathcal{R}_i^{\text{cG-}\psi_y}$ term is very similar.

For the element residual (44) fully discrete quantities are obtained by evaluating the integrals with numerical quadrature with the hypothesis of linear variation of all the quantities involved. In particular, the flux-divergence terms are evaluated directly as contour integrals, the hydrostatic term is evaluated exactly w.r.t a linear variation of H and η, while the dispersive terms are computed by evaluating exactly

$$\begin{aligned}\Phi_{\psi_x}^{\mathrm{K}} = -B \int_{\mathrm{K}} h^2 \Big(\partial_{tx} w_{x\mathrm{h}}^{q_x} + \frac{1}{2} (\partial_{tx} w_{y\mathrm{h}}^{q_y} + \partial_{ty} w_{x\mathrm{h}}^{q_y}) \Big) \\ - \frac{1}{6} \int_{\mathrm{K}} h \Big(\partial_x h \, (2\, \partial_{tx} q_{x\mathrm{h}} + \partial_{ty} q_{y\mathrm{h}}) + \partial_y h \, \partial_{tx} q_{y\mathrm{h}} \Big) - \beta g \int_{\mathrm{K}} h^2 \partial_x w_{\mathrm{h}}^{\eta}\end{aligned} \tag{48}$$

where, as in the Galerkin term (47), the second order mixed derivatives of q_y have been implemented, for symmetry , as an average of the x and y derivatives of the components of $\vec{w}^{q_y}$. A similar expression is used for $\Phi_{\psi_y}^{\mathrm{K}}$. The auxiliary variables $\vec{w}^{q_x} \equiv (w_x^{q_x}, w_y^{q_x})$ and $\vec{w}^{q_y} \equiv (w_x^{q_y}, w_y^{q_y})$ are defined as

$$\vec{w}^{q_x} = \nabla q_x \,, \quad \vec{w}^{q_y} = \nabla q_y$$

Nodal values of these quantities are obtained following [2, 5, 57] as

$$\begin{aligned} w_i^{\eta} &= - \frac{1}{\sum\limits_{\mathrm{K} \in \mathrm{K}_i} |\mathrm{K}|} \sum_{\mathrm{K} \in \mathrm{K}_i} \int_{\mathrm{K}} h \nabla \eta_{\mathrm{h}} \cdot \nabla \varphi_i \\ \vec{w}_i^{q_x} &= \frac{1}{\sum\limits_{\mathrm{K} \in \mathrm{K}_i} |\mathrm{K}|} \sum_{\mathrm{K} \in \mathrm{K}_i} |K| \nabla q_x^{\mathrm{K}} \\ \vec{w}_i^{q_y} &= \frac{1}{\sum\limits_{\mathrm{K} \in \mathrm{K}_i} |\mathrm{K}|} \sum_{\mathrm{K} \in \mathrm{K}_i} |K| \nabla q_y^{\mathrm{K}} \end{aligned} \tag{49}$$

where $|\mathrm{K}|$ is the area of the element. Note that, while all of the above expressions can be obtained by an L^2 projection with mass lumping on the left hand side, the last two boil down to a Green-Gauss reconstruction of the nodal gradients of $\vec{q}$ [5, 57].

8.3 Upwind stabilization

The one-dimensional analysis and the numerical tests have shown us that the Petrov-Galerkin SUPG scheme provides an accurate approximation of the enhanced Boussinesq system considered in this paper. Element based multidimensional upwinding, however, is a more complex procedure. In the following sections we present a one-to-one comparison of two different approaches, both reducing in the 1d case to the SUPG scheme presented in section §2.

The first approach reduces, when considering the NLSW limit, to the standard multidimensional P^1 SUPG finite element scheme [43]. In this case, we set in (43)

$$\Gamma_i^K = (\mathcal{A}^{\mathrm{K}} \cdot \nabla\varphi_i)\tau_{\mathrm{K}}$$

where $\nabla\varphi_i$ is the local value of the gradient of the shape function of node i in element K, which is constant and given by

$$\nabla\varphi_i = \frac{\vec{n}_i}{2|K|}$$

with $\vec{n}_i$ the inward pointing scaled normal to the edge facing node i. The matrix τ_{K} is a scaling matrix parameter, and $\mathcal{A}^{\mathrm{K}} \equiv (A_x,\, A_y)^{\mathrm{K}}$ is a local mean value of the Jacobian of the x and y components of the NLSW conservative fluxes. We recall that $\forall\, \vec{\xi} \in \mathbb{R}^2$ the matrix $\mathcal{A}^{\mathrm{K}} \cdot \vec{\xi}$ admits a full set of linearly independent eigenvectors associated to the thee eigenvalues $\vec{u} \cdot \vec{\xi}$ and $\vec{u} \cdot \vec{\xi} \pm C\|\vec{\xi}\|$ with $C^2 = gH$ (see e.g. [61] for detailed expressions). This decomposition can be used to define the scaling parameter τ_{K}, which has been set to [11, 23, 1, 61] (see also equation (17) section §3.4),

$$\tau_{\mathrm{K}} = \frac{2}{3}\left(\sum_{j\in\mathrm{K}} |\mathcal{A}^{\mathrm{K}} \cdot \nabla\varphi_j|\right)^{-1}$$

with the absolute value of a matrix computed using standard eigenvalues decomposition.

A different route to construct upwind biased discretizations relies on the Multidimensional Upwinding procedure introduced by P.L. Roe and H. Deconinck in the context of Residual Distribution. The interested reader can refer to [62, 70, 19, 22] and references therein for an overview and an historical perspective on this approach. Here, we will compare the SUPG to the Petrov-Galerkin form of the Multidimensional-Upwind Residual Distribution scheme known as LDA [22]. The earliest Petrov-Galerkin finite element formulation of multidimensional upwind RD schemes for time dependent problems is due to J. Maerz and G. Degrez for scalar advection [51], later generalized to the Euler equations by Ferrante and Deconinck [32]. Since then, several versions of this finite element analogy have been presented in literature [24, 3, 22, 59]. The simplest form of this analogy allows to rewrite RD schemes as in (43), with a Multidimensional Upwind stabilization obtained by setting

$$\Gamma_i^K = \beta_i^K - \frac{\mathrm{I}_3}{3}$$

with β_i^K the RD distribution matrix. In particular, in the case of the Multidimensional Upwind LDA scheme [22, 59] we have

$$\beta_i^K = \left(\mathcal{A}^{\mathrm{K}} \cdot \nabla\varphi_i\right)^+ \left(\sum_{j\in\mathrm{K}} \left(\mathcal{A}^{\mathrm{K}} \cdot \nabla\varphi_j\right)^+\right)^{-1}$$

where the positive part of a matrix is computed by standard eigenvalues decomposition. For further details concerning upwind RD schemes in the multidimensional case the interested reader can consult the references given above.

8.4 Implementation details

As in one dimension, the numerical simulations presented in the next sections make use of Crank-Nicholson integration in time, obtained by simply replacing the time derivatives $\partial_t(\cdot)$ by divided differences $\Delta^{n+1}(\cdot)/\Delta t = ((\cdot)^{n+1} - (\cdot)^n)/\Delta t$, and evaluating the remaining quantities at the $n+1/2$ level (cf. section §5). The model and the Newton algorithm discussed in section §5 has been implemented in the Shallow Water `C` code developed in [61], and modified in two dimensions as follows

1. Set $W_0 = (\eta_0^{n+1},\, q_0^{n+1})^T = (\eta^n\, q^n)^T$
2. Evaluate the *frozen* Jacobian matrix
$$\mathcal{M} = \frac{\partial F}{\partial W_{\rm h}}(W_{\rm h} = W_0 | \eta_{\rm h}^n,\, q_{\rm h}^n)$$
3. Compute a LU factorization of $\mathcal{M}$
4. for k $= 1,$ k$_{\max}$ do
   (a) Evaluate $F(W_{k-1}|\eta_{\rm h}^n,\, q_{\rm h}^n)$
   (b) If $\|F\| \le \epsilon$ set k $=$ k$_{\max}$ and exit, else evaluate $W_k = W_{k-1} - \mathcal{M}^{-1}F(W_{k-1}|\eta_{\rm h}^n,\, q_{\rm h}^n)$
5. Set $W_{\rm h} = W_{\rm k_{max}}$

In two dimensions, besides freezing the Jacobian, we also pre-compute and store its LU factorization. Thus, not only step 2. is the only one in which the sparse matrix $\mathcal{M}$ is assembled, but step 3. is the only step in which the actual "inversion" of $\mathcal{M}$ is performed. Thus, the underlined operation in step 4.(b) is reduced to a double matrix vector product.

All these steps are implemented making use of standard functionalities of the MUMPS algebraic library [8, 9]. In particular, the cost of step 4(b) is further reduced due to the use of optimized BLAS libraries made in "solve" function of MUMPS, while the computation and storage of the decomposition of $\mathcal{M}$ is completely transparent to the user. The convergence threshold ϵ is in practice replaced by the $\min(\epsilon,\, \epsilon_{\rm rel}\|F(F(W_0|\eta_{\rm h}^n,\, q_{\rm h}^n))\|$, with $\epsilon_{\rm rel}$ a relative tolerance. The typical convergence behavior over a few time steps for the tests considered in the paper is reported on figure 19 for values of the relative convergence of 10^{-4} and 10^{-6}, and for an absolute convergence threshold of 10^{-12}. Three sub iterations are in general enough to achieve a relative tolerance of four orders of magnitude, while four or five sub iterations are enough to go in between 6 orders of magnitude w.r.t. the initial residual value and machine accuracy.

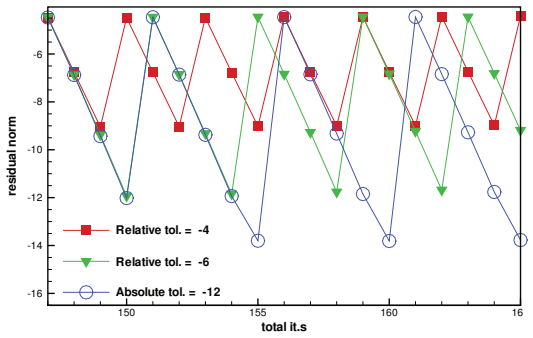

Figure 19: Typical Newton convergence of the 2d implemetnation

To measure the savings obtained with our 2d implementation of the Newton loop, we perform a test similar to the one of section §6. In particular, on the rectangle $[0,\, 400] \times [0,\, 0.4]$ we consider the generation of a planar periodic signal, using the generator of section §5.3 with $T = 2.525\, s$ and $A = 0.027\, m$. We compute 5 periods in time on several meshes, and then measure the average unit CPU time per node, time step and Newton iteration, and the average CPU time per time step and node. These are denoted by CPU and $\text{CPU}_{\Delta t}$ in table 2. The superscript ^{0} is added for the values obtained with the implementation discussed above, while the values without superscript refer to computations in which the Jacobian is re-computed and inverted at each Newton step. In table 2 we summarize the results obtained for the cG, SUPG and LDA schemes. The tolerance used in the Newton loop is 10^{-13}, and all the computations have been run on a portable 2.66 Ghz Intel Dual Core PC with 4 GB of RAM memory

	CPU^0	CPU	CPU^0/CPU	$\overline{\text{k}}_{\max}$	$\text{CPU}^0_{\Delta t}$	$\text{CPU}_{\Delta t}$	$\text{CPU}^0_{\Delta t}/\ \text{CPU}_{\Delta t}$
cG	3.0536	8.7984	0.3470	3.883	11.8838	34.3165	0.346
SUPG	3.6156	9.8664	0.3664	3.958	14.2446	39.0502	0.365
LDA	3.6376	9.8952	0.3676	3.962	14.4028	39.1848	0.367

Table 2: Performance evaluation of the 2d implementation of the schemes. CPU time units : $10^{-5}\,[s]$ (see text for a complete description).

The figures reported in table 2 allow to slightly correct the CPU over-head estimated in 1d, when comparing the performance of the Petrov-Galerkin schemes with Galerkin. In particular, the fifth and sixth columns allow to estimate this overhead to roughly 20% in the froze Jacobian case, which we believe being a more realistic estimate than the 60% obtained in 1d. The other interesting aspect is the gain obtained with our implementation of the Newton algorithm. We can see from the third column that at each Newton iteration we are able (in average) to save 75% of the CPU time by freezing the Jacobian matrix and storing its LU decomposition. Note that, since we are anyways using a direct solver, these gains in CPU time do not come at any additional memory requirements, since the memory used is that normally already required by the linear solver. Note also, that these requirements might be more demanding in the P^2 and P^3 cases, however will remain much more moderate for a C^0 interpolation than for a discontinuous one, and might anyways be mitigated by the need of a lower number of elements, thanks to the increase of accuracy.

Concerning the implementation of the solver, the most tricky part of the implementation is the inclusion in the Jacobian of the nonlocal terms. These are evaluated in to steps, first by computing nodal Jacobians of the auxiliary variables, which are then used in the Jacobian assembly en each element. In addition, the reflective boundary conditions

$$\vec{q}_{\text{h}} \cdot \vec{n} = 0\,, \quad \nabla \eta_{\text{h}} \cdot \vec{n} = 0$$

have has been implemented in a strong form by modifying the two components of the right hand side of the $\vec{q}_{\text{h}}$ equations, the auxiliary variable equations, and the related Jacobian entries such that

$$\mathcal{R}^{\vec{q}}_i \cdot \vec{n} = 0$$

at each Newton iteration.

9 Two-dimensional results

9.1 Wave diffraction over a semi-circular shoal

We consider here the reproduction of the tests carried out in [74] involving the study of the focusing effect induced by a semicircular shoal on wave trains of different periods. The experiments were carried out in a wave tank 6.096m wide and 25.6m long, its middle portion consisted in a semi-circular shoal leading the water depth to decrease from $h_0 = 0.4572$m (at the wave maker) to 0.1524m at the end of the tank. The bottom topography is described by the equation

$$z = \begin{cases} 0 & \text{if} \quad 0 \leq x < 10.67 - G(y) \\ (10.67 - G(y) - x)/25 & \text{if} \quad 10.67 - G(y) \leq x < 18.29 - G(y) \\ 0.30480 & \text{if} \quad 10.29 - G(y) \leq x \end{cases}$$

with $G(y) = \sqrt{y(6.096 - y)}$.The depth h is obtained as $h = h_0 - z$. A contour plot of the bathymetry is reported on the left on figure 20. This test is a standard benchmark for 2D dispersive models (see e.g. [55, 12, 72, 54, 31, 68, 46] and references therein).

Three cases are considered here, with wave trains characterized by [74]

(a) T=1 s, A = 0.0195 m, $h_0/\lambda = 0.306$;

(b) T=2 s, A = 0.0075 m, $h_0/\lambda = 0.117$;

(c) T=3 s, A = 0.0068 m, $h_0/\lambda = 0.074$;

having denoted by T the period and by λ the wavelength. For all the cases, the harmonic analysis of free surface elevation measurements taken along the tank centerline are available, and are used to verify the capabilities of a model to reproduce nonlinear refraction and diffraction.

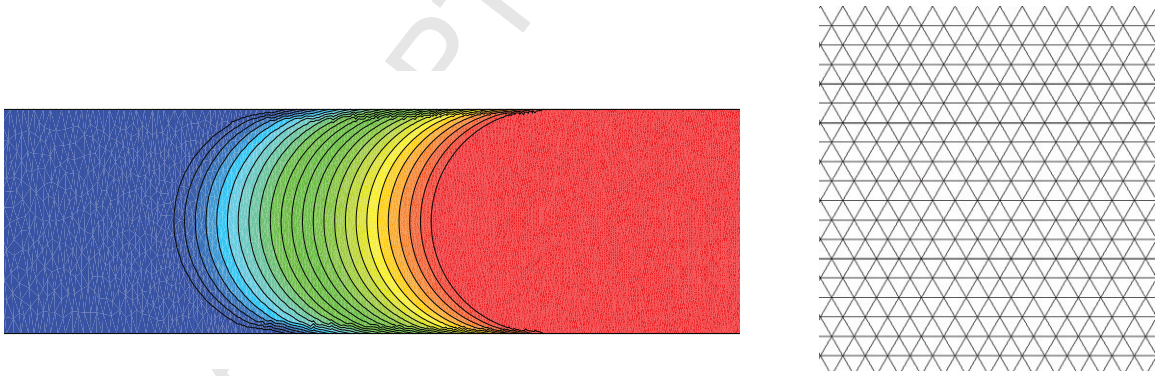

Figure 20: Wave diffraction over a semi-circular shoal. Left : bathymetry contours. Right: close up view of the structured grid

The computational domain is the rectangle $[-10, 36]m \times [0, 6.096]m$. Periodic waves are generated by means of the internal generator described in section §6.3, centered at $x = -2$m. Sponge layers (cf. section §6.1) of a length of 5m are set at the left and right ends of the domain, while reflective boundary conditions are imposed along the top and bottom boundaries. Following [72, 46], we run the tests on a regular triangulation (cf. right picture on figure 20)

of size approximately h ≈ 0.1, in the x direction. The mesh size in the y direction is instead approximately 0.2m. The mesh contains 64470 triangles. The time step is set to $\Delta t = T/100$.

For all the cases, we compare the solution obtained with the SUPG and LDA schemes with the experimental data. In particular, we report in figures 21, 22, and 23 a three dimensional visualization of the wave patterns obtained (top pictures), a comparison of the centerline data obtained with the SUPG and LDA schemes (bottom left picture), and between comparison of the harmonic components of the $\eta(t)$ along the channel centerline obtained with our model with the experiments. In particular, in order to make sure that a steady periodic state is obtained, we start sampling the solution after 15 periods of oscillations. and then perform a DFT of the centerline data collected over 15 periods.

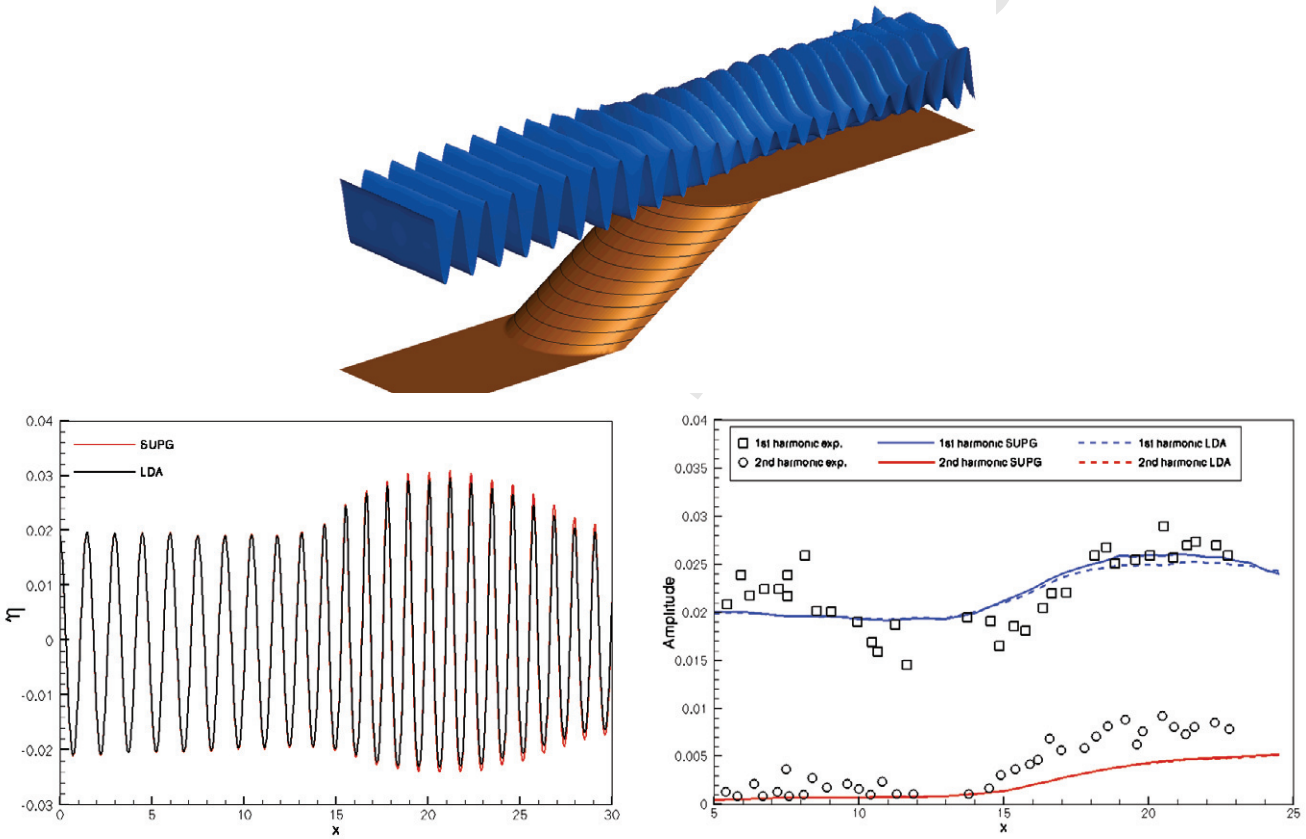

Figure 21: Wave diffraction over a semi-circular, case (a). Top : exaggerated 3D view of the free surface. Bottom-left : centerline data comparison between the LDA and SUPG solutions at $t \approx 79.3\,[s]$. Bottom-right : comparison of DFT of computed centerline data with experiments.

Looking at the results we can draw the following conclusions. Quite consistently, the SUPG gives a stronger shoaling than the LDA scheme. This is particularly visible in the first harmonics in the cases (a) and (b), especially in the first one. The SUPG also show a richer content in higher harmonics in case (b). This is probably a sign of a higher numerical dissipation introduced by the multidimensional upwind stabilization of the LDA scheme. In case (c) the 2 schemes give otherwise nearly identical results, the SUPG still giving a stronger shoaling which is however barely visible in the harmonic comparison.

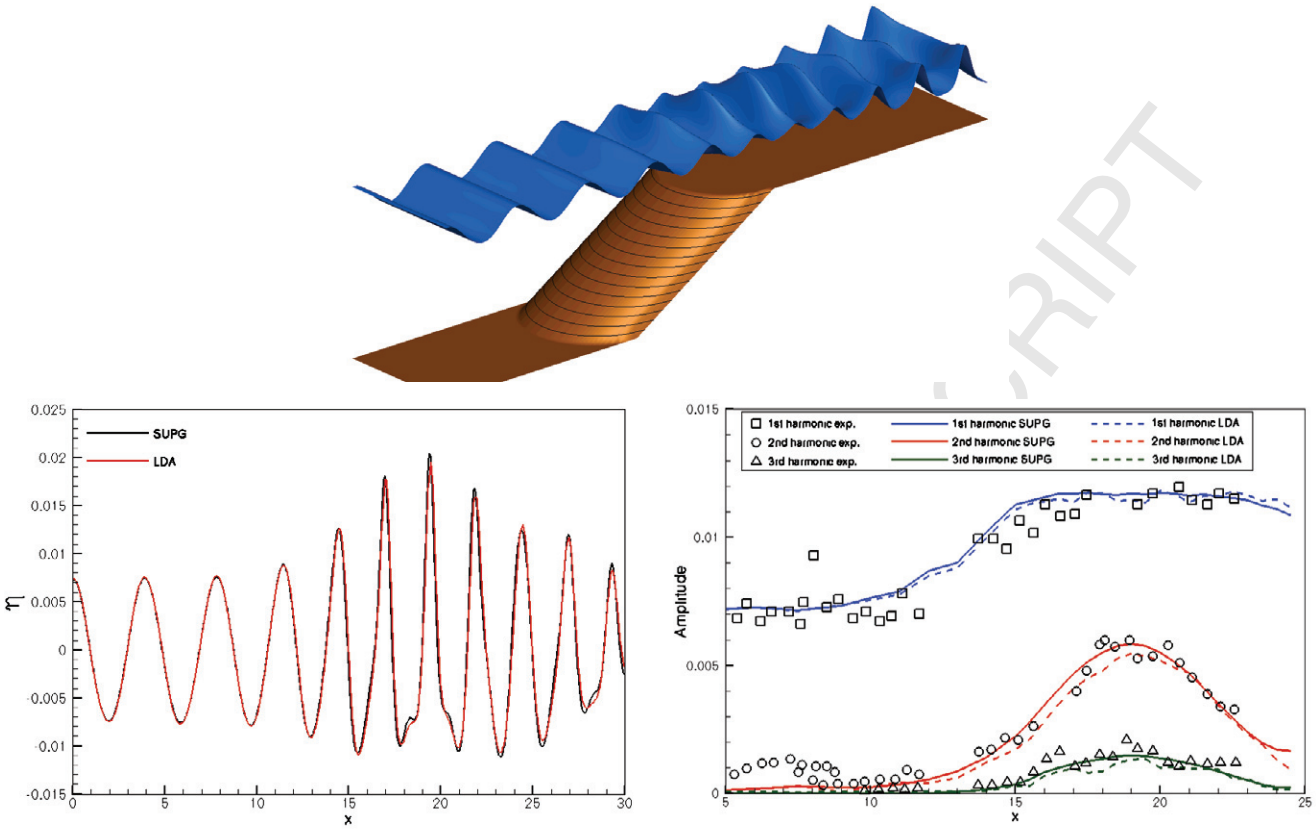

Figure 22: Wave diffraction over a semi-circular, case (b). Top : exaggerated 3D view of the free surface. Bottom-left : centerline data comparison between the LDA and SUPG solutions at $t \approx 75.55\,[s]$. Bottom-right : comparison of DFT of computed centerline data with experiments.

Concerning the comparison with the experiments, the results reported are similar to those presented by others (cf. [55, 12, 72, 54, 31, 68, 46] for example). In particular, case (c), which we consider as being resolved concerning both the spatial and temporal scales, we have an overestimation of the first harmonic which is common to many computational results presented in literature, while both the second and third harmonics are underestimated, as also is quite common in literature. In case (b), which we also consider well resolved, we have a good match between the harmonic content of the computed signal and the experiments. Finally, in case (a), we have a good estimation of the first harmonic, while we have a considerable underestimation of the second. This might be related to an under-resolution problem, as observed e.g. in [31] where a denser mesh was used for this case with a second order DG method, but also related to the "low" order of the time integration scheme.

Overall we judge these results very encouraging. In particular we believe that they confirm our observation in one space dimension.

For completeness, we also report a comparison of the SUPG results obtained on the regular triangulation with those obtained on an unstructured grid. This second mesh is an unstructured triangulation with $h_x \approx 0.2$m at the left and right boundaries. As shown on the right in figure 24, h_x is progressively reduced when approaching the shoal from the left end boundary

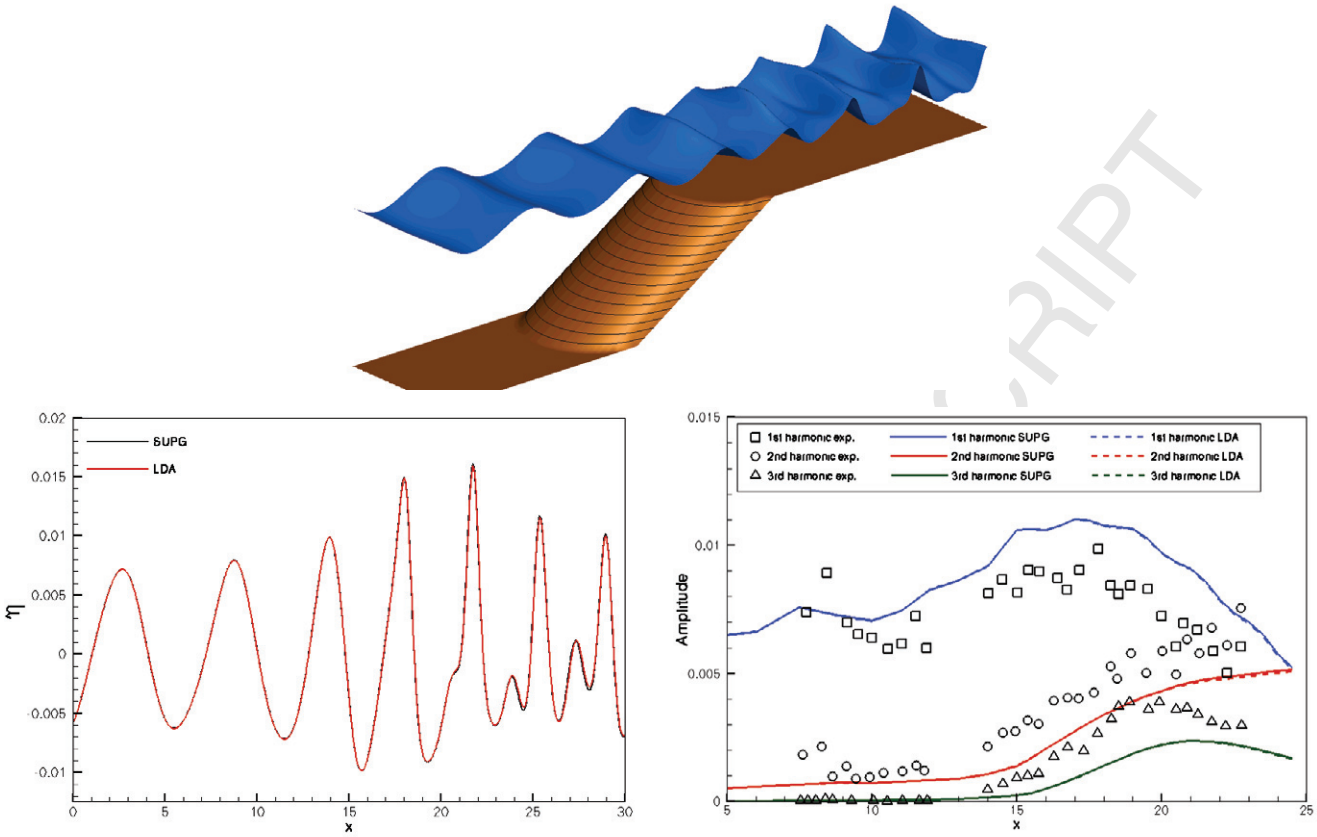

Figure 23: Wave diffraction over a semi-circular, case (c). Top : exaggerated 3D view of the free surface. Bottom-left : centerline data comparison between the LDA and SUPG solutions at $t \approx 120\,[s]$. Bottom-right : comparison of DFT of computed centerline data with experiments.

to reach the value o $h_x \approx 0.1$m before the end of the shoal. As for the regular mesh, also this one has a characteristic size in the y direction twice as large as that in the x direction. The adapted mesh contains 30705 (less than half the number of elements of the regular mesh).

The results are summarized on figure 25 in terms of centerline distribution of the free surface, and harmonic content of the computed signal. It is clear that the unstructured mesh result matches quite well the regular grid one. There are noticeable differences in the harmonic content. In particular, toward the end of the shoal the unstructured grid computation shows a higher content in the second harmonic, which we are still unable to explain. Both solutions, however, match quite well the experimental data.

9.2 Wave diffraction over an elliptic shoal

This test reproduces the experiment of [14] studying the refraction and diffraction of monochromatic waves over a complex bathymetry. This is a standard test to verify models based on the mild-slope equations but it is often also used as a test for extended Boussinesq models (see *e.g.* and [68, 72, 71, 34] references therein). A sketch of the experiment is reported on the left picture on figure 26. The actual wave tank is 20m wide and 22m long. The bathymetry consists of an elliptic shoal mounted on a ramp of constant slope, forming a 20° angle with the

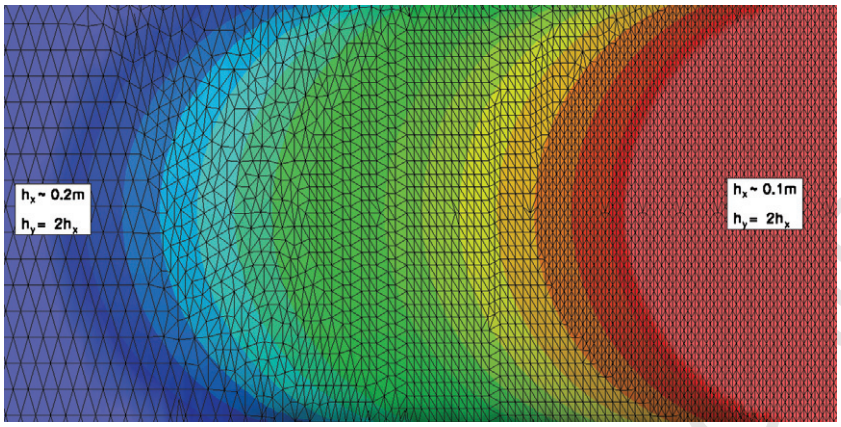

Figure 24: Wave diffraction over a semi-circular : close-up of the unstructured grid.

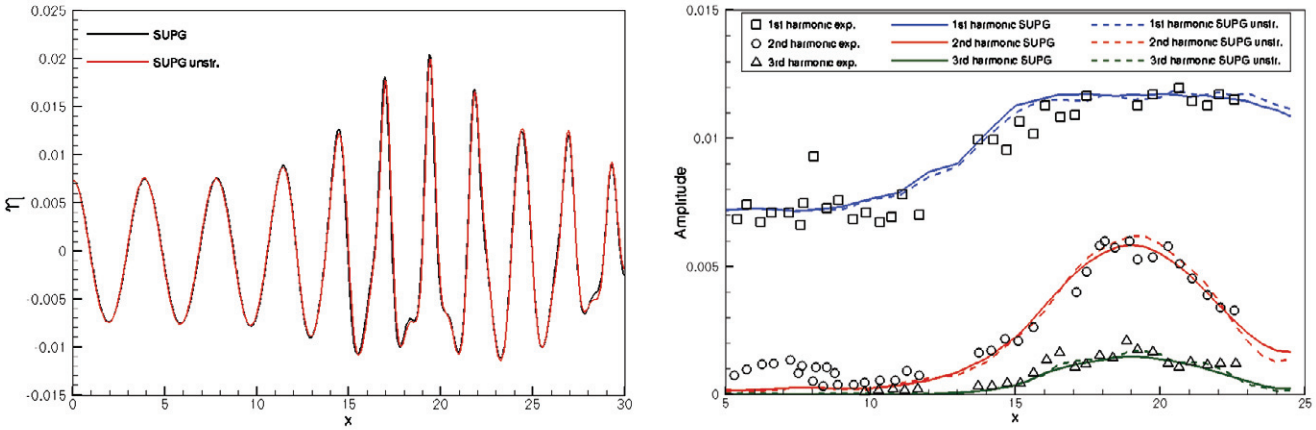

Figure 25: Wave diffraction over a semi-circular, comparison of SUPG results on structured and unstructured triangulation. Left : centerline data at $t \approx 75.55\,[s]$. Right : DFT of the time dependent data along the centerline.

x axis. The maximum water depth is $h_0 = 0.45$m at the wave maker, while the bathymetry is given by the formula $z = z_0 + z_s$, where

$$z_0 = \begin{cases} (5.82 + y_r)/50 & \text{if} \quad y_r \leq -5.82 \\ 0 & \text{otherwise} \end{cases}$$

$$z_s = \begin{cases} -0.3 + \dfrac{1}{2}\sqrt{1 - \left(\dfrac{x_r}{5}\right)^2 - \left(\dfrac{4y_r}{15}\right)^2} & \text{if} \quad \left(\dfrac{x_r}{4}\right)^2 + \left(\dfrac{y_r}{3}\right)^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where the transformed coordinates $(x_r,\, y_r)$ are defined as

$$x_r = x\cos(20^\circ) - y\sin(20^\circ)\,, \quad y_r = x\sin(20^\circ) + y\cos(20^\circ)$$

The water depth is computed as $h = h_0 - z$. The incoming periodic wave has period $T = 1$s, and amplitude $A = 0.0232$m. In [14], the wave elevation was measured in 8 different sections

along which the normalized time average wave height distribution has been computed. A sketch of the problem with the indicative position of the measurement sections is reported in the left picture on figure 26 (cf. [14] for details).

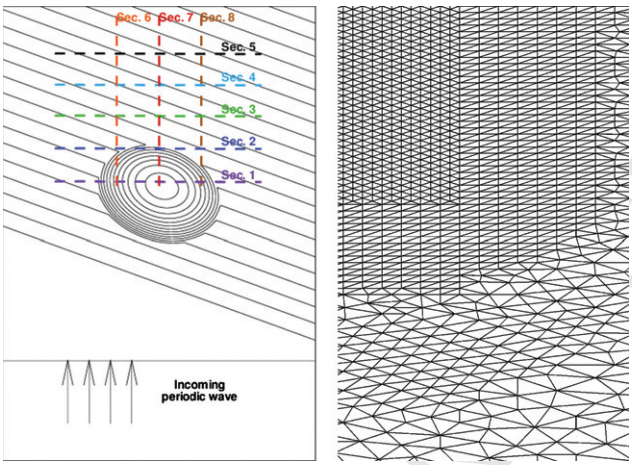

Figure 26: Wave diffraction over an elliptic shoal. Left : sketch of the problem with bathymetry contours and position of the experimental sections. Right : close up view of the mesh.

The computational domain is the square $[-10,\ 10]m \times [-17,\ 15]m$. The incoming periodic wave is obtained by means of the internal wave generator described in section §6.3, centered at $y = -13$m, while sponge layers of 4m thickness are placed at the bottom and top ends of the domain. As in [68], reflective boundary conditions are imposed on the left and right boundaries. Concerning the mesh size, we have run this case on an unstructured grid. The grid has been refined around the region of the interaction. In particular, the grid size in the y direction varies from $\mathrm{h}_y \approx 0.1$ on the bottom and top boundaries to $\mathrm{h}_y \approx 0.05$ in the region around the shoal where the interaction takes place (a grid size of 0.1 m is used in [72], while a mesh size of 0.03m in the main propagation direction is used in [68]). In the x direction, the grid size is $\mathrm{h}_x = 2\mathrm{h}_y$. A close up view of the mesh refinement in the bottom left region of the computational domain is reported in the right picture on figure 26. Computations have been run until time $t = 50$s with $\Delta t \approx 0.02$s.

Three dimensional visualizations of the water elevation computed at time $t = 50$s are reported on figure 27. The left picture shows the region of the interaction with the mesh superimposed. Both pictures provide a visualization of the complex pattern obtained from the diffraction of the incoming periodic wave on the elliptic shoal. To compare our results with the experiments of [14] time dependent data have been extracted from $t = 25$s to $t = 50$s from the sections (cf. leftmost picture on figure 26 and [14]) : section $1 \equiv \{y = 1m| - 5m \leq x \leq 5m\}$ section $2 \equiv \{y = 3m| - 5m \leq x \leq 5m\}$, section $3 \equiv \{y = 5m| - 5m \leq x \leq 5m\}$, section $4 \equiv \{y = 7m| - 5m \leq x \leq 5m\}$,, section $5 \equiv \{y = 9m| - 5m \leq x \leq 5m\}$, section $6 \equiv \{x = -2m|0m \leq y \leq 10m\}$, section $7 \equiv \{x = 0m|0m \leq y \leq 10m\}$, and section $8 \equiv \{x = 2m|0m \leq y \leq 10m\}$. The data obtained has been analyzed using the *zero up-crossing* principle to isolate single waves and compute the average wave height distributions, the wave height of a single wave being defined as $\eta_{\max} - \eta_{\min}$. The results, normalized by the incoming wave amplitude $A = 0.0232$m, are reported on figure 28.

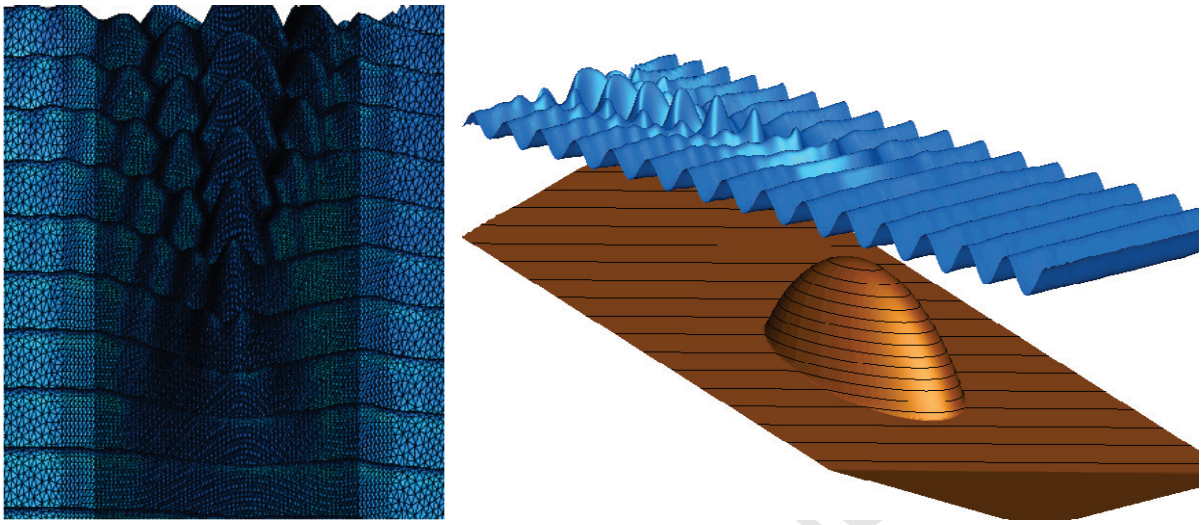

Figure 27: Wave diffraction over an elliptic shoal, results at $t = 50$s. Left : top view of the free surface with mesh. Right : exaggerated 3D view of the free surface with bathymetry.

The agreement of our computational results with the experiments is very satisfactory, and compares very favorably with what can be found in published literature (see *e.g.* [68, 72] and references therein). Once more the SUPG scheme gives a slightly stronger shoaling, as it is visible clearly in sections 2, 5, and 7. Again we assume this behavior to be related to a less dissipative character of the streamline upwind stabilization term w.r.t. the multidimensional upwind one.

10 Conclusions and future work

In this paper we have presented an investigation of the applicability of low order upwind stabilized continuous finite elements and residual distribution schemes for the discretization of the enhanced Boussinesq equations of [55, 65] to simulate wave propagation on complex bathymetries. In the first half of the paper, we have thoroughly studied the properties of the standard continuous Galerkin scheme, of its upwind stabilized variant, the SUPG scheme, and of the centered and Upwind Residual Distribution schemes. In both the finite element and residual distribution frameworks upwinding is performed along the characteristics of the nonlinear shallow water equations. The theoretical analysis of the truncation and dispersion error of the schemes, has shown that both the Galerkin and SUPG schemes provide an accuracy comparable or even better than those of the standard finite difference scheme of Wei and Kirby [34]. These observation have been largely confirmed by the extensive numerical validation presented in one space dimension.

Extensions to the two-dimensional case of the SUPG analyzed in 1d have then been studied. In particular, we have considered two forms of the upwind Petrov-Gelerkin stabilization. The first is that of the standard SUPG scheme [43], while the second is the one leading to the Multidimensional Upwind LDA residual distribution scheme, formulated in a Petrov-Galerkin context, as suggested e.g. in [32, 3, 59]. These schemes have been tested and compared on standard two-dimensional benchmarks. The results show that, on meshes similar to those

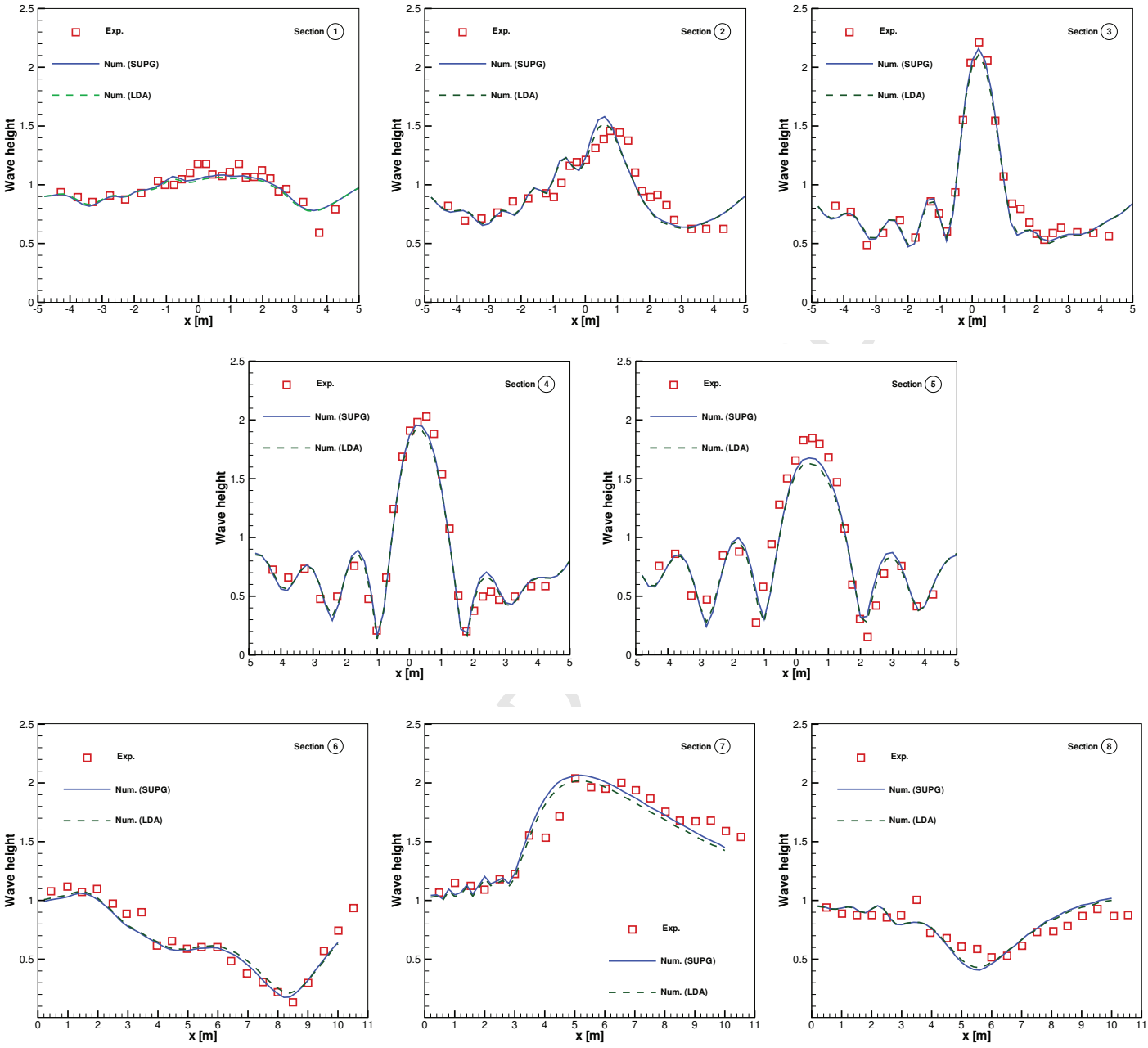

Figure 28: Wave diffraction over an elliptic shoal. Comparison of the computed average wave height with the experimental data of [14] in sections 1 to 8.

used in the reference literature, these schemes provide results comparable to those presented by other authors.

Concerning the computational cost, we have reported a detailed discussion of the unit costs (per node, time step, and Newton iteration, and per node and time step) of all the schemes considered. We have shown that, while in our (perhaps inefficient) one dimensional Matlab®

implementation the upwind stabilized Galerkin SUPG scheme requires 70% more CPU time than the scheme of Wei and Kirby, and 60% more than the "unstabilized" Galerkin scheme, in our improved two-dimensional implementation the overhead of the stabilized schemes w.r.t. the Galerkin scheme is only of 20%, which we claim being a much more realistic figure. Moreover, we have also shown that the frozen Newton approach we propose allows gains up to 75% of the CPU time, if the LU decomposition of the Jacobian is also stored and frozen. In particular, since we are anyways using a direct solver, these gains in CPU time do not come at any additional memory requirements, since the memory used is that normally already required by the linear solver. Note also, that these requirements might be more demanding in the P^2 and P^3 cases, however will remain much more moderate for a C^0 interpolation than for a discontinuous one, and might anyways be mitigated by the need of a lower number of elements, thanks to the increase of accuracy.

In conclusion, we believe we have given very strong evidence that low order upwind stabilized continuous Galerkin schemes can be applied to the discretization of enhanced Boussinesq equations with results comparable to those presented in the published literature, and definitely comparable to those obtained with the basic "unstabilized" Galerkin scheme. This justifies the continuation of our work in several future directions :

- the use of a hybrid approach to handle wave breaking by reverting to the nonlinear Shallow Water equations in breaking regions. This development, which constitutes the motivation behind the present paper, has already been successfully tested in [68, 69, 16, 47]. Aside from the upwind stabilization allowing to stably handle the hyperbolic Shallow Water equations, additional care will have to be taken in resolving monotonically the bores obtained in the NLSW region. Our preliminary one dimensional implementation shows that mass lumping alone allows to locally recover a monotonicity preserving scheme in 1d. In the multidimensional case, this will not be enough, and will have to be coupled with other shock capturing techniques, such as those proposed in [61], or on more standard finite element discontinuity capturing operators [39, 42].
- we are currently extending the study presented here to other models, including the enhanced Boussinesq equations of Nwogu [53, 72, 31], and genuinely nonlinear dispersive model such as the Green-Naghdi equations proposed in [49, 16]. Our preliminary results show that for the equations of Nwogu we obtain a very similar behavior to that demonstrated here for the model of Madsen and Sørensen. In particular, the SUPG scheme has remarkably low dispersion errors. These results will be reported elsewhere ;
- even though the numerical results shown here are very satisfactory, we believe that a higher order implicit time stepping scheme should be employed, to guarantee a better resolution of high frequency content in the solution, while allowing to choose the time step without being constrained by e.g. geometrical stiffness related to local mesh adaptation. The form f the Boussinesq equations used in conjunction with such techniques will have to be properly studied, as suggested in [28] ;
- even though this paper has discussed results concerning the P^1 case, in practice the use of higher order approximations might provide an additional gain in efficiency, in terms of CPU time per node and per time step, for a given level of resolution (error). This is also the topic of our current work, which, in particular, focuses on the study of both hierarchical approaches [75, 7, 30], and of Bezier-type approximations [41, 6]

- concerning the efficiency, a comparison between the MUMPS based implementation discussed here and a Newton algorithm using an iterative linear solver will be performed, and, of course, the optimal parallel implementation of the schemes on heterogeneous architectures will have to be considered.

Acknowledgments

The authors are indebted to Mathieu Colin and Stevan Bellec (Inria Bordeaux Sud-Ouest) for providing a SSP-RK integrator to compute improved solitary wave solutions. Argiris Delis and Maria Kazolea (Technical University of Crete) are also warmly thanked for all the discussions concerning internal wave generation. We are grateful to the anonymous reviewers who helped so much improving the paper.

A FD equations and dispersion parameters

We report here the finite difference expressions of all the schemes analyzed in section §4 when applied to the linearized equations (2), including their discrete dispersion parameters. All the derivation is done by assuming an equally spaced mesh of size h $= \Delta x$, $C_0^2 = gh_0$ denotes the square of the linearized Shallow Water celerity, while the notation $\mu_{\Delta x} = k\Delta x$ has been used in the expressions of the dispersion parameters.

A.1 Second order central finite differences

The second order centered finite difference scheme applied to system (2) yields

$$\begin{aligned}
&\frac{d\eta_i}{dt} + \frac{h_0}{2\Delta x}\left(u_{i+1} - u_{i-1}\right) = 0 \\
&\frac{du_i}{dt} - \frac{Bh_0^2}{\Delta x^2}\left(\frac{du_{i-1}}{dt} - 2\frac{du_i}{dt} + \frac{du_{i+1}}{dt}\right) \\
&+ \frac{g}{2\Delta x}\left(\eta_{i+1} - \eta_{i-1}\right) - \beta\frac{gh_0^2}{2\Delta x^3}\left(-\eta_{i-2} + 2\eta_{i-1} - 2\eta_{i+1} + \eta_{i+2}\right) = 0
\end{aligned} \tag{50}$$

Replacing in the scheme the time dependent nodal values by a Fourier mode in space, we obtain for the discrete dispersion parameters the following expressions :

$$\xi_{\mathrm{FD2}} = 0$$

$$\omega_{\mathrm{FD2}}^2 = \frac{N_{\mathrm{FD2}}}{D_{\mathrm{FD2}}}$$

where

$$N_{\mathrm{FD2}} = C_0^2 k^2 \frac{\sin^2 \mu_{\Delta x}}{\mu_{\Delta x}^2} + 2\beta g h_0^3 k^4 \frac{\sin^2 \mu_{\Delta x}}{\mu_{\Delta x}^4}(1 - \cos \mu_{\Delta x})$$

$$D_{\mathrm{FD2}} = 1 + 2B\frac{h_0^2 k^2}{\mu_{\Delta x}^2}(1 - \cos \mu_{\Delta x})$$

A.2 Central finite difference scheme by Wei and Kirby

Applying the spatial discretization described in [34] to the linearized system of Madsen and Sørensen enhanced Boussinesq equations (2) yields

$$\begin{aligned}
&\frac{d\eta_i}{dt} + \frac{h_0}{12\Delta x}\left(u_{i-2} - 8u_{i-1} + 8u_{i+1} - u_{i+2}\right) = 0 \\
&\frac{du_i}{dt} - \frac{Bh_0^2}{\Delta x^2}\left(\frac{du_{i-1}}{dt} - 2\frac{du_i}{dt} + \frac{du_{i+1}}{dt}\right) \\
&\quad + \frac{g}{12\Delta x}\left(\eta_{i-2} - 8\eta_{i-1} + 8\eta_{i+1} - \eta_{i+2}\right) \\
&\quad - \beta\frac{gh_0^2}{2\Delta x^3}\left(-\eta_{i-2} + 2\eta_{i-1} - 2\eta_{i+1} + \eta_{i+2}\right) = 0
\end{aligned} \tag{51}$$

Replacing in the scheme the time dependent nodal values by a Fourier mode in space, we obtain for the discrete dispersion parameters the following expressions :

$$\xi_{\mathrm{FDWK}} = 0$$

$$\omega^2_{\mathrm{FDWK}} = \frac{N_{\mathrm{FDWK}}}{D_{\mathrm{FDWK}}}$$

where

$$N_{\mathrm{FDWK}} = C_0^2 k^2 \frac{\sin \mu_{\Delta x}{}^2}{\mu^2_{\Delta x}}(4 - \cos\mu_{\Delta x})\left(\frac{1}{9}(4 - \cos\mu_{\Delta x}) + \frac{2}{3}\frac{\beta h_0^2 k^2}{\mu^2_{\Delta x}}(1 - \cos\mu_{\Delta x})\right)$$

$$D_{\mathrm{FDWK}} = 1 + 2B\frac{h_0^2 k^2}{\mu^2_{\Delta x}}(1 - \cos\mu_{\Delta x})$$

A.3 Continuous Galerkin scheme

When applied to system (2) on a uniform mesh with definition (22) of the auxiliary variable w_i^η yielding

$$w_i^\eta = \frac{\eta_{i+1} - 2\eta_i + \eta_{i-1}}{\Delta x^2}$$

the continuous Galerkin scheme (7) reduces to

$$\begin{aligned}
&\frac{\Delta x}{6}\frac{d\eta_{i-1}}{dt} + \frac{2\Delta x}{3}\frac{d\eta_i}{dt} + \frac{\Delta x}{6}\frac{d\eta_{i+1}}{dt} + \frac{h_0}{2}\left(u_{i+1} - u_{i-1}\right) = 0 \\
&\frac{\Delta x}{6}\frac{du_{i-1}}{dt} + \frac{2\Delta x}{3}\frac{du_i}{dt} + \frac{\Delta x}{6}\frac{du_{i+1}}{dt} + \frac{g}{2}\left(\eta_{i+1} - \eta_{i-1}\right) - \frac{Bh_0^2}{\Delta x}\left(\frac{du_{i-1}}{dt} - 2\frac{du_i}{dt} + \frac{du_{i+1}}{dt}\right) \\
&\quad - \beta\frac{gh_0^2}{2\Delta x^2}\left(-\eta_{i-2} + 2\eta_{i-1} - 2\eta_{i+1} + \eta_{i+2}\right) = 0
\end{aligned} \tag{52}$$

Note that the only difference w.r.t the second order finite difference scheme is the presence of a non-diagonal mass matrix multiplying the $\partial_t\eta$ and $\partial_t u$ terms, otherwise the two schemes are identical.

Replacing in the scheme the time dependent nodal values by a Fourier mode in space, we obtain for the discrete dispersion parameters the following expressions :

$$\xi_{\mathrm{cG}} = 0$$

$$\omega_{\mathrm{cG}}^2 = C_0^2 k^2 \frac{\sin^2(\mu_{\Delta x})}{\mu^2} \frac{1 + 2\beta\mu^2 \dfrac{1-\cos(\mu_{\Delta x})}{\mu_{\Delta x}^2}}{\dfrac{2+\cos(\mu_{\Delta x})}{3}\left(\dfrac{2+\cos(\mu_{\Delta x})}{3} + 2B\mu^2 \dfrac{1-\cos(\mu_{\Delta x})}{\mu_{\Delta x}^2}\right)}$$

A.4 Centered residual distribution

When applied to system (2), definitions (22) of w^η yields

$$w_i^\eta = \frac{\eta_{i+1} - 2\eta_i + \eta_{i-1}}{\Delta x^2} \quad \text{and}$$

while w^q is replaced by a $w^u \approx \partial_x u$ similarly defined, which under the present hypotheses is computed as

$$w_i^u = \frac{u_{i+1} - u_{i-1}}{2\Delta x}$$

the centered residual distribution scheme reduces to

$$\frac{\Delta x}{4}\frac{d\eta_{i-1}}{dt} + \frac{\Delta x}{2}\frac{d\eta_i}{dt} + \frac{\Delta x}{4}\frac{d\eta_{i+1}}{dt} + \frac{h_0}{2}(u_{i+1} - u_{i-1}) = 0$$

$$\frac{\Delta x}{4}\frac{du_{i-1}}{dt} + \frac{\Delta x}{2}\frac{du_i}{dt} + \frac{\Delta x}{4}\frac{du_{i+1}}{dt} + \frac{g}{2}(\eta_{i+1} - \eta_{i-1}) - \frac{Bh_0^2}{2\Delta x}\left(\frac{du_{i-2}}{dt} - 2\frac{du_i}{dt} + \frac{du_{i+2}}{dt}\right) \tag{53}$$

$$- \beta\frac{gh_0^2}{2\Delta x^2}(-\eta_{i-2} + 2\eta_{i-1} - 2\eta_{i+1} + \eta_{i+2}) = 0$$

Note that the difference w.r.t the second order finite difference scheme is the presence of a non-diagonal mass matrix multiplying the $\partial_t\eta$ and $\partial_t u$ terms, and the fact that the second order derivative of u in space is approximated with a larger stencil.

Replacing the time dependent nodal values by a Fourier mode in space, we obtain for the discrete dispersion parameters the following expressions :

$$\xi_{\mathrm{cRD}} = 0$$

$$\omega_{\mathrm{cRD}}^2 = C_0^2 k^2 \frac{\sin^2(\mu_{\Delta x})}{\mu^2} \frac{1 + 2\beta\mu^2 \dfrac{1-\cos(\mu_{\Delta x})}{\mu_{\Delta x}^2}}{\dfrac{1+\cos(\mu_{\Delta x})}{2}\left(\dfrac{1+\cos(\mu_{\Delta x})}{2} + \dfrac{B\mu^2}{2} \dfrac{1-\cos(2\mu_{\Delta x})}{\mu_{\Delta x}^2}\right)}$$

A.5 SUPG scheme

In the linearized case, the matrix associated to the Shallow Water quasi-linear form is

$$A_0 = \begin{bmatrix} 0 & h_0 \\ g & 0 \end{bmatrix}$$

Straightforward eigenvalue decomposition show that the sign of this matrix can be written as

$$\text{sign}(A_0) = \begin{bmatrix} 0 & C_0/g \\ g/C_0 & 0 \end{bmatrix}$$

Using this expression, and the pointwise expressions of the auxiliary variables recalled above, we can deduce the form of the SUPG scheme (18) applied to system (2) :

$$\begin{aligned}
&\frac{\Delta x}{6}\frac{d\eta_{i-1}}{dt} + \frac{2\Delta x}{3}\frac{d\eta_i}{dt} + \frac{\Delta x}{6}\frac{d\eta_{i+1}}{dt} + \frac{h_0}{2}\left(u_{i+1} - u_{i-1}\right) - \frac{C_0}{2g}\left\{\frac{\Delta x}{2}\left(\frac{du_{i+1}}{dt} - \frac{du_{i-1}}{dt}\right)\right. \\
&\qquad - \frac{Bh_0^2}{\Delta x}\left(\frac{du_{i+2}}{dt} - 2\frac{du_{i+1}}{dt} + 2\frac{du_{i-1}}{dt} - \frac{du_{i-2}}{dt}\right) - g(\eta_{i+1} - 2\eta_i + \eta_{i-1}) \\
&\qquad\qquad \left. - \frac{\beta g h_0^2}{\Delta x^2}(\eta_{i+2} - 4\eta_{i+1} + 6\eta_i - 4\eta_{i-1} + \eta_{i-2})\right\} = 0 \\
&\frac{\Delta x}{6}\frac{du_{i-1}}{dt} + \frac{2\Delta x}{3}\frac{du_i}{dt} + \frac{\Delta x}{6}\frac{du_{i+1}}{dt} + \frac{g}{2}\left(\eta_{i+1} - \eta_{i-1}\right) - \frac{Bh_0^2}{\Delta x}\left(\frac{du_{i-1}}{dt} - 2\frac{du_i}{dt} + \frac{du_{i+1}}{dt}\right) \\
&\qquad - \beta\frac{gh_0^2}{2\Delta x^2}\left(-\eta_{i-2} + 2\eta_{i-1} - 2\eta_{i+1} + \eta_{i+2}\right) \\
&\qquad - \frac{g}{2C_0}\left\{\frac{\Delta x}{2}\left(\frac{d\eta_{i+1}}{dt} - \frac{d\eta_{i-1}}{dt}\right) + h_0(u_{i+1} - 2u_i + u_{i-1})\right\} = 0
\end{aligned} \tag{54}$$

Replacing the time dependent nodal values by a Fourier mode in space, we obtain for the discrete dispersion parameters the following expressions :

$$\xi_{\text{SUPG}} = -\frac{A_\nu}{2A_{\nu^2}}$$
$$\omega^2_{\text{SUPG}} = \frac{A_s}{A_{\nu^2}} - \xi^2$$

where

$$\begin{aligned}
A_{\nu^2} &= \frac{(2+\cos\mu_{\Delta x})}{3}\left(\frac{(2+\cos\mu_{\Delta x})}{3} + 2B\mu^2\frac{(1-\cos(\mu_{\Delta x}))}{\mu^2_{\Delta x}}\right) + \\
&\quad + \frac{\sin^2\mu_{\Delta x}}{4}\left(1 + 2B\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu^2_{\Delta x}}\right) > 0;
\end{aligned}$$

$$A_\nu = C_0 k\frac{(1-\cos\mu_{\Delta x})^2}{\mu_{\Delta x}}\left(\frac{1}{6} + \left(1 + 2\beta\mu^2\frac{1-\cos\mu_{\Delta x}}{\mu^2_{\Delta x}}\right)\left(\frac{1}{6} + B\mu^2\frac{1-\cos\mu_{\Delta x}}{\mu^2_{\Delta x}}\right)\right) \geq 0;$$

$$A_s = C_0^2k^2\frac{(1-\cos\mu_{\Delta x})^2}{\mu^2_{\Delta x}}\left(1 + 2\beta\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu_{\Delta x^2}}\right) + \ C_0^2k^2\frac{\sin^2\mu_{\Delta x}}{\mu^2_{\Delta x}}\left(1 + 2B\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu^2_{\Delta x}}\right)$$

Note that from the inequalities $A_{\nu^2} > 0$ and $A_\nu \geq 0$ we immediately deduce $\xi_{\text{SUPG}} \leq 0$ which shows the dissipative character of the method.

A.6 Upwind residual distribution

Using the expressions of the sign of the Shallow Water Jacobian in the linearized case, and the expressions of the auxiliary variables recalled above, the upwind RD scheme applied to the linearized system (2) becomes

$$\begin{aligned}
&\frac{\Delta x}{4}\frac{d\eta_{i-1}}{dt}+\frac{\Delta x}{2}\frac{d\eta_i}{dt}+\frac{\Delta x}{4}\frac{d\eta_{i+1}}{dt}+\frac{h_0}{2}\left(u_{i+1}-u_{i-1}\right)-\frac{C_0}{2g}\left\{\frac{\Delta x}{2}\left(\frac{du_{i+1}}{dt}-\frac{du_{i-1}}{dt}\right)\right.\\
&\quad-\frac{Bh_0^2}{\Delta x}\left(\frac{du_{i+2}}{dt}-2\frac{du_{i+1}}{dt}+2\frac{du_{i-1}}{dt}-\frac{du_{i-2}}{dt}\right)-g(\eta_{i+1}-2\eta_i+\eta_{i-1})\\
&\quad\left.-\frac{\beta g h_0^2}{\Delta x^2}(\eta_{i+2}-4\eta_{i+1}+6\eta_i-4\eta_{i-1}+\eta_{i-2})\right\}=0\\
&\frac{\Delta x}{4}\frac{du_{i-1}}{dt}+\frac{\Delta x}{2}\frac{du_i}{dt}+\frac{\Delta x}{4}\frac{du_{i+1}}{dt}+\frac{g}{2}\left(\eta_{i+1}-\eta_{i-1}\right)-\frac{Bh_0^2}{2\Delta x}\left(\frac{du_{i-2}}{dt}-2\frac{du_i}{dt}+\frac{du_{i+2}}{dt}\right)\\
&\quad-\beta\frac{gh_0^2}{2\Delta x^2}\left(-\eta_{i-2}+2\eta_{i-1}-2\eta_{i+1}+\eta_{i+2}\right)\\
&\quad-\frac{g}{2C_0}\left\{\frac{\Delta x}{2}\left(\frac{d\eta_{i+1}}{dt}-\frac{d\eta_{i-1}}{dt}\right)+h_0(u_{i+1}-2u_i+u_{i-1})\right\}=0
\end{aligned}\tag{55}$$

Replacing the time dependent nodal values by a Fourier mode in space, we obtain after long calculations :

$$\xi_{\mathrm{URD}}=0$$

$$\omega^2_{\mathrm{URD}}=\frac{A_s}{A_{\nu^2}}$$

where

$$\begin{aligned}
A_{\nu^2}&=\frac{(1+\cos\mu_{\Delta x})}{2}\left(\frac{(1+\cos\mu_{\Delta x})}{2}+B\mu^2\frac{(1-\cos(2\mu_{\Delta x}))}{2\mu^2_{\Delta x}}\right)+\\
&\quad+\frac{\sin^2\mu_{\Delta x}}{4}\left(1+2B\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu^2_{\Delta x}}\right);\\
A_s&=C_0^2k^2\frac{(1-\cos\mu_{\Delta x})^2}{\mu^2_{\Delta x}}\left(1+2\beta\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu_{\Delta x^2}}\right)+\\
&\quad+C_0^2k^2\frac{\sin^2\mu_{\Delta x}}{\mu^2_{\Delta x}}\left(1+2B\mu^2\frac{(1-\cos\mu_{\Delta x})}{\mu^2_{\Delta x}}\right).
\end{aligned}$$

References

[1] R. Abgrall. Essentially non oscillatory residual distribution schemes for hyperbolic problems. *J. Comput. Phys*, 214(2):773–808, 2006.

[2] R. Abgrall, G. Baurin, A. Krust, D. de Santis, and M. Ricchiuto. Numerical approximation of parabolic problems by residual distribution schemes. *International Journal for Numerical Methods in Fluids*, 71(9):1191–1206, 2013.

[3] R. Abgrall and M. Mezine. Construction of second-order accurate monotone and stable residual distribution schemes for unsteady flow problems. *J. Comput. Phys.*, 188:16–55, 2003.

[4] R. Abgrall and P.L. Roe. High-order fluctuation schemes on triangular meshes. *J. Sci. Comput.*, 19(3):3–36, 2003.

[5] R. Abgrall, D. De Santis, and M. Ricchiuto. Construction of a high order Residual Distribution scheme for complex viscous flows. In *Seventh International Conference on Computational Fluid Dynamics (ICCFD7)*, Big Island, Hawaii, USA, July 2012.

[6] R. Abgrall and J. Trefilik. An example of high order residual distribution scheme using non-lagrange elements. *Journal of Scientific Computing*, 45:3–25, 2010.

[7] S. Adjerid, M. Aiffa, and J.E. Flaherty. Hierarchical finite element bases for triangular and tetrahedral elements. *Computer Methods in Applied Mechanics and Engineering*, 190:2925 – 2941, 2001.

[8] P. Amestoy, I.S. Duff, J.-Y. L'Excellent, and J. Koster. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM J. Matrix Anal. Appl.*, 23(1):15–41, January 2001.

[9] P. Amestoy, A. Guermouche, J.-Y. L'Excellent, and S. Pralet. Hybrid scheduling for the parallel solution of linear systems. *Parallel Computing*, 32(2):136–156, 2006.

[10] W. El Asmar and O. Nwogu. Finite volume solution of boussinesq-type euations on an unstructured grid. in: J. McKee Smith (ed.), Proc. of the 30$^{\text{th}}$ Int. Conf. on Coastal Engineering, San Diego, CA, 2006.

[11] T.J. Barth. Numerical methods for gasdynamic systems on unstructured meshes. In Kröner, Ohlberger, and Rohde, editors, *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws*, volume 5 of *Lecture Notes in Computational Science and Engineering*, pages 195–285. Springer-Verlag, Heidelberg, 1998.

[12] S. Beji and K. Nadaoka. A formal derivation and numerical modeling of the improved boussinesq equations for varying depth. *Ocean Engineering*, 23, 1996.

[13] M.J. Berger and R.J. Leveque. Adaptive mesh refinement using wave-propagation algorithms for hyperbolic systems. *SIAM J. Numer. Anal.*, 35, 1998.

[14] J.C.W. Berkhoff, N. Booy, and A.C. Radder. Verification of numerical wave propagation models for simple harmonic linear water waves. *Coastal Engineering*, 6:255–279, 1982.

[15] P. Bonneton. Modeling of periodic wave transformation in the inner surf zone. *Ocean Engineering*, 34, 2007.

[16] P. Bonneton, F. Chazel, D. Lannes, F. Marche, and M. Tissier. A splitting approach for the fully nonlinear and wakly dispersive green-naghdi model. *Journal of Computational Physics*, 230, 2011.

[17] P. Brufau and P. Garcia-Navarro. Unsteady free surface flow simulation over complex topography with a multidimensional upwind technique. *Journal of Computational Physics*, 186(2):503 – 526, 2003.

[18] E. Burman. Consistent supg-method for transient transport problems: Stability and convergence. *Computer Methods in Applied Mechanics and Engineering*, 199(17-20):1114 – 1123, 2010.

[19] J.-C. Carette, H. Deconinck, H. Paillère, and P.L. Roe. Multidimensional upwinding: its relation to finite elements. *International Journal for Numerical Methods in Fluids*, 20:935–955, 1995.

[20] R. Cienfuegos, E. Barthelemy, and P. Bonneton. A fourth order compact finite volume scheme for fully nonlinear and weakly dispersive boussinesq type equations. parti II: Boundary conditions and validation. *Int. J. Numer. Meth. Fluids*, 53, 2006.

[21] L. Quartapelle D. Ambrosi. A taylor-galerkin method for simulating nonlinear dispersive water waves. *Journal of computational physics*, 146:546–569, 1998.

[22] H. Deconinck and M. Ricchiuto. Residual distribution schemes: foundation and analysis. In E. Stein, R. de Borst, and T.J.R. Hughes, editors, *Encyclopedia of Computational Mechanics*. John Wiley & Sons, Ltd., 2007. DOI: 10.1002/0470091355.ecm054.

[23] H. Deconinck, M. Ricchiuto, and K. Sermeus. Introduction to residual distribution schemes and stabilized finite elements. *VKI LS 2003-05, 33rd Computational Fluid dynamics Course, von Karman Institute for Fluid Dynamics*, 2003.

[24] H. Deconinck, K. Sermeus, and R. Abgrall. Status of multidimensional upwind residual distribution schemes and applications in aeronautics. AIAA paper 2000-2328, June 2000. AIAA CFD Conference, Denver (USA).

[25] D. Dutykh, T. Katsaonis, and D. Mitsotakis. Finite volume schemes for dispersive wave propagation and runup. *Journal of Computational Physics*, 230, 2011.

[26] A.P. Engsig-Karup. Unstructured nodal dg-fem solution of high order boussinesq type equations. PhD Thesis, Technical University of Denmark, 2006.

[27] A.P. Engsig-Karup, J.S. Hesthaven, H.B. Bingham, and P.A. Madsen. Nodal dg-fem sollution of high order boussinesq-type equations. *J. Eng. Math*, 56, 2006.

[28] C. Eskilsson and A.P. Engsig-Karup. On devosing boussinesq-type models with bounded eigenspectra: one horizontal dimension. *Journal of computational physics*, 2013. to appear.

[29] C. Eskilsson and S.J. Sherwin. An hp/spectral element model for efficient long-time integration of boussinesq-type equations. *J. of Coastal Engineering*, 45(2), 2003.

[30] C. Eskilsson and S.J. Sherwin. Spectral/hp discontinuous galerkin methods for modelling 2d boussinsq equations. *Journal of Computational Physics*, 210, 2006.

[31] C. Eskilsson, S.J. Sherwin, and L. Bergdhal. An unstructured spectral/*hp* element model for enhanced boussinesq-type euations. *Coastal Engineering*, 53, 2006.

[32] A. Ferrante and H. Deconinck. Solution of the unsteady Euler equations using residual distribution and flux corrected transport. Technical Report VKI-PR 97-08, von Karman Institute for Fluid Dynamics, 1997.

[33] D.R. Fuhrman and H.B Bingham. Numerical solution of fully nonlinear and highly dispersive boussinesq equations in two horizontal dimensions. *Int. J. Numer. Meth. Fluids*, 44, 2004.

[34] J.T. Kirby G. Wei. A time-dependent numerical code for extended boussinesq equations. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 120:251–261, 1995.

[35] M.F. Gobbi, J.T. Kirby, and G.E. Wei. A fully nonlinear boussinesq model for surface waves. part 2: Extension to o$(kh)^4$. *J.Fluid.Mech*, 405, 2000.

[36] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability preserving high-order time discretization methods. *SIAM review*, 43(1):89–112, 2001.

[37] N. Goutal, M.O. Bristeau, and J. Sainte-Marie. Numerical simulation of a non-hydrostatic shallow water model. *Computers&Fluids*, 47(1), 2011.

[38] Ralf Hartmann and Paul Houston. Error estimation and adaptive mesh refinement for aerodynamic flows. In H. Deconinck, editor, *VKI LS 2010-01: 36th CFD/ADIGMA course on hp-adaptive and hp-multigrid methods, Oct. 26-30, 2009*. Von Karman Institute for Fluid Dynamics, Rhode Saint Genèse, Belgium, 2009.

[39] G. Hauke. A symmetric formulation for computing transient shallow water flows. *Computer Methods in Applied Mechanics and Engineering*, 163(1-4):111–122, 1998.

[40] T.J.R. Hughes and A. Brook. Streamline upwind Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comp. Meth. Appl. Mech. Engrg.*, 32:199–259, 1982.

[41] T.J.R. Hughes, J.A. Cottrell, and Y. Bazilevs. Isogeometric analysis: Cad, finite elements, nurbs, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Eng.*, 194(39-41):4135–4195, 2005.

[42] T.J.R. Hughes and M. Mallet. A new finite element formulation for CFD IV: a discontinuity-capturing operator for multidimensional advective-diffusive systems. *Comp. Meth. Appl. Mech. Engrg.*, 58:329–336, 1986.

[43] T.J.R. Hughes, G. Scovazzi, and T. Tezduyar. Stabilized methods for compressible flows. *J. Sci. Comp.*, 43:343–368, 2010.

[44] P.H. Taylor J. Orszaghova, A.G.L. Borthwick. From the paddle to the beach - a bousinesq shallow water numerical wave tank based on madsen and sørensen's equations. *Journal of Computational Physics, 231, 328-344*, 2012.

[45] M. Kazolea, A.I. Delis, I.K. Nikolos, and C.E. Synolakis. An unstructured finite volume boussinesq model for near shore applications. presented at "Frontiers of Computational Physics : Modeling the Earth system", Boulder (CO), December 2012.

[46] M. Kazolea, A.I. Delis, I.K. Nikolos, and C.E. Synolakis. An unstructured finite volume numerical scheme for extended 2d boussinesq-type equations. *Coastal Engineering*, 69:42–66, 2012.

[47] M. Kazolea, A.I. Delis, and C. Synolakis. Numerical treatment of wave breaking on unstructured finite volume approximations for extended boussinesq type equations. *J.Comput.Phys.*, 2013. to appear.

[48] G. Kim, C. Lee, and K.-D. Shu. Internal generation of waves : Delta source function method and source terms addition method. *Coastal Engineering*, 34:2551–2264, 2007.

[49] D. Lannes and P. Bonneton. Derivation of asymptotic two-dimensional time-dependent equations for surface water wave propagation. *Physics of Fluids*, 21, 2009. 016601 doi:10.1063/1.3053183.

[50] Y.S. Li, S.X. Liu, Y.X. Yu, and G. Lai. Numerical modelling of boussinesq equations by finite element method. *Coastal Engineering*, 37, 1999.

[51] J. Maerz and G. Degrez. Improving time accuracy of residual distribution schemes. Technical Report VKI-PR 96-17, von Karman Institute for Fluid Dynamics, 1996.

[52] H. Nishikawa. A first-order system approach for diffusion equation. i: Second-order residual-distribution schemes. *Journal of Computational Physics*, 227(1):315 – 352, 2007.

[53] O. Nwogu. An alternative form of the boussinesq equations for near-shore wave propagation. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, pages 618–638, 1994.

[54] L. Sørensen O.R. Sørensen, H. Schaffer. Boussinesq-type modelling using an unstructured finite element technique. *Coastal Engineering*, 50, 2004.

[55] O.R. Sørensen P.A. Madsen. A new form of the boussinesq equations with improved linear dispersion characteristics. part 2: a slowing varying bathymetry. *Coastal Engineering*, 18:183–204, 1992.

[56] D.H. Peregnine. Long waves on a beach. *J.Fluid.Mech*, 27, 1967.

[57] R. Remi, D. De Santis, and M. Ricchiuto. High order preserving residual distribution schemes for advection-diffusion scalar problems on arbitrary grids. Research Report RR-8157, INRIA, November 2012. submitted to SINUM.

[58] M. Ricchiuto. On the c-property and generalized c-property of residual distribution for the shallow water equations. *Journal of Scientific Computing*, 48:304–318, 2011.

[59] M. Ricchiuto and R. Abgrall. Explicit runge-kutta residual distribution schemes for time dependent problems: Second order case. *Journal of Computational Physics*, 229(16):5653 – 5691, 2010.

[60] M. Ricchiuto, R. Abgrall, and H. Deconinck. Application of conservative residual distribution schemes to the solution of the shallow water equations on unstructured meshes,. *J. Comput. Phys.*, 222:287–331, 2007.

[61] M. Ricchiuto and A. Bollermann. Stabilized residual distribution for shallow water simulations. *J. Comput. Phys*, 228(4):1071–1115, 2009.

[62] P. L. Roe. Linear advection schemes on triangular meshes. Technical Report CoA 8720, Cranfield Institute of Technology, 1987.

[63] D.F Rogers. *An Introduction to NURBS: with Historical Perspectives*. Morgan Kaufman, 2001.

[64] J.A. Battjes S. Beji. Numerical simulations of nonlinear-wave propagation over a bar. *Coastal Engineering, 23, 1-16*, 1994.

[65] H.A. Schaffer and P.A. Madsen. Further enhancements of boussinesq-type equations. *Coastal Engineering*, 26:1–14, 1995.

[66] H.A. Schaffer and R Sorensen O. On the internal wave generation in boussinesq and mild-slope equations. *Coastal Engineering*, 53:319–323, 2006.

[67] M. Tissier, P. Bonneton, F. Marche, F. Chazel, and D. Lannes. A new approach to handle wave breaking in fully nonlinear boussinesq models. *Coastal Engineering*, 67, 2012.

[68] M. Tonelli and M. Petti. Hybrid finite volume - finite difference scheme for 2dh improved boussinesq equations. *Coastal Engineering*, 56, 2009.

[69] M. Tonelli and M. Petti. Simulation of wave breaking over complex bathymetries by a boussinesq model. *J. Hydraulic Res.*, 49, 2011.

[70] E. van der Weide and H. Deconinck. Positive matrix distribution schemes for hyperbolic systems. In *Computational Fluid Dynamics*, pages 747–753, New York, 1996. Wiley.

[71] M.A. Walkley. A numerical method for extended boussinesq shallow-water wave equations. PhD Thesis, University of Leeds, 1999.

[72] M.A. Walkley and M. Berzins. A finite element method for the two-dimensional extended boussinesq equations. *Int. J. Numer. Meth Fluids*, 39, 2002.

[73] G. Wei, J.T. Kirby, and A. Sinha. Generation of waves in boussinesq models using a source function method. *Coastal Engineering*, 36:271–299, 1999.

[74] R.W. Whalin. The limit of applicability of linear wave refraction theory in a convergence zone. Res.Rep.H-71-3, USACE,Waterways Expt. Station, Vicksburg, MS, 1971.

[75] C.H. Whiting and K.E. Jansen. A stabilized finite element formulation for the incompressible navier-stokes equations using a hierarchical basis. *International Journal for Numerical Methods in Fluids*, 35, 2001.