# A new Symmetric Interior Penalty Discontinuous Galerkin formulation for the Serre-Green-Naghdi equations

Meriem Zefzouf, Fabien Marche

# A new Symmetric Interior Penalty Discontinuous Galerkin formulation for the Serre-Green-Naghdi equations

Meriem Zefzouf [*1] and Fabien Marche[†1]

[1]University of Montpellier, Institut Montpelliérain Alexander Grothendieck, 34095 Montpellier, France

**Abstract**

In this work, we investigate the construction of a new discontinuous Galerkin discrete formulation to approximate the solution of Serre-Green-Naghdi (SGN) equations in the one-dimensional horizontal framework. Such equations describe the time evolution of shallow water free surface flows in the fully nonlinear and weakly dispersive asymptotic approximation regime. A new non conforming discrete formulation belonging to the family of Symmetric Interior Penalty discontinuous Galerkin methods (SIP-DG) is introduced to accurately approximate the solutions of the second order elliptic operator occurring in the SGN equations. We show that the corresponding discrete bilinear form enjoys some consistency and coercivity properties, thus ensuring that the corresponding discrete problem is well-posed. The resulting global discrete formulation is then validated through an extended set of benchmarks, including convergence studies and comparisons with data taken from experiments.

*Keywords:* Serre-Green–Naghdi equations, discontinuous Galerkin, Interior Penalty methods, high-order schemes, free surface flows, shallow water equations, dispersive equations

## 1 Introduction

In an incompressible, homogeneous and inviscid fluid, the propagation of surface waves is governed by the Euler equations with nonlinear boundary conditions at the surface and at the bottom. As this problem is still complicated to solve in its full generality, both mathematically and numerically, several simpler models have been derived to describe the behavior of the solution in some physical specific regimes (see, e.g., [34] for a review). In what follows, we focus on the *shallow water* regime:

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

where $H_0$ refers to the typical water depth, $\lambda$ the typical wave length. In this regime, the classical Nonlinear Shallow Water (NSW) equations [11] can be derived from the full water waves equations by neglecting all the terms of order $\mathcal{O}(\mu)$, see for instance [33]. This model provides an accurate description of important unsteady processes in the surf and swash zones, such as nonlinear wave transformations, run-up and flooding due to storm waves, see for instance [4], but it neglects the dispersive effects which are fundamental for the study of wave transformations in the shoaling area and, possibly, slightly deeper water areas. Keeping these dispersive effects in the equations, and neglecting only the $\mathcal{O}(\mu^2)$ terms, one obtains a more accurate - but mathematically and numerically more complicated - set of equations known as the Serre [51] in the horizontal surface dimension, or Green-Naghdi [27, 55], or fully nonlinear Boussinesq [60] equations for the two-dimensional case. We refer to these models here as the Serre-Green-Naghdi (SGN) equations. No smallness assumption is made on the size of the surface perturbations, the corresponding regime is said to be *fully nonlinear*:

$$(\textit{fully nonlinear / large amplitude regime}) \qquad \varepsilon := \frac{a}{H_0} = \mathcal{O}(1), \tag{2}$$

*Meriem.Zefzouf@umontpellier.fr
†Fabien.Marche@umontpellier.fr

where $a$ is the typical wave's amplitude. More precisely, let $x$, $z$, and $t$ denote, respectively, the horizontal, vertical, and time coordinates. We denote by $\zeta(t,x)$ the free surface elevation with respect to its rest state, by $-H_0+b(x)$ a parametrization of the bottom's variations and by $H := H_0+\zeta-b$ the water depth and by $\eta = H+b$ the total free surface elevation, as shown in Figure 1. Denoting by $u_{\rm hor}$ the horizontal component of the velocity field in the fluid domain, we define the vertically averaged horizontal velocity $u \in \mathbb{R}$ as

$$u(t,x) := \frac{1}{H}\int_{-H_0+b}^{\zeta} u_{\rm hor}(t,x,z)dz,$$

and we denote by $q := Hu$ the corresponding horizontal momentum.

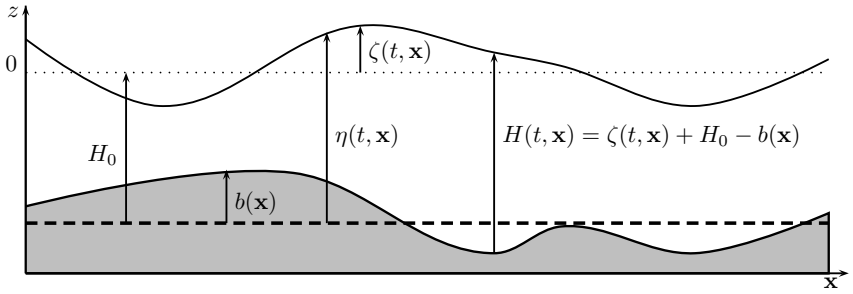


Figure 1: Free surface flow description and notations.

The SGN equations, as formulated in [5], read as follows:

$$\partial_t \eta + \partial_x q = 0, \tag{3a}$$
$$\partial_t q + \partial_x(uq) = -\left(1+\mathbf{T}[H,b]\right)^{-1}\mathbb{Q}[H,b](\eta,u), \tag{3b}$$

where the linear operator $\mathbf{T}[H,b]\cdot$ and the nonlinear operator $\mathbb{Q}[H,b]$ are defined for all smooth enough scalar-valued functions $w$ by

$$\mathbf{T}[H,b]w := H\mathcal{T}[H,b]\frac{w}{H}, \tag{4}$$
$$\mathbb{Q}[H,b](\eta,u) = gH\partial_x\eta + H\mathcal{Q}_1[H,b](u), \tag{5}$$
$$\mathcal{Q}_1[H,b](w) := -2\mathcal{R}_1[H,b]((\partial_x w)^2) + \mathcal{R}_2[H,b]((w\partial_x)^2 b), \tag{6}$$

the linear operator $\mathcal{T}[H,b]\cdot$ and the quadratic form $\mathcal{Q}[H,b](\cdot)$ are defined by

$$\mathcal{T}[H,b]w := \mathcal{R}_1[H,b](\partial_x w) + \mathcal{R}_2[H,b](w\partial_x b), \tag{7}$$

where

$$\mathcal{R}_1[H,b]w := -\frac{1}{3H}\partial_x(H^3 w) - \frac{H}{2}w\partial_x b, \qquad \mathcal{R}_2[H,b]w := \frac{1}{2H}\partial_x(H^2 w) + w\partial_x b.$$

Considering the discrete formulations, the SGN equations have recently received attention and various numerical methods have been introduced, mostly in the surface horizontal $d=1$ case, like Finite-Difference (FD) approaches [2], Finite-Volume (FV) [5, 8], WENO [6, 12], pseudo-spectral (PS) [22], (continuous) Finite-Element (FEM) [25, 43, 44], FV and FEM methods on hyperbolic relaxed approximating models [24, 29] and discontinuous Galerkin approaches (possibly mixed with FEM) in [16, 17, 20, 38, 45, 52]. In $d=2$ case, several methods have been developed mostly on cartesian meshes: FD [3, 60, 65], FV [37], hybrid FV-FD [47, 53] and WENO-FD [35], PS (in the rotating case) [46] and more recently a Hybridizable-DG method [50] and a Central DG-FE method in [39]. Numerical approximations of SGN equations on general unstructured (simplicial) meshes are considered in [21, 42].
We have introduced in [16] some high-order fully discontinous Galerkin discrete formulations for the (classical) SGN equations, as well as for the asymptotically enhanced SGN models of [35] in order to optimize the dispersive properties of the classical SGN equations. In these formulations, relying on the

Symmetric Weighted Interior Penalty DG approach (SWIP-DG), the fully nonlinear and weakly dispersive equations are written as coupled nonlinear (pseudo) hyperbolic-elliptic problems, relying on the (non-dispersive) NSW equations supplemented by additional algebraic source terms, which fully accounts for the $O(\mu^2)$ nonlinear dispersive correction. These source terms are themselves computed as the solutions of auxiliary linear second order elliptic problems associated with the elliptic operator $1+\mathcal{T}[H,b]$ defined in (**??**). We observe from this work, as well as in [21, 42], that high-order DG methods are well-suited for the approximation of the solutions of SGN equations. It is generally acknowledged that DG methods exhibit several appealing features (*i.e.* local conservation, stability, a straightforward ability to handle arbitrary high-order polynomial approximations, a great geometrical discretization flexibility, compact stencils and minimal inter-element communications). Beyond these general features, using DG method provide a general and unified discrete framework allowing to accurately approximate both the hyperbolic and elliptic parts of the SGN equations. Keeping nearshore oceanography applications in mind, the robustness of DG methods in the vicinity of sharp gradients also appears to be well suited for the description of wave steepening and breaking, see for instance [49]. Such features may also allow to easily introduce adaptive algorithms (refining or coarsening a grid can be achieved without enforcing the continuity property commonly associated with the conforming FEM), together with the use of high order polynomial approximations away from breaking and run-up areas.

However, considering the SWIP-DG method of [16] for the classical SGN equations (the Model 1 of [16]), we are not able to ensure the discrete coercivity for the discrete counterpart of the elliptic operator $1+\mathcal{T}[H,b]$ defined in (**??**) (see Remark 3 of [16] for details). As an answer to this drawback, we focus in this work on the construction and validation of a new Symmetric Interior Penalty discontinuous Galerkin formulation that is well-posed for the elliptic operator $H\left(1+\mathcal{T}[H,b]\right)$. This new formulation shares with the formulation introduced in [16] the appealing symmetry property and it is especially designed so that a discrete coercivity property is ensured, with a negligible computational overhead. As in [16], the stability threshold on the penalty parameter is still independent of the interface values of both $H$ and $b$ and it naturally allows to deal with the discontinuous nature of the discrete elliptic problem's coefficients in a stable and consistent way. Our numerical investigations show that this approach leads to an excellent agreement with data taken from several experiments.
The rest of this work is organized as follows: the next section is devoted to an alternative reformulation of the SGN equations, which proves to be useful in the numerical analysis that follows. The proposed new numerical method is introduced in third section and the numerical validations are provided in the last section.

## 2 Another reformulation of the SGN equations

In [16], in order to build a symmetric discrete formulation associated with $1+\mathbf{T}[H,b]$, it is observed that, assuming that the water depth $H$ is bounded away from zero, it holds that, for any sufficiently smooth scalar-valued function $v$:

$$(1+\mathbf{T}[H,b])v = H(1+\mathcal{T}[H,b])\frac{v}{H},$$

where, for any sufficiently smooth scalar-valued function $w$:

$$H(1+\mathcal{T}[H,b])w = -\frac{1}{3}\partial_x(H^3\partial_x w) + \left(H+\frac{1}{2}\partial_x(H^2\partial_x b)+H(\partial_x b)^2\right)w. \tag{8}$$

The approximation of the right-hand side of (3b) may consequently be computed as the solution of a non-homogeneous diffusion-reaction-like elliptic problem, with a time-dependent reaction-like coefficient defined as $H+\frac{1}{2}\partial_x(H^2\partial_x b)+H(\partial_x b)^2$. A Symmetric Weighted Interior Penalty discontinuous Galerkin (SWIP-DG) discrete bilinear form is proposed in [16], with non-homogeneous and time-dependent diffusion-like and reaction-like coefficients respectively defined from discrete approximations of $\nu[H]$ and $\beta[H,b]$. However, as already pointed out in [16], one main drawback of this formulation is that the sign of the reaction-like coefficient $H+\frac{1}{2}\partial_x(H^2\partial_x b)+H(\partial_x b)^2$ depends on the sign of $\partial_x(H^2\partial_x b)$. As a result, we are not able to ensure, for all topography configurations, a discrete coercivity property for the proposed discrete counterpart of the operator $H(1+\mathcal{T}[H,b])$.

In order to design a new symmetric DG formulation with better properties, let recall first that SGN equations are rigorously justified in [1], and a well-posedness result is proved for the general two-dimensional case with varying bottom using a Nash-Moser scheme. This result has also been obtained using a Picard iterative scheme in the one-dimensional case in [31].
More precisely, in the $d = 1$ case, assuming that the initial data $(\eta_0, u_0) \in H^s(\mathbb{R}) \times H^{s+1}(\mathbb{R})$ with $s > \frac{3}{2}$ (where $H^s(\mathbb{R})$ is the Sobolev space of functions $v \in L^2(\mathbb{R})$ such that their weak derivatives up to order $s$ have a finite $L^2$-norm) and that $b \in \mathcal{C}_b^\infty(\mathbb{R})$ (where $\mathcal{C}_b^\infty(\mathbb{R})$ is the space of infinitely differentiable functions that are bounded together with all their derivatives), then there exists a maximal time $t_{\max} > 0$, uniformly bounded with respect to $\mu$, such that SGN admits a unique solution $(\eta, u) \in \mathcal{C}([0, t_{\max}], H^s(\mathbb{R}) \times H^{s+1}(\mathbb{R}))$. This result can be extended to finite domains and periodic boundary conditions. In particular, we recall the following result:

**Proposition 1.** Assuming $b \in \mathcal{C}_b^\infty(\mathbb{R})$ and $\eta \in W^{1,\infty}(\mathbb{R})$ such that

$$\exists H_\varepsilon,\ \inf_{x\in\mathbb{R}} H = \eta - b > H_\varepsilon.$$

then the operator

$$H(1 + \mathcal{T}[H, b]) \,:\, H^2(\mathbb{R}) \ \to\ L^2(\mathbb{R}),$$

is well-defined, one-to-one and onto.

*Proof.* The detailed proof is given in [31] for the nondimensionalized equations. Let just recall here that the following identity holds:

$$H\mathcal{T}[H, b] = (S_1[H, b])^\star H\, S_1[H, b] + (S_2[b])^\star H\, S_2[b], \tag{9}$$

where the first and zero order differential operators $S_1[H, b]$ and $S_2[b]$ are defined as follows:

$$S_1[H, b] = \frac{H}{\sqrt{3}}\partial_x - \frac{\sqrt{3}}{2}\partial_x b, \qquad S_2[b] = \frac{1}{2}\partial_x b,$$

with adjoint operators

$$(S_1[H, b])^\star = -\frac{1}{\sqrt{3}}\partial_x(H\cdot) - \frac{\sqrt{3}}{2}\partial_x b, \qquad (S_2[b])^\star = S_2[b].$$

Considering the bilinear form defined and continuous on $H^1(\mathbb{R}) \times H^1(\mathbb{R})$ by:

$$a(v, w) = \big(HS_1[H, b]v, S_1[H, b]w\big) + \big(HS_2[H, b]v, S_2[H, b]w\big) + \big(Hv, w\big), \tag{10}$$

where $(\cdot, \cdot)$ refers to the $L^2$-scalar product, we have:

$$a(w, w) = (H(1 + \mathcal{T}[H, b])w, w) \geq H_\varepsilon \left( \|w\|^2 + \|S_1[H, b]w\|^2 + \|S_2[b]w\|^2 \right),$$

so that $H(1 + \mathcal{T}[H, b])$ is one to one. We additionally have

$$H_\varepsilon \left( \|w\|^2 + \|S_1[H, b]w\|^2 + \|S_2[b]w\|^2 \right) \geq C \|w\|^2_{H^1(\mathbb{R})},$$

proving that $a(\cdot, \cdot)$ is coercive on $H^1(\mathbb{R})$ and from Lax-Milgram theorem, for all $f \in L^2(\mathbb{R})$, there exists unique $u \in H^1(\mathbb{R})$ such that, for all $w \in H^1(\mathbb{R})$, $a(u, w) = (f, w)$ . □

In the following, we aim at designing a new DG formulation for SGN equations that mimics these properties observed on the continuous model. While we choose in [16] to design the discrete version of $H\mathcal{T}[H, b]$ from the formulation (8), we rather propose here to take inspiration from the proof of the previous Proposition and use the fact that, for any $w \in H^2(\mathbb{R})$, we have:

$$(1+\mathbf{T}[H, b])w = \big(H + (S_1[H, b])^\star H\, S_1[H, b] + (S_2[b])^\star H\, S_2[b]\big)\frac{w}{H}.$$

As a consequence, system (3) can be equivalently reformulated as follows:

$$\partial_t \mathbf{w} + \partial_x \mathbb{f}(\mathbf{w}, b) + \mathbb{d}(\mathbf{w}, b) = \mathbb{b}(\mathbf{w}, \partial_x b), \tag{11a}$$

$$\big(H + (S_1[H,b])^\star H\, S_1[H,b] + (S_2[b])^\star H\, S_2[b]\big)\Psi = \mathbb{Q}[H,b](\eta,u), \tag{11b}$$

where $\mathbf{w} := (\eta, q)$ gathers the flow variables, $\Psi$ is an auxiliary coupling variable between (11a) and (11b) and

$$\mathbb{f}(\mathbf{w},b) := \begin{pmatrix} q \\ \mathfrak{f}(\mathbf{w},b) \end{pmatrix}, \quad \mathbb{b}(\mathbf{w},\partial_x b) := \begin{pmatrix} 0 \\ -g\eta\partial_x b \end{pmatrix}, \quad \mathbb{d}(\mathbf{w},b) := \begin{pmatrix} 0 \\ H\Psi - gH\partial_x\eta \end{pmatrix}, \tag{11c}$$

with

$$\mathfrak{f}(\mathbf{w},b) := uq + \frac{1}{2}g(\eta^2 - 2\eta b).$$

Additionally observing that

$$\Big(H + S_1[H,b]^\star H\, S_1[H,b] + S_2[b]^\star H\, S_2[b]\Big)w \\ = \partial_x(-\kappa[H]\partial_x w) - \beta[H,\partial_x b]\partial_x w + \partial_x(\beta[H,\partial_x b]\, w) + \delta[H,\partial_x b]w,$$

where

$$\kappa[H] := \frac{1}{3}H^3, \qquad \beta[H,\partial_x b] := \frac{\sqrt{3}}{2}\kappa[H]^{\frac{1}{2}}\chi[H,\partial_x b], \qquad \delta[H,\partial_x b] := \chi[H,\partial_x b]^2 + H, \tag{12a}$$

and

$$\chi[H,\partial_x b] = H^{\frac{1}{2}}\partial_x b, \tag{12b}$$

system (11) may be finally written as

$$\partial_t \mathbf{w} + \partial_x \mathbb{f}(\mathbf{w}, b) + \mathbb{d}(\mathbf{w}, b) = \mathbb{b}(\mathbf{w}, \partial_x b), \tag{13a}$$

$$\partial_x\big(-\kappa[H]\partial_x\Psi\big) - \beta[H,\partial_x b]\partial_x\Psi + \partial_x(\beta[H,\partial_x b]\Psi) + \delta[H,\partial_x b]\Psi = \mathbb{Q}[H,b](\eta,u). \tag{13b}$$

**Remark 1.** In practice, the numerical solution of (13) is sought in a bounded spatial domain $\Omega \subset \mathbb{R}$. To close (13), we have to prescribe suitable boundary conditions for the primal variables $\mathbf{w} := (\eta, q)$ but also for the auxiliary variable $\Psi$, for the corresponding discrete problem to be well-defined on $\Omega$. In the following, we consider only *reflexive* boundary conditions, which aim at mimicking the presence of a solid-wall, and which may be obtained by enforcing :

$$\partial_x \eta_{|\partial\Omega} = 0, \qquad q_{|\partial\Omega} = 0, \qquad \partial_x \Psi_{|\partial\Omega} = 0. \tag{14}$$

All the following analysis may be reproduced in the case of *periodic* boundary conditions, adapting the discrete formulation accordingly.

**Remark 2.** It is classical that the frequency dispersion of (13) can be improved by adding some terms of order $O(\mu^2)$ to the momentum equation, see for instance [5, 41, 61]. With such optimization, system (13) becomes:

$$\partial_t \mathbf{w} + \partial_x \mathbb{f}(\mathbf{w}, b) + \mathbb{d}_\alpha(\mathbf{w}, b) = \mathbb{b}(\mathbf{w}, \partial_x b), \tag{15a}$$

$$\partial_x\big(-\kappa[H]\partial_x\Psi\big) - \beta[H,\partial_x b]\partial_x\Psi + \partial_x(\beta[H,\partial_x b]\Psi) + \delta[H,\partial_x b]\Psi = \mathbb{Q}_\alpha[H,b](\eta,u), \tag{15b}$$

where

$$\mathbb{d}_\alpha(\mathbf{w},b) := \begin{pmatrix} 0 \\ H\Psi - \frac{1}{\alpha}gH\partial_x\eta \end{pmatrix}, \tag{16}$$

and

$$\kappa[H] := \frac{\alpha}{3}H^3, \qquad \beta[H,\partial_x b] := \frac{\sqrt{3}}{2}\kappa[H]^{\frac{1}{2}}\chi[H,\partial_x b], \qquad \delta[H,\partial_x b] := \chi[H,\partial_x b]^2 + H, \tag{17}$$

$$\chi[H,\partial_x b] = \alpha^{\frac{1}{2}}H^{\frac{1}{2}}\partial_x b, \tag{18}$$

$$\mathbb{Q}_\alpha[H,b](\eta,u) = \frac{1}{\alpha}gH\partial_x\eta + H\mathcal{Q}_1[H,b](u), \tag{19}$$

where $\alpha$ may be set to 1.159 in order to provide a better description of the wave in intermediate water depth. This is the formulation upon which our new discrete bilinear form is designed in the next section. Note that choosing $\alpha = 1$ allows to recover (13) and that the subscript index $\alpha$ is removed in the following for the sake of simplicity.

# 3 Discrete formulations

In this section, we design a new discontinuous Galerkin (DG) discrete formulation for the SGN equations, as written in (15). Although we work here in one space dimension, we keep the notations as close as possible to the classical one for DG methods in higher space dimensions, both to facilitate the reader familiar with DG methods, and to make easier the extension to two space dimensions.

## 3.1 Setting and notations

Let $\Omega \subset \mathbb{R}$ denote an open segment with boundary $\partial\Omega$. We consider a partition $\mathcal{T}_h$ of $\Omega$ in $|\mathcal{T}_h|$ open disjoint segments $T$ of boundary $\partial T$ such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$. The partition is characterized by the meshsize $h := \max_{T\in\mathcal{T}_h} h_T$, where $h_T$ is the length of the element $T$. For all $T \in \mathcal{T}_h$, we denote by $\mathrm{n}_T$ the unit outward normal taking values in $\{-1, 1\}$ on $\partial T$, and by $x_T$ its barycenter.

Mesh faces, reduced here to vertices, are collected in the set $\mathcal{F}_h$ partitioned as $\mathcal{F}_h = \mathcal{F}_h^{\mathrm{i}} \cup \mathcal{F}_h^{\mathrm{b}}$, where $\mathcal{F}_h^{\mathrm{i}}$ collects the Interior vertices and $\mathcal{F}_h^{\mathrm{b}}$ the (two) boundary vertices. The abscissa of a vertex $F \in \mathcal{F}_h$ is denoted by $x_F$, and we let $h_F$ denote the minimum length of the mesh elements to which $F$ belongs. For all $T \in \mathcal{T}_h$, $\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$ denotes the set of vertices in $\partial T$ and, for all $F \in \mathcal{F}_T$, $\mathrm{n}_{TF}$ is the unit normal to $F$ pointing out of $T$. For any Interior vertex $F \in \mathcal{F}_h^{\mathrm{i}}$, we choose an arbitrarily oriented but fixed unit normal $\mathrm{n}_F$, and we set $\mathrm{n}_F := \mathrm{n}_{TF}$ for all boundary vertices $F \subset \partial T \cap \partial\Omega$. The maximum number of mesh faces composing the boundary of mesh elements is denoted by

$$N_\partial = \max_{T\in\mathcal{T}_h} \mathrm{card}(\mathcal{F}_T),$$

(and we obviously have $N_\partial = 2$ in the present setting). Given an integer polynomial degree $k \geq 1$, we consider the broken polynomial space

$$\mathbb{P}^k(\mathcal{T}_h) := \left\{ v \in L^2(\Omega) \,|\, v_{|T} \in \mathbb{P}^k(T) \;\; \forall T \in \mathcal{T}_h \right\}, \tag{20}$$

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

For a given final computational time $t_{\max} > 0$, we consider a partition $(t^n)_{0\leq n\leq N}$ of the time interval $[0, t_{\max}]$ with $t^0 = 0$, $t^N = t_{\max}$ and $t^{n+1} - t^n =: \Delta t^n$. More details on the computation of the time step $\Delta t^n$ and on the time marching algorithms are given in Section 3.6. For any sufficiently regular scalar-valued function of time $w$, we let $w^n := w(t^n)$.

We introduce the following inner products for regular enough scalar-valued functions $v, w$:

$$(v,w)_\Omega := \int_\Omega v(x)w(x)dx, \quad (v,w)_T := \int_T v(x)w(x)dx \quad \forall T \in \mathcal{T}_h, \quad (v,w)_F := (vw)(x_F) \quad \forall F \in \mathcal{F}_h,$$

and we denote respectively by $\|v\|_\Omega = (v,v)_\Omega^{\frac{1}{2}}$, $\|v\|_T = (v,v)_T^{\frac{1}{2}}$ and $\|v\|_F = (v,v)_F^{\frac{1}{2}}$ the corresponding $L^2$ norms.
For all $T \in \mathcal{T}_h$, we denote $p_T^k$ the $L^2$-orthogonal projector onto $\mathbb{P}^k(T)$ and $p_{\mathcal{T}_h}^k$ the $L^2$-orthogonal projector onto $\mathbb{P}^k(\mathcal{T}_h)$. Similarly, we denote $I_T^k$ the element nodal interpolation into $\mathbb{P}^k(T)$. The corresponding nodal distributions in elements and edges are approximated optimal nodes introduced in [7], which have better approximation properties than equidistant distributions. The global $I_{\mathcal{T}_h}^k$ interpolation into $\mathbb{P}^k(\mathcal{T}_h)$ is obtained by gathering the local interpolating polynomials defined on each elements.
For further use, we recall the following classical discrete trace inequality (see, e.g., [15, Lemma 1.46] for proof and details related to the mesh regularity parameters):

$$\forall v_h \in \mathbb{P}^k(\mathcal{T}_h), \forall T \in \mathcal{T}_h, \forall F \in \mathcal{F}_T, \; h_T^{\frac{1}{2}} \|v_h\|_F \leq C_{tr} \|v_h\|_T, \tag{21}$$

where the constant $C_{tr}$ only depends on $k$, the horizontal surface dimension $d$, and the mesh regularity parameters.

### 3.2 Symmetric and Weighted Interior Penalty discrete bilinear form

Let $\kappa, \beta, \delta \in L^\infty(\Omega)$ denote uniformly bounded coefficients and set, for the sake of simplicity, $\kappa_T := \kappa_{|T}$, $\beta_T := \beta_{|T}$ and $\delta_T := \delta_{|T}$ for all $T \in \mathcal{T}_h$. Following [13, 18], we define the jump and weighted average operators such that, for a sufficiently regular function $\varphi$ and an interior vertex $F \in \mathcal{F}_h^i$ such that $F \subset \partial T_1 \cap \partial T_2$ for distinct mesh elements $T_1$ and $T_2$,

$$[\![\varphi]\!] := \varphi_{|T_1} - \varphi_{|T_2}, \qquad \{\!\{\varphi\}\!\}_{\omega,F} := \omega_2 \varphi_{|T_1} + \omega_1 \varphi_{|T_2}, \quad \omega_i := \frac{\kappa_{T_i}}{\kappa_{T_1} + \kappa_{T_2}} \quad \forall i \in \{1,2\}. \tag{22}$$

In what follows, and when no confusion can arise, we omit the subscript $F$ from both $[\![v]\!]_{\omega,F}$ and $\{\!\{v\}\!\}_{\omega,F}$. When $\kappa \equiv C$ in $\Omega$ for some real number $C > 0$, we have $\omega_1 = \omega_2 = \frac{1}{2}$, and also the subscript $\omega$ is omitted. The definition of the average and jump operators at boundary vertices depends on the selected variable, according to the prescribed boundary conditions. We refer the reader to [15, Section 4.5] for a discussion on the role of weighted averages and harmonic means in the context of heterogeneous diffusion problems.

For further use, let consider the following bilinear form $a_h(\kappa, \beta, \delta; \cdot, \cdot)$ defined on $\mathbb{P}^k(\mathcal{T}_h) \times \mathbb{P}^k(\mathcal{T}_h)$:

$$\begin{aligned} a_h(\kappa, \beta, \delta; v_h, w_h) := & \sum_{T \in \mathcal{T}_h} \left(\kappa \partial_x v_h, \partial_x w_h\right)_T + \sum_{F \in \mathcal{F}_h} \xi \left(\frac{\gamma_{\kappa,F}}{h_F} [\![v_h]\!], [\![w_h]\!]\right)_F \\ & - \sum_{F \in \mathcal{F}_h} \left(\left(\{\!\{\kappa \partial_x^h v_h\}\!\}_\omega, [\![w_h]\!]\right)_F + \left([\![v_h]\!], \{\!\{\kappa \partial_x^h w_h\}\!\}_\omega\right)_F\right) \\ & - \sum_{T \in \mathcal{T}_h} \left(\beta v_h, \partial_x w_h\right)_T - \sum_{T \in \mathcal{T}_h} \left(\partial_x v_h, \beta w_h\right)_T \\ & + \sum_{F \in \mathcal{F}_h} \left(\left(\{\!\{\beta v_h\}\!\}_\omega, [\![w_h]\!]\right)_F + \left([\![v_h]\!], \{\!\{\beta w_h\}\!\}_\omega\right)_F\right) \\ & + \sum_{T \in \mathcal{T}_h} \left(\delta v_h, w_h\right)_T, \end{aligned} \tag{23}$$

with a $\kappa$-dependent penalty parameter $\gamma_{\kappa,F}$ defined as follows:

$$\gamma_{\kappa,F} := \begin{cases} \frac{2\kappa_{T_1}\kappa_{T_2}}{\kappa_{T_1}+\kappa_{T_2}} & \text{if } F \in \mathcal{F}_h^{\mathrm{i}} \text{ is such that } F = \partial T_1 \cap \partial T_2, \\ \kappa_T & \text{if } F \in \mathcal{F}_h^{\mathrm{b}} \text{ is such that } F = \partial T \cap \partial\Omega. \end{cases}$$

In (23), $\xi$ denotes a user-defined parameter sufficiently large to ensure coercivity (see Proposition 3) and $\partial_x^h$ has to be intended as the partial derivative along $x$ localized to mesh elements of $\mathcal{T}_h$, that is, for a given function $v_h \in \mathbb{P}^k(\mathcal{T}_h)$, $(\partial_x^h v_h)_{|T} = \partial_x(v_{h|T})$.

**Remark 3.** Taking $\beta := 0$ and $\delta := 0$ in (23), one obtains the Symmetric Weighted Interior Penalty-DG bilinear form of [13, 18] associated with heterogeneous diffusion problems.

**Remark 4.** In [16], the discrete formulation associated with (8) reads as follows

$$\begin{aligned} a_h(\kappa, \delta; v_h, w_h) := & \sum_{T \in \mathcal{T}_h} \left(\kappa \partial_x v_h, \partial_x w_h\right)_T + \sum_{F \in \mathcal{F}_h} \xi \left(\frac{\gamma_{\kappa,F}}{h_F} [\![v_h]\!], [\![w_h]\!]\right)_F \\ & - \sum_{F \in \mathcal{F}_h} \left(\left(\{\!\{\kappa \partial_x^h v_h\}\!\}_\omega, [\![w_h]\!]\right)_F + \left([\![v_h]\!], \{\!\{\kappa \partial_x^h w_h\}\!\}_\omega\right)_F\right) \\ & + \sum_{T \in \mathcal{T}_h} \left(\delta v_h, w_h\right)_T. \end{aligned} \tag{24}$$

Hence, we emphasize the occurrence of two new symmetric first-order terms $-\sum_{T \in \mathcal{T}_h} \left(\beta v_h, \partial_x w_h\right)_T - \sum_{T \in \mathcal{T}_h} \left(\partial_x v_h, \beta w_h\right)_T$ and $\sum_{F \in \mathcal{F}_h} \left(\left(\{\!\{\beta v_h\}\!\}_\omega, [\![w_h]\!]\right)_F + \left([\![v_h]\!], \{\!\{\beta w_h\}\!\}_\omega\right)_F\right)$ in the new formulation (23).

### 3.3 Discrete gradient and Laplace operators

As in [16], in order to discretize the linear and nonlinear operators that appear in our models, we need discrete counterparts of the gradient and of the Laplacian applied to broken polynomial functions. For any $v_h \in \mathbb{P}^k(\mathcal{T}_h)$, we define the following global lifting of the jumps of $v_h$ (see, e.g. [15, Section 4.3]):

$$\mathcal{R}_h^k([\![v_h]\!]) := \sum_{F\in\mathcal{F}_h} r_F^k([\![v_h]\!]),$$

where, for all $F \in \mathcal{F}_h$, the local lifting operator $r_F^k([\![v_h]\!]) \in \mathbb{P}^k(\mathcal{T}_h)$ is defined as the unique solution of the following problem:

$$\big(r_F^k([\![v_h]\!]), \psi_h\big)_\Omega = \big([\![v_h]\!], \{\!\{\psi_h\}\!\}\mathrm{n}_F\big)_F \qquad \forall \psi_h \in \mathbb{P}^k(\mathcal{T}_h),$$

with $\{\!\{\psi_h\}\!\}$ standard average operators given by (22) with $\omega_1 = \omega_2 = \frac{1}{2}$ at Interior nodes and extended as described in the previous section to boundary nodes. Following [15, Section 2.3], we define the discrete gradient operator $\nabla_h^k : \mathbb{P}^k(\mathcal{T}_h) \to \mathbb{P}^k(\mathcal{T}_h)$ such that, for all $v_h \in \mathbb{P}^k(\mathcal{T}_h)$,

$$\nabla_h^k v_h := \partial_x^h v_h - \mathcal{R}_h^k([\![v_h]\!]).$$

This gradient has better consistency properties than the broken (element-by-element) gradient $\partial_x^h$, as it accounts for the jumps of its argument through the second contribution; see [14, Theorem 2.2] for further insight into this point. Taking inspiration from [32, Eq. (2.10)], we also introduce the discrete Laplace operator $\Delta_h^k : \mathbb{P}^k(\mathcal{T}_h) \to \mathbb{P}^k(\mathcal{T}_h)$ such that, for all $v_h \in \mathbb{P}^k(\mathcal{T}_h)$, $\Delta_h^k(v_h)$ solves

$$-\big(\Delta_h^k v_h, \psi_h\big)_\Omega = a_h^{\mathrm{SIP}}(v_h, \psi_h) \qquad \forall \psi_h \in \mathbb{P}^k(\mathcal{T}_h),$$

where the bilinear form $a_h^{\mathrm{SIP}}(v_h, \psi_h)$ is given by (23) with the particular choices $\kappa := 1$, $\beta := 0$ and $\delta := 0$. It can be proved that, for any $v \in H_0^1(\Omega) \cap H^{k+1}(\Omega)$, it holds

$$\inf_{v_h\in\mathbb{P}^k(\mathcal{T}_h)} \|\nabla v - \nabla_h^k v_h\| \lesssim h^k, \qquad \inf_{v_h\in\mathbb{P}^k(\mathcal{T}_h)} \|\triangle v - \Delta_h^k v_h\| \lesssim h^{k-1},$$

where $a \lesssim b$ means $a \le Cb$ with real number $C > 0$ independent of the meshsize $h$, and the second estimate further requires mesh quasi-uniformity, see [14, 32].

### 3.4 The discrete problem

Let now $b_h \in \mathbb{P}^k(\mathcal{T}_h)$ denote a piecewise polynomial approximation of the topography parameterization $b$ which can be obtained either by $L^2$-orthogonal projection (*i.e.* $b_h = \pi_h^k(b)$) or by interpolation (*i.e.* $b_h = \mathcal{I}_h^k(b)$). Note that any order of approximation may be used for the definition of $b_h$, and in what follows, we choose the same order $k$ as for the primal variables, for the sake of simplicity, and the following shortcuts are introduced: $\nabla b_h = \nabla_h^k b_h$, $\Delta b_h = \Delta_h^k b_h$ and $\nabla^3 b_h = \nabla_h^k(\Delta_h^k b_h)$. The semi-discrete in space discontinuous Galerkin approximation of (15) reads:

Find $\mathbf{w}_h = (\eta_h, q_h) \in (\mathbb{P}^k(\mathcal{T}_h))^2$ and $\Psi_h \in \mathbb{P}^k(\mathcal{T}_h)$ such that, for all $(\phi_h, \varphi_h) \in (\mathbb{P}^k(\mathcal{T}_h))^2$,

$$\big(\partial_t \mathbf{w}_h, \varphi_h\big)_\Omega + \big(\mathcal{A}_h(\mathbf{w}_h), \varphi_h\big)_\Omega = 0, \tag{25a}$$

$$a_h(\kappa[H_h], \beta[H_h, \nabla b_h], \delta[H_h, \nabla b_h]; \Psi_h, \phi_h) = \big(\mathbb{Q}_h[H_h, b_h](\eta_h, u_h), \phi_h\big)_\Omega, \tag{25b}$$

where

(i) the discrete nonlinear operator $\mathcal{A}_h$ in (25a) is defined by

$$\big(\mathcal{A}_h(\mathbf{w}_h), \varphi_h\big)_\Omega := -\sum_{T\in\mathcal{T}_h} \big(\mathbb{f}(\mathbf{w}_h, b_h), \partial_x\varphi_h\big)_T + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} \big(\widehat{\mathbb{f}}_{TF}, \varphi_h\big)_F + \big(\mathbb{d}_h, \varphi_h\big)_\Omega - \big(\mathbb{b}(\mathbf{w}_h, \nabla b_h), \varphi_h\big)_\Omega, \tag{26}$$

and the discrete dispersive correction is defined as follows

$$\left(\mathbb{d}_h, \varphi_h\right)_\Omega = \begin{pmatrix} 0 \\ \left(H_h \Psi_h - \frac{1}{\alpha} g H_h \nabla_h^k \eta_h, \varphi_h\right)_\Omega \end{pmatrix}, \tag{27}$$

$\widehat{\mathbb{f}}_{TF}$ being a suitable approximation of the normal face fluxes $\mathbb{f}(\mathbf{w}_h, b_h) \cdot \mathrm{n}_{TF}$, see Section 3.5 below.

(ii) the discrete operators $\kappa[H_h]$, $\beta[H_h, \nabla b_h]$ and $\delta[H_h, \nabla b_h]$ are obtained according to (17)-(18), and recalled here:

$$\kappa[H_h] := \frac{\alpha}{3} H_h^3, \tag{28a}$$

$$\beta[H_h, \nabla b_h] := \frac{\sqrt{3}}{2} \kappa[H_h]^{\frac{1}{2}} \chi[H_h, \nabla b_h], \tag{28b}$$

$$\delta[H_h, \nabla b_h] := \chi[H_h, \nabla b_h]^2 + H_h, \tag{28c}$$

with

$$\chi[H_h, \nabla b_h] := \alpha^{\frac{1}{2}} H_h^{\frac{1}{2}} \nabla b_h. \tag{28d}$$

(iii) the discrete nonlinear operator $\mathbb{Q}_h[H_h, b_h]$ in (25b) is defined by

$$\mathbb{Q}_h[H_h, b_h](\eta_h, u_h) := \frac{1}{\alpha} g H_h \nabla_h^k \eta_h + H_h \mathcal{Q}_{1,h}[H_h, b_h](u_h).$$

where, for any $w_h \in \mathbb{P}^k(\mathcal{T}_h)$,

$$\begin{aligned} \mathcal{Q}_{1,h}[H_h, b_h](w_h) := {} & 2 H_h \nabla_h^k (H_h + \frac{b_h}{2}) (\nabla_h^k w_h)^2 + \frac{4}{3} H_h^2 \nabla_h^k w_h \Delta_h^k w_h \\ & + H_h \Delta b_h (\nabla_h^k w_h) w_h + \Big( \nabla_h^k \eta_h \Delta b_h + \frac{H_h}{2} \nabla^3 b_h \Big) w_h^2. \end{aligned}$$

(iv) the discrete water height $H_h$ and velocity $u_h$ are defined by $H_h := \eta_h - b_h$ and $u_h := p_{\mathcal{T}_h}^k (\frac{q_h}{H_h})$.

**Remark 5.** Note that the discrete operator $\delta[H_h, \nabla b_h] = \chi[H_h, \nabla b_h]^2 + H_h$ is non-negative, providing that the global discrete formulation preserves the positivity of $H_h$ (see §3.7). In [16], for the discrete formulation (24), we set

$$\delta[H_h, \nabla_h^k b_h] := H_h + \frac{1}{2} \nabla_h^k (H_h^2 \nabla_h^k b_h) + H_h (\nabla_h^k b_h)^2, \tag{29}$$

which is a major difference with the new discrete operator (28c), as it is not of constant sign due to the occurence of the term $\nabla_h^k (H_h^2 \nabla_h^k b_h)$.

### 3.5 Interface fluxes and well-balancing

The high-order reconstructed numerical flux detailed in [19] is a good default choice to approximate the interface fluxes $\mathbb{f}(\mathbf{w}, b) \cdot \mathrm{n}_T$, allowing to obtain a well-balanced scheme that preserves motionless steady states for SGN equations. Considering $T \in \mathcal{T}_h$ and $F \in \mathcal{F}_T \cap \mathcal{F}_h^i$ and denoting by $\mathbf{w}^-$, $\mathbf{w}^+$, respectively, the *interior* and *exterior* traces of $\mathbf{w}_h$ on $F$, with respect to the element $T$ and $b^-$ and $b^+$ the *interior* and *exterior* traces of $b_h$ on $F$, this numerical flux relies on reconstructed interface states $\check{\mathbf{w}}^-$, $\check{\mathbf{w}}^+$, $\check{b}$ such that

$$\widehat{\mathbb{f}}_{TF} = \mathbb{f}_h(\check{\mathbf{w}}^-, \check{\mathbf{w}}^+, \check{b}, \check{b}, \mathrm{n}_{TF}) + \widetilde{\mathbb{f}}_{TF}, \tag{30}$$

where $\mathbb{f}_h$ is the Lax-Friedrichs flux and $\widetilde{\mathbb{f}}_{TF}$ is a high-order correction term. The reader is referred to [16] for the detailed formulations.

### 3.6 Time discretization

Supplementing (13) with an initial datum $\mathbf{w}(0,\cdot)=\mathbf{w}^0$, the time stepping is carried out using the explicit SSP-RK schemes of [26]. For $k<3$, we consider RK-SSP schemes of order $(k+1)$. For instance, writing the semi-discrete equation (25a) in the operator form

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

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

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

where $\mathbf{w}_h^{n,i}$, $1\le i\le 2$, are the intermediate stages, $\Delta t^n$ is obtained from the CFL condition (31), and the discrete initial data $\mathbf{w}_h^0$ is defined either as the $L^2$-projection or interpolation on $(\mathbb{P}^k(\mathcal{T}_h))^2$ of $\mathbf{w}_0$. For $k\ge 3$, the five stages fourth order SSP-RK scheme of [54] is used (the details are omitted for the sake of simplicity) The corresponding time step $\Delta t^n$ is computed adaptively using the following CFL condition (see [9]):

$$\Delta t^n < \frac{1}{2k+1}\min_{T\in\mathcal{T}_h}\left(\frac{h_T}{\sigma_T}\right), \tag{31}$$

with

$$\sigma_T := \max_{\partial T}\left(\left|\frac{q_{h|T}}{H_{h|T}}\cdot \mathrm{n}_T\right| + \sqrt{gH_{h|T}}\right).$$

### 3.7 Positivity of the water height

While we choose not to focus on the issue of preserving the positivity of the water height in this paper, we emphasize that the proposed discrete formulation is compatible with the strategy of [20, 21] that ensures that the mean value of the water height remains positive, itself based on the ideas of [63, 64]. This approach is briefly recalled in the following. Let consider $H_{h|T}^n(x)$ obtained at the discrete time $t^n$ from the fully discrete previous dG formulation and $\bar{H}_T^n$ its average. Assuming that $\forall T\in\mathcal{T}_h$, $\bar{H}_T^n\ge 0$, we want to ensure that $\forall T\in\mathcal{T}_h$, $\bar{H}_T^{n+1}\ge 0$ without destroying the order of accuracy. For each element $T$:

1. let $S^T=\{r_j^T\}_{j=1,..,d}$ be the set of $d$ Legendre-Gauss-Lobatto (LGL) points on the element $T$, and $\{\hat{\omega}_j\}_{j=1,..,d}$ the corresponding quadrature weights. $d$ is chosen such that the associated quadrature rule is exact for polynomials of degree $k$ (*i.e.* $2d-3\ge k$). We compute $m_T^n = \min\limits_{r_i^T\in S^T} H_{h|T}^n(r_j^T)$.

2. we modify $H_{h|T}^n(x)$ in order to ensure that it is positive at the previous set of $d$ LGL nodes. This is done using the following conservative accuracy-preserving linear scaling around the cell average:

$$\check{H}_{h|T}^n(x) = \theta_T^n(H_{h|T}^n(x) - \bar{H}_T^n) + \bar{H}_T^n\,. \tag{32}$$

   where

$$\theta_T^n = \min\left(\frac{\bar{H}_T^n}{\bar{H}_T^n - m_T^n}, 1\right).$$

We deduce from (32) a modified polynomial $\check{\eta}_{h|T}^n(x)$ of order $k$, which is used to compute the numerical fluxes (30). Thus, following [63,64], the positivity of the mean water height $\bar{H}_T^{n+1}$, as well as the positivity of $H_h^n$ at chosen quadrature nodes, is ensured under the condition :

$$\sigma_T\frac{\Delta_t^n}{|T|}\le \hat{w}_1. \tag{33}$$

In practice, we have $\hat{w}_1=\frac{1}{6}$ for $N=2,3$ and $\hat{w}_1=\frac{1}{12}$ for $N=4,5$. Note that new *a posteriori* strategies based on Finite-Volume subcells and local flux reconstructions are currently under investigations.

### 3.8 Well-posedness of the discrete elliptic sub-problem

We show, in this section, that the proposed discrete bilinear form in (25b) is symmetric, consistent with the elliptic operator $H(1+\mathcal{T}[H,b])$ and enjoys some discrete coercivity property, provided that the penalty coefficient $\xi$ is large enough, so that the corresponding discrete problem (25b) is well-posed. From definitions (23) and (28), we have:

$$\begin{aligned}
a_h(\kappa[H_h],\beta[H_h,\nabla b_h],\delta[H_h,\nabla b_h];v_h,w_h) :=& \sum_{T\in\mathcal{T}_h}\big(\kappa[H_h]\partial_x v_h,\partial_x w_h\big)_T + \sum_{F\in\mathcal{F}_h}\xi\big(\tfrac{\gamma_{\kappa,F}}{h_F}[\![v_h]\!],[\![w_h]\!]\big)_F \\
&- \sum_{F\in\mathcal{F}_h}\Big(\big(\{\!\{\kappa[H_h]\partial_x^h v_h\}\!\}_\omega,[\![w_h]\!]\big)_F + \big([\![v_h]\!],\{\!\{\kappa[H_h]\partial_x^h w_h\}\!\}_\omega\big)_F\Big) \\
&- \frac{\sqrt{3}}{2}\sum_{T\in\mathcal{T}_h}\Big(\big(\kappa[H_h]^{\frac12}\chi[H_h,\nabla b_h]v_h,\partial_x w_h\big)_T + \big(\kappa[H_h]^{\frac12}\chi[H_h,\nabla b_h]\partial_x v_h,w_h\big)_T\Big) \\
&+ \frac{\sqrt{3}}{2}\sum_{F\in\mathcal{F}_h}\Big(\big(\{\!\{\kappa[H_h]^{\frac12}\chi[H_h,\nabla b_h]v_h\}\!\}_\omega,[\![w_h]\!]\big)_F + \big([\![v_h]\!],\{\!\{\kappa[H_h]^{\frac12}\chi[H_h,\nabla b_h]w_h\}\!\}_\omega\big)_F\Big) \\
&+ \sum_{T\in\mathcal{T}_h}\big((\chi[H_h,\nabla b_h]^2+H_h)v_h,w_h\big)_T.
\end{aligned} \tag{34}$$

This discrete bilinear form is obviously symmetric. In the following, we consider some smooth and uniformly bounded initial data $\mathbf{w}_0=(\eta_0,q_0)\in\mathcal{C}_b^\infty(\Omega)\times\mathcal{C}_b^\infty(\Omega)$, together with $b\in\mathcal{C}_b^\infty(\mathbb{R})$, so that the associated exact solution $\mathbf{w}=(\eta,q)$ of (15a), and the exact solution $\Psi$ of the elliptic problem (15b) satisfy the following jump conditions:

$$\forall F\in\mathcal{F}_h,\quad [\![\Psi]\!]_F=0,\quad [\![H]\!]_F=0,\quad [\![\partial_x b]\!]_F=0. \tag{35}$$

**Proposition 2.** (Consistency)

$$\forall\phi_h\in\mathbb{P}^k(\mathcal{T}_h),\quad a_h(\kappa[H],\beta[H,\partial_x b],\delta[H,\partial_x b];\Psi,\phi_h)=\big(\mathbb{Q}[H,b](\eta,u),\phi_h\big)_\Omega.$$

*Proof.* We set $\alpha=1$ in the following, for the sake of simplicity. To check the consistency of the discrete elliptic sub-problem (25b), we set $v_h:=\Psi$, $H_h:=H$, $b_h:=b$, $\nabla b_h:=\partial_x b$, $\kappa:=\kappa[H]$, $\beta:=\beta[H,\partial_x b]$, $\delta:=\delta[H,\partial_x b]$ in (34), to obtain:

$$\begin{aligned}
a_h(\kappa[H],\beta[H,\partial_x b],\delta[H,\partial_x b];\Psi,w_h) =& \sum_{T\in\mathcal{T}_h}\big(\tfrac13 H^3\partial_x\Psi,\partial_x w_h\big)_T - \sum_{F\in\mathcal{F}_h}\big(\tfrac13 H^3\partial_x\Psi,[\![w_h]\!]\big)_F \\
&- \frac12\sum_{T\in\mathcal{T}_h}\big(H^2\partial_x b\,\Psi,\partial_x w_h\big)_T + \frac12\sum_{F\in\mathcal{F}_h}\big(H^2\partial_x b\,\Psi,[\![w_h]\!]\big)_F \\
&- \frac12\sum_{T\in\mathcal{T}_h}\big(H^2\partial_x b\,\partial_x\Psi,w_h\big)_T + \sum_{T\in\mathcal{T}_h}\big(H((\partial_x b)^2+1)\Psi,w_h\big)_T,
\end{aligned} \tag{36}$$

and after integration by parts:

$$\begin{aligned}
a_h(\kappa[H],\beta[H,\partial_x b],\delta[H,\partial_x b];\Psi,w_h) =& -\sum_{T\in\mathcal{T}_h}\big(\partial_x(\tfrac13 H_h^3\partial_x\Psi),w_h\big)_T + \frac12\sum_{T\in\mathcal{T}_h}\big(\partial_x(H^2\partial_x b\,\Psi),w_h\big)_T \\
&- \frac12\sum_{T\in\mathcal{T}_h}\big(H^2\partial_x b\,\partial_x\Psi,w_h\big)_T + \sum_{T\in\mathcal{T}_h}\big(H((\partial_x b)^2+1)\Psi,w_h\big)_T,
\end{aligned} \tag{37}$$

so that $a_h(\kappa,\beta,\delta;\cdot,\cdot)$ is consistent in the following sense:

$$a_h(\kappa[H],\beta[H,\partial_x b],\delta[H,\partial_x b];\Psi,w_h)=\big(gH\partial_x\eta+\mathcal{Q}_1[H,b](u),w_h\big)_\Omega,\quad\forall w_h\in\mathbb{P}^k(\mathcal{T}_h).$$

□

In order to formulate the discrete stability result, let introduce the following norm, for all $v_h \in \mathbb{P}^k(\mathcal{T}_h)$,

$$|||v_h||| := \Big( \Big\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \Big\|_\Omega^2 + \|\chi[H_h, \nabla b_h] v_h\|_\Omega^2 + \Big\| H_h^{\frac{1}{2}} v_h \Big\|_\Omega^2 + |v_h|_{j,\kappa}^2 \Big)^{1/2}, \tag{38}$$

with the jumps seminorm

$$|v_h|_{j,\kappa} = \Big( \sum_{F \in \mathcal{F}_h} \frac{\gamma_{k,F}}{h_F} \|[\![v_h]\!]\|_F^2 \Big)^{1/2}. \tag{39}$$

Before addressing the discrete coercivity, we need some bounds on the boundary terms, which are addressed in the following Lemma.

**Lemma 1.** For all $(v_h, w_h) \in (\mathbb{P}^k(\mathcal{T}_h))^2$,

$$\Big| \sum_{F \in \mathcal{F}_h} \big( \{\!\{ \kappa[H_h] \partial_x^h v_h \}\!\}_\omega, [\![w_h]\!] \big)_F \Big| \le \Big( \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F \Big\| \kappa[H_h]_{|T}^{\frac{1}{2}} \partial_x^h v_{h|T}.n_F \Big\|_F^2 \Big)^{1/2} |w_h|_{j,\kappa}, \tag{40a}$$

$$\Big| \sum_{F \in \mathcal{F}_h} \big( \{\!\{ \kappa[H_h]^{\frac{1}{2}} \chi[H_h, \nabla b_h] v_h \}\!\}_\omega, [\![w_h]\!] \big)_F \Big| \le \Big( \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F \big\| \chi[H_h, \nabla b_h]_{|T} v_{h|T}.n_F \big\|_F^2 \Big)^{1/2} |w_h|_{j,\kappa}. \tag{40b}$$

*Proof.* Inequality (40a) is shown in e.g., [15, proof of Lemma 4.50]). For all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$, let introduce the following convenient shortcuts $\omega_i = \omega_{T_i,F}, \kappa_i = \kappa[H_h]_{|T_i}, \chi_i = \chi[H_h, \nabla b_h]_{|T_i}$ and $a_i = \chi_i v_{h|T_i}.n_F, i \in \{1,2\}$. The Cauchy-Schwartz inequality yields

$$\begin{aligned} \big( \{\!\{ \kappa[H_h]^{\frac{1}{2}} \chi[H_h, \nabla b_h] v_h \}\!\}_\omega, [\![w_h]\!] \big)_F &= \big( \omega_2 \kappa_1^{\frac{1}{2}} a_1 + \omega_1 \kappa_2^{\frac{1}{2}} a_2, [\![w_h]\!] \big)_F \\ &\le \Big( \frac{1}{2} h_F (\|a_1\|_F^2 + \|a_2\|_F^2) \Big)^{\frac{1}{2}} \\ &\times \Big( 2(\omega_2^2 \kappa_1 + \omega_1^2 \kappa_2) \frac{1}{h_F} \|[\![w_h]\!]\|_F^2 \Big)^{\frac{1}{2}}, \end{aligned}$$

and since $2(\omega_2^2 \kappa_1 + \omega_1^2 \kappa_2) = \gamma_{\kappa,F}$, we infer

$$\big( \{\!\{ \kappa[H_h]^{\frac{1}{2}} \chi[H_h, \nabla b_h] v_h \}\!\}_\omega, [\![w_h]\!] \big)_F \le \Big( \frac{1}{2} h_F (\|a_1\|_F^2 + \|a_2\|_F^2) \Big)^{\frac{1}{2}} \times \Big( \frac{\gamma_{\kappa,F}}{h_F} \Big)^{\frac{1}{2}} \|[\![w_h]\!]\|_F$$

Moreover, for all $F \in \mathcal{F}_h^b$ with $F = \partial T_1 \cap \partial \Omega$,

$$\big( \{\!\{ \kappa[H_h]^{\frac{1}{2}} \chi[H_h, \nabla b_h] v_h \}\!\}_\omega, [\![w_h]\!] \big)_F \le h_F^{\frac{1}{2}} \big\| \chi_T v_{|T}.n_F \big\|_F \times \Big( \frac{\gamma_{\kappa,F}}{h_F} \Big)^{\frac{1}{2}} \|[\![w_h]\!]\|_F .$$

Summing over mesh faces, using the Cauchy-Schwartz inequality, and regrouping the face contributions for each mesh element yields the assertion. □

**Proposition 3.** (Discrete coercivity) For all $\xi > \underline{\xi} := 2\sqrt{3}(1 + \frac{2}{\sqrt{3}} C_{tr}^2 N_\partial)(2 - \sqrt{3})^{-1}$, where $C_{tr}$ results from the discrete trace inequality (21), the bilinear form defined by (34) is coercive on $\mathbb{P}^k(\mathcal{T}_h)$ with respect to the $|||\cdot|||$-norm, i.e.,

$$\exists C_\xi > 0, \ \ \forall v_h \in \mathbb{P}^k(\mathcal{T}_h), \quad a_h(\kappa[H_h], \beta[H_h, \nabla b_h], \delta[H_h, \nabla b_h]; v_h, v_h) \ge C_\xi |||v_h|||^2,$$

where $C_\xi$ depends on $\xi$ and $C_{tr}$.

*Proof.* Again, we set $\alpha = 1$ in the following. Let $v_h \in V_h$, owing to the discrete trace inequality (21), we have

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F \Big\| \kappa[H_h]_{|T}^{\frac{1}{2}} \partial_x^h v_{h|T}.n_F \Big\|_F^2 &\le \quad C_{tr}^2 N_\partial \Big\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \Big\|_\Omega^2, \\ \sum_{T \in \mathcal{T}_h} \sum_{F \in F_T} h_F \big\| \chi[H_h, \nabla b_h]_{|T} v_{|T}.n_F \big\|_F^2 &\le \quad C_{tr}^2 N_\partial \|\chi[H_h, \nabla b_h] v_h\|_\Omega^2, \end{aligned}$$

and we infer from (40a)-(40b) that

$$\left| \sum_{F\in\mathcal{F}_h} \left( \{\!\{ \kappa[H_h]^{\frac{1}{2}} \chi[H_h,\nabla b_h] v_h \}\!\}_\omega, [\![ v_h ]\!] \right)_F \right| \leq C_{tr} N_\partial^{\frac{1}{2}} \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega |v_h|_{j,\kappa},$$

$$\left| \sum_{F\in\mathcal{F}_h} \left( \{\!\{ \kappa[H_h] \partial_x^h v_h \}\!\}_\omega, [\![ v_h ]\!] \right)_F \right| \leq C_{tr} N_\partial^{\frac{1}{2}} \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega |v_h|_{j,\kappa}.$$

Using Cauchy-Schwartz and Young inequalities, we also have

$$\left| \sum_{T\in\mathcal{T}_h} \left( \kappa[H_h]^{\frac{1}{2}} \chi[H_h,\nabla b_h] \partial_x v_h, v_h \right)_T \right| \leq \frac{1}{2} \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega^2 + \frac{1}{2} \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega^2.$$

As a result,

$$\begin{aligned} a_h(\kappa[H_h],\beta[H_h,\nabla b_h],\delta[H_h,\nabla b_h];v_h,v_h) \geq & \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega^2 - 2C_{tr} N_\partial^{\frac{1}{2}} \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega |v_h|_{j,\kappa} + \xi |v_h|_{j,\kappa}^2 \\ & + \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega^2 - \sqrt{3} C_{tr} N_\partial^{\frac{1}{2}} \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega |v_h|_{j,\kappa} \\ & - \frac{\sqrt{3}}{2} \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega^2 - \frac{\sqrt{3}}{2} \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega^2 + \left\| H^{\frac{1}{2}} v_h \right\|_\Omega^2. \end{aligned}$$

We now use the following inequality: let $\sigma$ be a positive real numbers, let $\xi \geq 2\sigma^2$, then

$$\forall (x,y) \in \mathbb{R}^2, \qquad x^2 - 2\sigma xy + \frac{\xi}{2} y^2 \geq \frac{\xi - 2\sigma^2}{2+\xi} (x^2 + y^2). \tag{41}$$

Applying this inequality twice, one for $\sigma_1 := C_{tr} N_\partial^{\frac{1}{2}}$, $x := \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega$, $y := |v_h|_{j,\kappa}$, and the other for $\sigma_2 := \frac{\sqrt{3}}{2} C_{tr} N_\partial^{\frac{1}{2}}$, $x := \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega$, $y := |v_h|_{j,\kappa}$, introducing $C_1 := (\xi - 2\sigma_1^2)(2+\xi)^{-1}$, $C_2 := (\xi - 2\sigma_2^2)(2+\xi)^{-1}$, and choosing $\xi$ such that $C_i - \frac{\sqrt{3}}{2} > 0$, $i = 1,2$, we infer,

$$\begin{aligned} a_h(\kappa[H_h],\beta[H_h,\nabla b_h],\delta[H_h,\nabla b_h];v_h,v_h) \geq & \left(C_1 - \frac{\sqrt{3}}{2}\right) \left\| \kappa[H_h]^{\frac{1}{2}} \partial_x^h v_h \right\|_\Omega^2 + \left(C_2 - \frac{\sqrt{3}}{2}\right) \left\| \chi[H_h,\nabla b_h] v_h \right\|_\Omega^2 \\ & + (C_1 + C_2) |v_h|_{j,\kappa}^2 + \left\| H^{\frac{1}{2}} v_h \right\|_\Omega^2, \end{aligned}$$

and finally

$$a_h(\kappa[H_h],\beta[H_h,\nabla b_h],\delta[H_h,\nabla b_h];v_h,v_h) \geq C |\!|\!| v_h |\!|\!|^2,$$

with $C = \min(C_1 - \frac{\sqrt{3}}{2}, C_2 - \frac{\sqrt{3}}{2}, C_1 + C_2, 1) > 0$. We remark that choosing the penalty coefficient $\xi$ large enough to enforce $C_i - \frac{\sqrt{3}}{2} > 0$ leads to $\xi > \frac{2\sqrt{3}+4\sigma_1^2}{2-\sqrt{3}}$. □

A straightforward consequence of the Lax-Milgram lemma is that the discrete problem (25b) is well-posed.

### 3.9 Arbitrary order well-balancing property

We have the following well-balancing result, obtained as a straightforward consequence of the well-balancing property for the (hyperbolic) shallow water equations (see for instance [19]) and the previous well-posedness result for the elliptic sub-problem:

**Proposition 4.** The discrete formulation (25) together with the numerical fluxes (30) and a first order Euler time-marching algorithm preserves the motionless steady states, providing that the integrals of (25a) are exactly computed for the motionless steady states: For all $n \in \mathbb{N}$ and all $\eta^e \in \mathbb{R}$,

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

*Proof.* Assuming $\mathbf{w}_h \equiv \mathbf{w}_h^e := (\eta^e, 0)$, we infer homogeneous conditions for (25b) (that is $\mathbb{Q}_h[H_h, b_h](\eta_h, u_h) = 0$ and homogeneous Neumann boundary conditions) and as (25b) is well-posed (see Propositions 2 and 3), we get $\Psi_h = 0$ as unique solution of (25b) and therefore $\left(\mathbb{d}_h, \varphi_h\right)_\Omega = 0$. The proof is then similar to the one of [19] for the (non-dispersive) shallow water equations. □

This analysis can also be extended to the high order SSP schemes of Section 3.6 exploiting the fact that the intermediate stages $\mathbf{w}_h^{n,i}$ are obtained as convex combinations of forward Euler substeps; see, e.g., [62].

# 4 Numerical validations

In this section, we validate the previous discrete formulation through several benchmark problems. In all the test cases, the time step restriction is computed according to the most restrictive condition between (31) and (33). Exploiting the symmetry of the discrete bilinear form, the sparse linear systems associated with the discretization of $H(1 + \alpha\mathcal{T}[H, b])$ are solved using a Cholesky methods. We set $\alpha = 1$ in the first three test cases, so that solitary waves of the form (42) are exact solutions. In the other test cases, we use the optimized value $\alpha = 1.159$. We consider *solid-wall* boundary conditions on the two domain's boundaries. Some numerical accuracy and convergence analysis are performed in test 1, using the $L^2$ norm defined, for any arbitrary scalar valued piecewise polynomial function $w_h \in \mathbb{P}^k(\mathcal{T}_h)$, as follows:

$$\|w_h\|^2_{\mathcal{T}_h} = \left(w_h, w_h\right)_{\mathcal{T}_h},$$

with

$$\left(v, w\right)_{\mathcal{T}_h} := \sum_{T \in \mathcal{T}_h} \left(v, w\right)_T.$$

## 4.1 Solitary wave propagation

In this first test, we investigate the propagation of a solitary wave as defined by (42), over a flat topography. In the flat bottom case, SGN equations admit solitary wave solutions of amplitude $\varepsilon H_0$, which have known formulae in a closed form:

$$\eta(t, x) = H_0 + \varepsilon H_0 \operatorname{sech}^2\left(\kappa(x - ct)\right), \qquad q(t, x) = c\left(\eta(t, x) - H_0\right), \forall t \geq 0, \forall x \in \Omega, \tag{42}$$

with $\kappa := \sqrt{\frac{3\varepsilon}{4H_0^2(1+\varepsilon)}}$ and $c := \sqrt{gH_0(1+\varepsilon)}$. The computational domain is $200\,m$ long, the reference water depth is set to $H_0 = 1m$ and the relative amplitude to $\epsilon = 0.1$. The solitary wave is initially centered at $x_0 = 80\,m$ and we compute the $L^2$-error between numerical results and reference solution for $\eta$ and $q$ at $t = 0.1\,s$ on a sequence of progressively refined uniform meshes. We start from mesh containing $|\mathcal{T}_h| = 800$ elements, and the computation is performed in double precision arithmetic. The corresponding $L^2$-errors are gathered in Figure 2 and we observe a scaling that lies between $\mathcal{O}(h^{k+\frac{1}{2}})$ and $\mathcal{O}(h^{k+1})$.

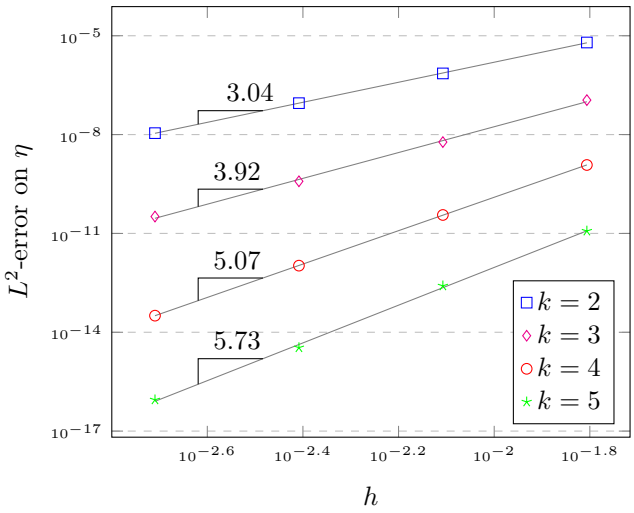


Figure 2: Test 1 - Solitary wave propagation: $L^2$-error for the total free surface elevation $\eta$ at $t_{\max} = 0.1\,s$ vs. $h$, for polynomial orders $k = 2, 3, 4, 5$.

## 4.2 Head-on collision of solitary waves

Let consider now the head-on collision of two identical solitary waves propagating in opposite directions (see [10] for an extensive study). The collision of the two waves is associated with a change of the nonlinear dispersion characteristics and the discrete formulation has to ensure the equilibrium between amplitude and frequency dispersion in order to allow the propagation at constant shape and speed. The computational domain is defined as $\Omega = (-200\,m, 200\,m)$. The initial condition is defined with two solitary waves (42) of relative amplitude $\varepsilon = 0.2$ initially located at $x = -50\,m$ and $x = 50\,m$ and with opposite velocities. The number of mesh elements is set to $|\mathcal{T}_h| = 800$, corresponding to a uniform meshsize of $h = 0.5\,m$, and the polynomial order to $k = 2$. We show on Figure 3 some snapshots of the free surface at several times during the propagation, including a zoom on the dispersive tail generated after the collision. We observe that the maximum wave amplitude during the collision is slightly larger than twice the initial amplitude, in agreement with the results of [10, 43, 48]. The dispersive tail is very well reproduced, even with this low number of mesh elements.

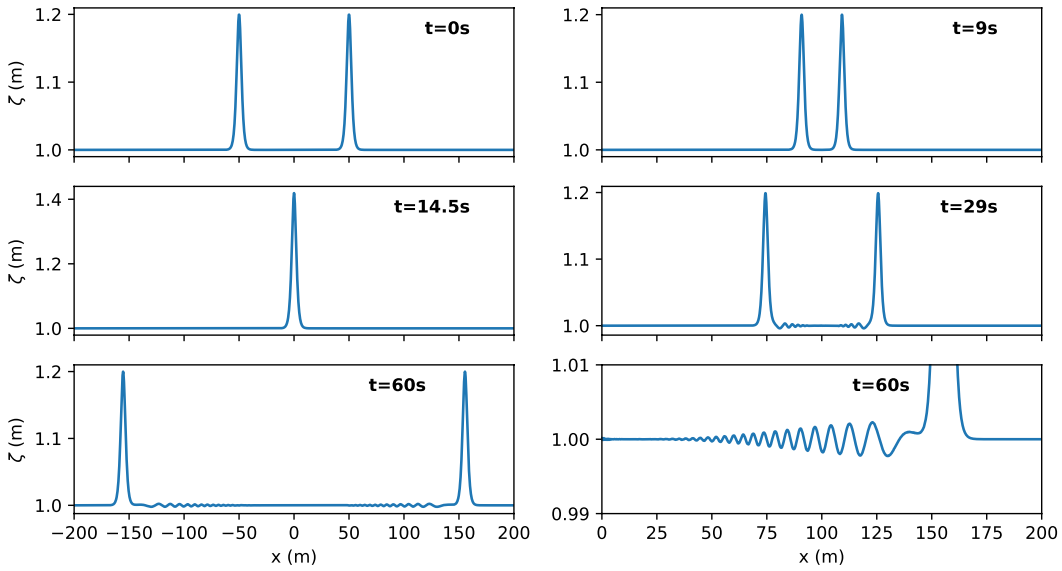


Figure 3: Test 2 - Head-on collision of solitary waves: snapshots of the free surface.

### 4.3 One dimensional dispersive dam-break problem

We now study the time evolution of a dispersive dam break problem over a flat bottom. We consider the computational domain $\Omega = (-300\,m,\, 300\,m)$ and the initial data is defined as :

$$H(0,x) = \frac{H_L - H_R}{2}(1 - \tanh(\frac{x}{\chi})), \quad q(0,x) = 0,$$

with $H_L = 1.8$, $H_R = 1$ and $\chi = 0.4$, and the velocity is initially uniformly set to zero, reproducing the set-up of [37] and emulating a piecewise constant initial data with a sharp variation initially located at $x = 0$. The number of mesh elements is set to $|\mathcal{T}_h| = 1500$, corresponding to a uniform meshsize of $h = 0.4\,m$, and the polynomial order to $k = 2$. As expected, the initial discontinuity divides into a dispersive shock wave propagating to the right and a rarefaction wave propagating to the left. We show on Figures 4 and 5 the structure of the water height and the velocity at time $t = 47.5\,s$. As pointed out in [23], the analysis of Riemann invariants of the shallow-water system, coupled with the analysis of the Witham system for SGN equations, allows to approximate the values $(H^*, u^*)$ of the mean flow dividing the rarefaction wave and the dispersive shock zones:

$$H^* = \frac{(\sqrt{H_L} + \sqrt{H_R})^2}{4}, \quad u^* = 2(\sqrt{gH^*} - \sqrt{gH_R}).$$

An asymptotic approximation of the amplitude of the lead soliton, denoted $a^+$ in the following, may also be obtained :

$$a^+ = \sigma_0 - \frac{1}{12}\sigma_0^2 + \mathcal{O}(\sigma_0^3),$$

where $\sigma_0 = H_L - H_R$ denotes the initial jump value (note that the values provided in [23] are obtained with nondimensionalized equations). Numerical results are compared with these values and show very good agreement, as highlighted on Figures 4 and 5. We emphasize that, for instance, the results shown in [37] or [58] are obtained respectively with 24000 and 8000 mesh elements and second order Finite-Volume schemes, while the use of a higher-order scheme allows us to obtain very satisfying results with only 1500 mesh elements.

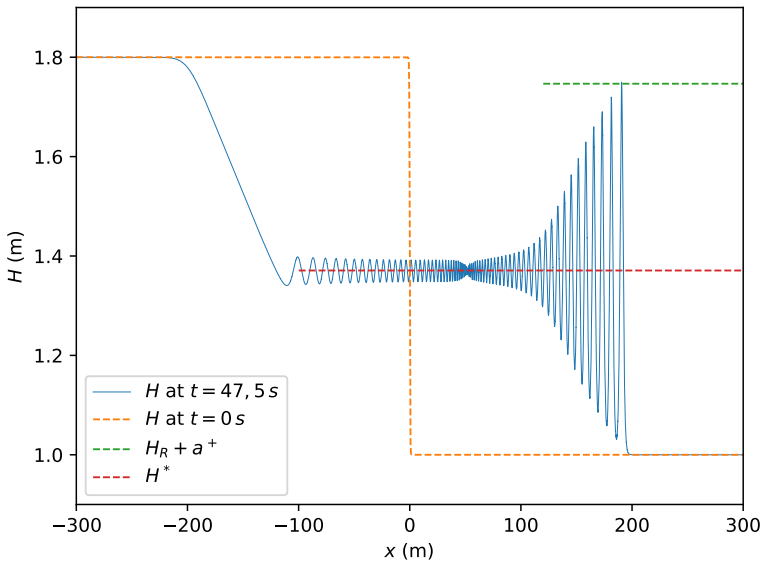


Figure 4: Test 3 - Dispersive dam-break: water depth profile at $t = 47.5\,s$.

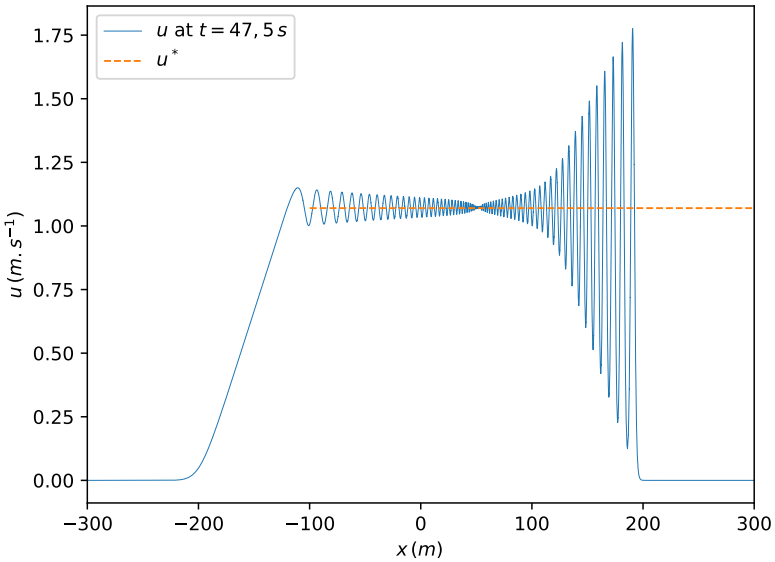


Figure 5: Test 3 - Dispersive dam-break: velocity profile at $t = 47.5\,s$.

## 4.4 Reflection of solitary waves at vertical walls

In this section, we focus on the propagation and reflexion of solitary waves against a vertical wall. We compare numerical results with two different sets of experimental data, coming for two different experiments and involving varying bottoms.
In the first experiment, the spatial domain is $60\,m$ long, the depth profile is piecewise linear, with a slope

$s$ defined as follows:

$$s(x) = \begin{cases} 0 & \text{if} \quad x \leq 40, \\ 1/50 & \text{if} \quad 40 \leq x \leq 60, \end{cases} \tag{43}$$

and terminated by a vertical solid wall located at $x = 60\,m$. The reader is referred to [59] for a complete description. The initial solitary wave, define from (42), is centered at $x = 10\,m$ m and is propagating from left to right. The still water depth is $H_0 = 0.7\,m$. Two runs are performed with two different initial solitary wave amplitudes, provided with relative amplitudes $\epsilon = 0.1$ and $\epsilon = 0.171$. The computational domain is uniformly discretized using $|\mathcal{T}_h| = 100$ mesh elements and the polynomial order is set to $k = 1$. Experimental data are compared with numerical results in Figure 6, in which time series of the surface elevation measured at a location near the solid wall ($x = 57.75$) are shown. We can observe the two expected peaks corresponding respectively to the incident and reflected waves and the discrete formulation provides a very accurate matching between simulation and experimental data, especially concerning the amplitude of the waves.

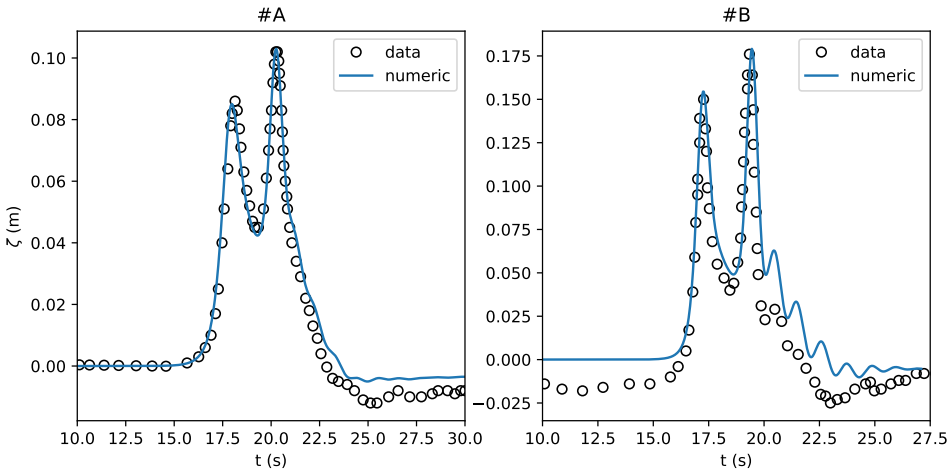


Figure 6: Test 4 - Reflexion of a solitary wave against a vertical wall (first test): time series of the free surface at x = 57.75 m for $\epsilon = 0.1$ (left) and $\epsilon = 0.171$ (right) - Comparison between numerical results (-) and experimental data (o).

In the second test case, we study the propagation of solitary waves over a composite beach which mimics the geometrical dimensions of the Revere Beach. The original experiment was performed in a tank by the U.S. Army corps of Engineers at the Coastal Engineering Research Center in Vicksburg, Mississippi and is depicted in details in [57]. The spatial domain is $33,23\,m$ long and the constructed beach consists of three piece-wise linear segments, terminated with a vertical wall on the left and the solitary wave is initially centered at $x = 0$, see Figure 7 for a sketch of the corresponding initial set-up.

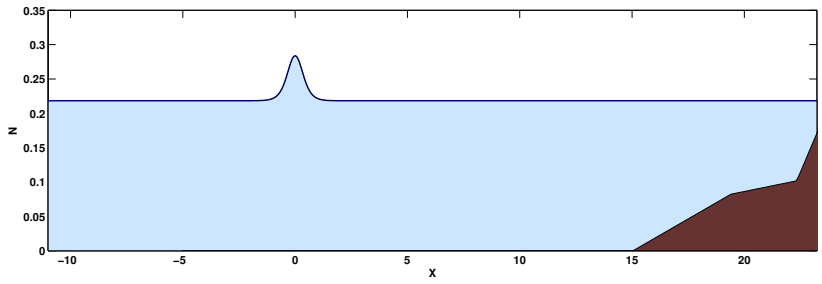


Figure 7: Test 5 - Reflection of a solitary wave on a composite beach: topography and initial free surface.

The slope $s$ of the topography is defined as follows:

$$s(x) = \begin{cases} 0 \quad \text{if} \quad x \leq 15.04, \\ 1/53 \quad \text{if} \quad 15.04 \leq x \leq 19.4, \\ 1/150 \quad \text{if} \quad 19.4 \leq x \leq 22.33, \\ 1/13 \quad \text{if} \quad 22.33 \leq x \leq 23.23. \end{cases} \tag{44}$$

We focus here on the case (B) of the experiment $\varepsilon = 0.28$ and we provide the solitary wave of targeted height, centered at $x = 0$, as the initial condition. The computational domain is uniformly discretized using $|\mathcal{T}_h| = 100$ mesh elements and the polynomial order is set to $k = 1$. We observe the propagation, reflection on the wall at the right boundary before traveling back to the left boundary. Experimental data are provided as time series of the wave elevation at several gauges located along the wave flume. We show on Figure 8 the comparison between data and computed results at wave gauges 5, 7 and 9, respectively located at $x_5 = 15.04\,m$, $x_7 = 19.4\,m$ and $x_9 = 22.33\,m$ (exactly at the locations corresponding to the slope variations). We observe a very accurate agreement, the wave amplitude and celerity during the propagation and reflection are very well reproduced, even with a very low number of mesh elements.

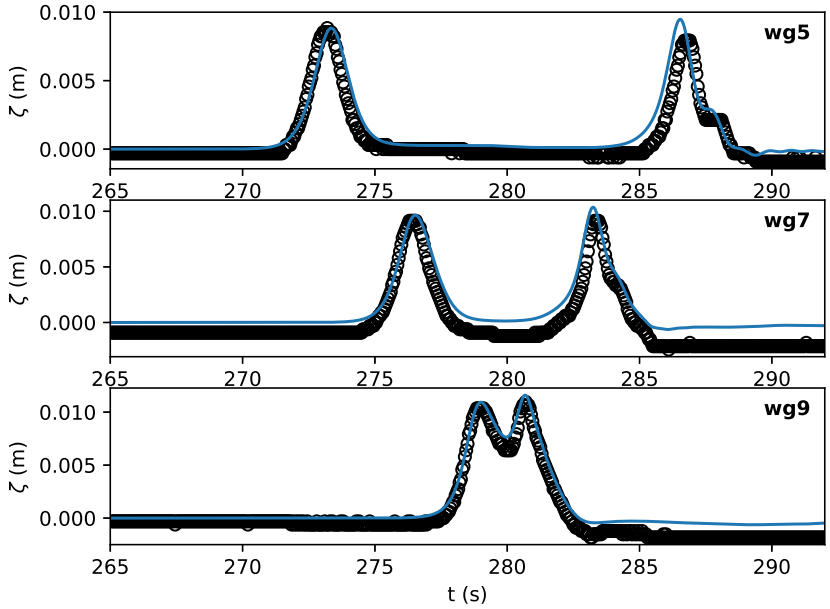


Figure 8: Test 5 - Reflection of a solitary wave on a composite beach: comparaison between experimental (o) data and numerical solution (-) at gauges 5,7 and 9 for case B ($\epsilon = 0.28$).

## 4.5 Shoaling of solitary waves

In this section, we investigate the dispersive properties of our model through the study of the nonlinear shoaling of solitary waves over constant bed slopes. In what follows, we numerically reproduce three different experimental configurations.
The first one relies on the data issued from the laboratory study performed at LEGI (Grenoble, France) and detailed in [30]. We consider in this test at $36\,m$ channel with constant bed slope $1:30$ and a solitary waves propagating from the left boundary, with a water level at rest $H_0 = 0.25\,m$. Measurements of the free surface are available in the vicinity of the breaking point during the simulation, at several wave gauges. We consider 4 series of experiments, involving increasing waves relative amplitudes, from $\epsilon = 0.096$ to $\epsilon = 0.534$. For each experiment, we compare the numerical results with time series of the free surface elevation at several wave gauges (the exact locations of the wave gauges are reported in Table 1).

| Incident wave amplitude: $\varepsilon = 0.096$ | | | |
|---|---|---|---|
| Gauge location (m) | 2.430 | 2.215 | 1.960 |
| Incident wave amplitude: $\varepsilon = 0.298$ | | | |
| Gauge location (m) | 3.980 | 3.765 | 3.510 |
| Incident wave amplitude: $\varepsilon = 0.456$ | | | |
| Gauge location (m) | 4.910 | 4.695 | 4.440 |
| Incident wave amplitude: $\varepsilon = 0.534$ | | | |
| Gauge location (m) | 5.180 | 4.965 | 4.710 |

Table 1: Test 6 - Shoaling of solitary waves: LEGI experiment - Location of wave gauges for solitary waves shoaling on a 1:30 sloped beach (relative to the shoreline).

Numerical results, obtained with $|\mathcal{T}_h| = 400$ and $k = 2$, are shown on Figure 9, together with the corresponding experimental data. We observe that for the 4 experiments, the corresponding shoaling and wave tranformations near the breaking points are very well reproduced.

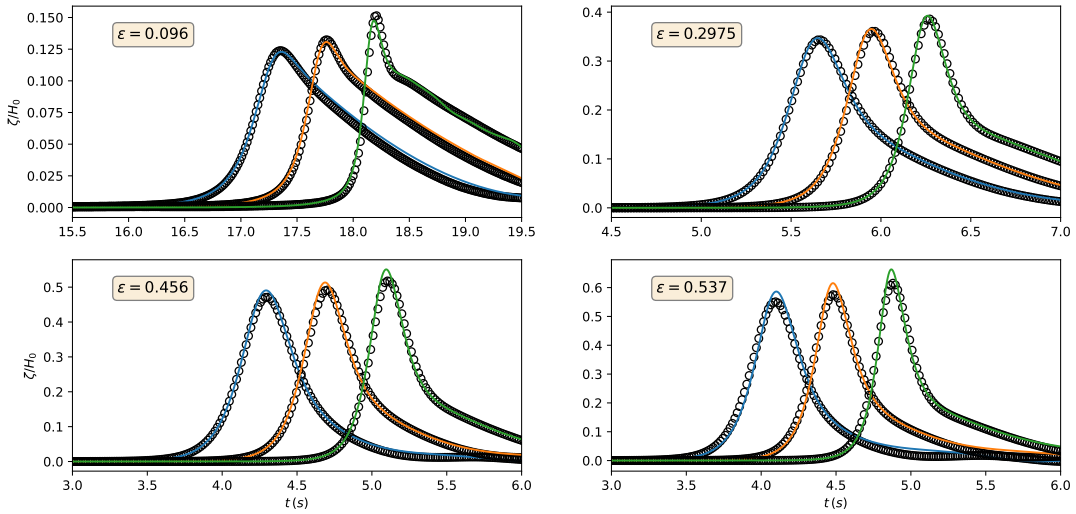


Figure 9: Test 6 - Shoaling of solitary waves: LEGI experiment: comparison between numerical (-) and experimental (o) time series of total free surface at three different locations before the breaking point.

The second test is based on an experimental set-up detailed in [28], for an incident solitary wave of relative amplitude $\epsilon = 0.2$, which propagates and breaks over a planar beach with a slope of $1 : 35$ in a computational domain of $85\,m$ long and a water level at rest $H_0 = 1\,m$.
Experimental data are provided as time series of the wave elevation at several gauges located along the wave flume. We show on Figure 10 the comparison between numerical results (obtained with $|\mathcal{T}_h| = 400$ and $k = 2$) and experimental data measured from five different wave gauges located at $x_1 = 20.96\,m, x_2 = 22.55\,m, x_3 = 23.68\,m, x_4 = 24.68\,m$, and $x_5 = 25.91\,m$, the last one being very close to the breaking point. We observe an excellent agreement between numerical and measured free surface elevation up to the breaking point.

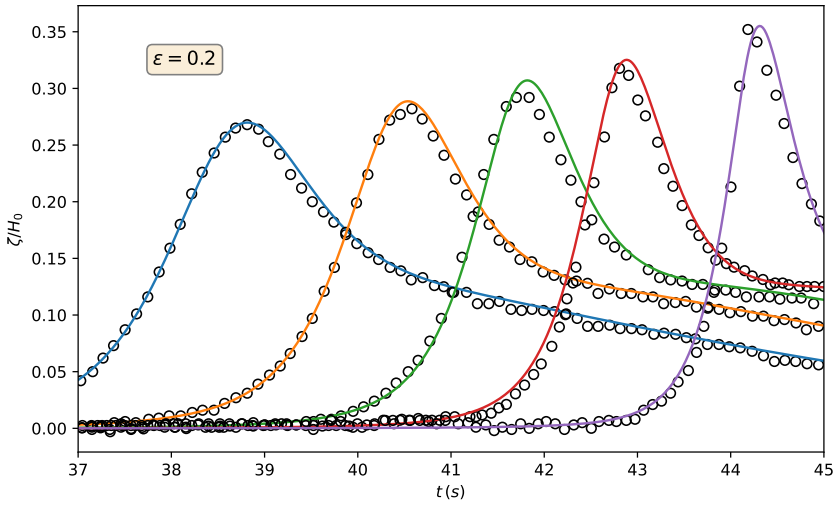


Figure 10: Test 7 - Shoaling of solitary waves: Grilli *et al.* experiment - Time series of the free surface elevation for the solitary wave propagating over the 1:35 sloping beach. (-) numerical results, (o) experimental data.

The third test allows to study the ability of the proposed discrete formulation to deal with the occurence of dry areas. We focus on the shoaling and run-up of a solitary wave over a beach with constant slope $1:19.85$, following the experiments of [56]. The incident wave is supplied using (42), with a water level at rest $H_0 = 1m$ and $\varepsilon = 0.28m$. The mesh is set to $|\mathcal{T}_h| = 250$ elements and the polynomial order to $k = 2$. The resulting numerical results are compared with the experimental data at several times during the propagation (with a nondimensionalized time $t^* = t\sqrt{g/H_0}$) and we observe a very close agreement, as shown on Figure 11. In particular, the run-up is reproduced very well, and we do not have to artificially cancel the dispersive correction in the vicinity of dry areas.

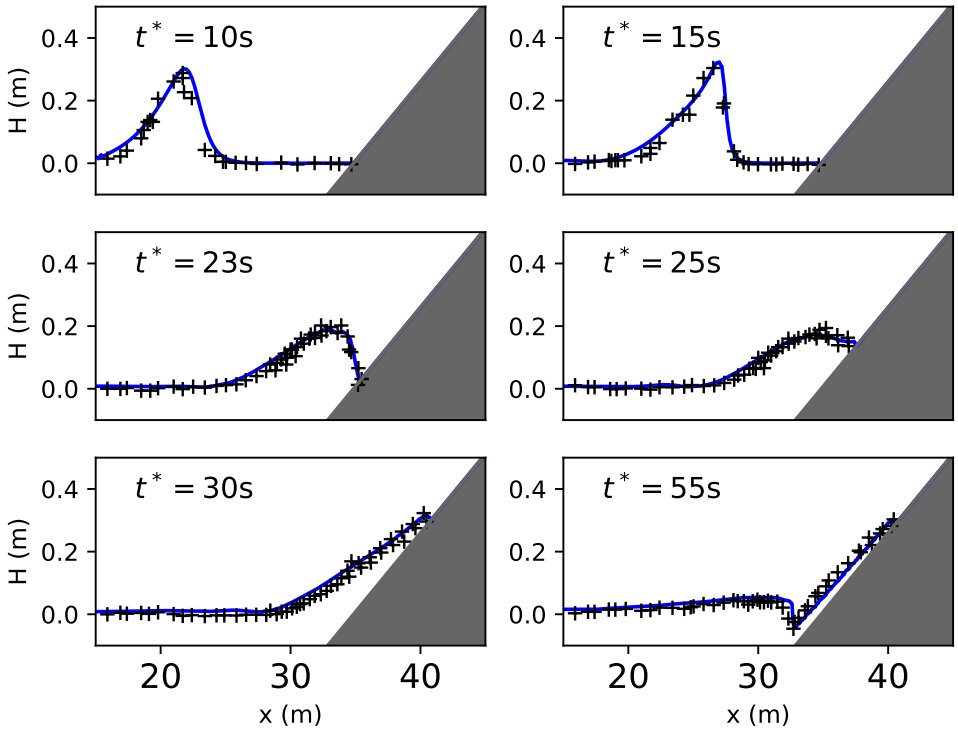


Figure 11: Test 8 - Shoaling of solitary waves: Synolakis *et al.* experiment - Free surface profiles (numerical results in solid lines, experimental data in cross) at several times during the propagation.

## 5 Conclusion

We introduce in this paper a new discontinuous Galerkin discrete formulation to approximate the solutions of Serre-Green-Naghdi (SGN) equations. A new non conforming discrete formulation belonging to the family of Symmetric Interior Penalty discontinuous Galerkin methods (SIP-DG) is introduced to accurately approximate the solutions of the second order elliptic operator occurring in the SGN equations in the surface horizontal $d = 1$ case. We show that the corresponding discrete bilinear form enjoys some consistency and coercivity properties. The global discrete formulation is extensively validated through an extended set of benchmarks, including some convergence studies for various polynomial orders of approximation and comparisons with experimental data for the reflexion of solitary waves over uneven topographies and nonlinear shoaling. Further studies will be devoted to the extension to SGN equations with vorticity based on the models of [36], to the surface horizontal $d = 2$ case, and to *a posteriori* limitations methods allowing to accurately handle the occurrence of dry areas within the framework of very high-order DG methods.

## Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgements** The authors wish to thank Pr. D.A. Di Pietro for fruitful discussions.

# References


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

[2] J.S. Antunes do Carmo. Boussinesq and Serre type models with improved linear dispersion characteristics: Applications. *J. Hydraulic Res.*, 51(6):719–727, 2013.

[3] J.S. Antunes do Carmo, F.J. Seabra-Santos, and A.B. Almeida. Numerical solution of the generalized Serre equations with the Mac-Cormack finite-difference scheme. *Int J Numer Methods Fluids*, 16(725-738), 1993.

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

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

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

[7] Q. Chen and I. Babuska. Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Comput. Methods Appl. Mech. Engrg*, 128:405–417, 1995.

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

[9] B. Cockburn and C.-W. Shu. Runge-Kutta Discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16(3):173–260, 2001.

[10] W. Craig, P. Guyenne, J. Hammack, D. Henderson, and C. Sulem. Solitary water wave interactions. *Physics of fluids*, 18(5), 2006.

[11] A.J.-C. de Saint-Venant. Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit. *C.R. Acad. Sci. Paris, Section Mécanique*, 73:147–154, 1871.

[12] M. Dehghan and M. Abbaszadeh. The solution of nonlinear Green-Naghdi equation arising in water sciences via a meshless method which combines moving kriging interpolation shape functions with the weighted essentially non–oscillatory method. *Nonlinear science and numerical simulation*, 68:220–239, 2018.

[13] D. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection. *SIAM J. Numer. Anal.*, 46(2):805–831, 2008.

[14] D. A. Di Pietro and A. Ern. Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations. *Math. Comp.*, 79(271):1303–1330, 2010.

[15] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69 of *Mathématiques and Applications*. Springer, 2012.

[16] D. A. Di Pietro and F. Marche. Weighted interior penalty discretization of fully nonlinear and weakly dispersive free surface shallow water flows. *J. Comput. Phys.*, 355:285–309, 2018.

[17] H. Dong and M. Li. A reconstructed central discontinuous galerkin-finite element method for the fully nonlinear weakly dispersive Green-Naghdi model. *Applied Numerical Mathematics*, 110:110–127, 2016.

[18] M. Dryja. On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85, 2003.

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

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

[21] A. Duran and F. Marche. A discontinuous Galerkin method for a new class of Green-Naghdi equations on unstructured simplicial meshes. *Appl. Math. Modelling*, 45:840–864, 2017.

[22] D. Dutykh, D. Clamond, P. Milewski, and D Mitsotakis. Finite volume and pseudo-spectral schemes for fully-nonlinear 1d Serre equations. *Eur. J. Appl. Math.*, 24:761–787, 2013.

[23] G.A. El, R.H.J. Grimshaw, and N.F. Smyth. Unsteady undular bores in fully nonlinear shallow-water theory,. *Phys. Fluids*, 18, 2006.

[24] N. Favrie and S. Gavrilyuk. A rapid numerical method for solving serre–green–naghdi equations describing long free surface gravity waves. *Nonlinearity*, 30(7), 2017.

[25] A.G. Filippini, M. Kazolea, and M. Ricchiuto. A flexible genuinely nonlinear approach for wave propagation, breaking and runup. *J. Comput. Phys.*, 310:381–417, 2016.

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

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

[28] S.T. Grilli, R. Subramanya, I.A. Svendsen, and J. Veeramony. Shoaling of solitary waves on plane beaches. *J. Wtrwy. Port Coast. and Oc. Engrg.*, 120(6):609–628, 1994.

[29] J.-L. Guermond, B. Popov, E. Tovar, and C. Kees. Robust explicit relaxation technique for solving the Green-Naghdi equations. *J. Comput. Phys.*, 2019.

[30] S. Guibourg. *Modélisation numérique et expérimentale des houles bidimensionnelles en zone cotière,*. PhD Thesis, Université Joseph Fourier-Grenoble I, 1994.

[31] S. Israwi. Large time existence for 1D Green-Naghdi equations. *Nonlinear Analysis*, 74(81-93), 2011.

[32] D. Kay, V. Styles, and E. Sulli. Discontinuous Galerkin finite element approximation of the Cahn-Hilliard equation with convection. *SIAM J. Numer. Anal.*, 47(4):2660–2685, 2009.

[33] D. Lannes. *The water waves problem: mathematical analysis and asymptotics.* Number 188 in Mathematical Surveys and Monographs. American Mathematical Society, 2013.

[34] D. Lannes and P. Bonneton. Derivation of asymptotic two-dimensional time-dependent equations for surface water wave propagation. *Physics of fluids*, 21:016601, 2009.

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

[36] D. Lannes and F. Marche. Nonlinear wave-current interactions in shallow water. *Stud. Appl. Math.*, 136(4):382–423, 2016.

[37] O. Le Métayer, S. Gavrilyuk, and S. Hank. A numerical scheme for the Green-Naghdi model. *J. Comput. Phys.*, 34(229):2034–2045, 2010.

[38] M. Li, P. Guyenne, F. Li, and L. Xu. High order well-balanced CDG-FE methods for shallow water waves by a Green-Naghdi model. *J. Comput. Phys.*, 257:169 – 192, 2014.

[39] M. Li, L. Xu, and Y. Cheng. A CDG-FE method for the two-dimensional Green-Naghdi model with the enhanced dispersive property. *J. Comput. Phys.*, 399, 2019.

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

[41] Per A Madsen, Russel Murray, and Ole R Sorensen. A new form of the Boussinesq equations with improved linear dispersion characteristics. *Coastal Engineering*, 15(4):371 – 388, 1991.

[42] F. Marche. Combined Hybridizable Discontinuous Galerkin (HDG) and Runge-Kutta Discontinuous Galerkin (RK-DG) formulations for Green-Naghdi equations on unstructured meshes. *J. Comput. Phys.*, 418:109637, 2020.

[43] D. Mitsotakis, B. Ilan, and D. Dutykh. On the Galerkin/Finite-Element Method for the Serre equations. *J. Sci. Comput.*, 61(1):166–195, 2014.

[44] D. Mitsotakis, C.E. Synolakis, and M. McGuinness. A modified Galerkin/finite element method for the numerical solution of the Serre-Green-Naghdi system. *Int. J. Numer. Meth. Fluids*, 83(10):755–778, 2017.

[45] N. Panda, C. Dawson, Y. Zhang, A.B. Kennedy, J.J. Westerink, and A.S. Donahue. Discontinuous Galerkin methods for solving Boussinesq-Green-Naghdi equations in resolving non-linear and dispersive surface water waves. *J. Comput. Phys.*, 273:572–588, 2014.

[46] J.D. Pearce and J.G. Esler. A pseudo-spectral algorithm and test cases for the numerical solution of the two-dimensional rotating Green-Naghdi shallow water equations. *J. Comput. Phys.*, 229:7594–7608, 2010.

[47] S. Popinet. A quadtree-adaptive multigrid solver for the Serre-Green-Naghdi equations. *J. Comput. Phys.*, 302:336–358, 2015.

[48] M. Ricchiuto and A.G. Filippini. Upwind Residual discretization of enhanced Boussinesq equations for wave propagation over complex bathymetries. *J. Comput. Phys.*, 271:306–341, 2014.

[49] G. L. Richard, A. Duran, and B. Fabrèges. A new model of shoaling and breaking waves. Part 2. run-up and two-dimensional waves. *J. Fluid Mech.*, 867:146–194, 2019.

[50] A. Samii and C. Dawson. An explicit hybridized discontinuous Galerkin method for Serre–Green–Naghdi wave model. *Comput. Methods Appl. Mech. Engrg*, 330:447–470, 2018.

[51] F Serre. Contribution à l'étude des écoulements permanents et variables dans les canaux. *Houille Blanche*, 8:374–388, 1953.

[52] M.K. Sharifian, G. Kesserwani, and Y. Hassanzadeh. A discontinuous Galerkin approach for conservative modelling of fully nonlinear and weakly dispersive wave transformations. *Ocean Modelling*, 2018.

[53] F. Shi, J.T. Kirby, J.C. Harris, J.D. Geiman, and S.T. Grilli. A high-order adaptive time-stepping TVD solver for Boussinesq modeling of breaking waves and coastal inundation. *Ocean Modelling*, 43-44:36–51, 2012.

[54] R.J. Spiteri and S.J. Ruuth. Non-linear evolution using optimal fourth-order strong-stability-preserving Runge-Kutta methods. *Math. Comput. Simulation*, 62(1-2):125–135, 2003.

[55] C.H. Su and C.S. Gardner. Korteweg-de-Vries equation and generalizations. III. Derivation of the Korteweg-de-Vries equation and Burgers equation. *J. Math. Phys.*, 10(3):536–539, 1969.

[56] C.E. Synolakis. The runup of solitary waves. *J. Fluid. Mech.*, 185:523—545, 1987.

[57] C.E. Synolakis, E.N. Bernard, V.V. Titov, U Kanoglu, and F.I. Gonzalez. Standards, criteria, and procedures for noaa evaluation of tsunami numerical models. *NOAA Tech. Memo.*, OAR PMEL-135, 2007.

[58] S. Tkachenko. *Analytical and numerical study of a dispersive shallow water model.* PhD thesis, Université Aix-Marseille, 2020.

[59] M. Walkley and M. Berzins. A finite element method for the two-dimensional extended Boussinesq equations. *Internat. J. Numer. Methods Fluids*, 39(10):865–885, 2002.

[60] G. Wei, J. Kirby, S. Grilli, and R. Subramanya. A fully nonlinear Boussinesq model for surface waves. Part 1. Highly nonlinear unsteady waves. *Journal of Fluid Mechanics*, 294:71–92, 1995.

[61] J.M. Witting. ï¿¼a unified model for the evolution of nonlinear water waves. *J. Comput. Phys.*, 56:203–236, 1984.

[62] Y. Xing and X. Zhang. Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes. *J. Sci. Comput.*, 57:19–41, 2013.

[63] Y. Xing, X. Zhang, and C.-W. Shu. Positivity-preserving high order well-balanced discontinuous galerkin methods for the shallow water equations. *Advances in Water Resources*, 33(12):1476 – 1493, 2010.

[64] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comput. Phys.*, 229(9):3091 – 3120, 2010.

[65] Y. Zhang, A.B. Kennedy, T. Tomiczek, A.S. Donahue, and J.J. Westerink. Validation of Boussinesq–Green–Naghdi modeling for surf zone hydrodynamics. *Ocean Engineering*, 111:299–309, 2016.