A posteriori Finite-Volume local subcell correction of high-order discontinuous Galerkin schemes for the nonlinear shallow-water equations

Ali Haidar, Fabien Marche, François Vilar

A posteriori Finite-Volume local subcell correction of high-order discontinuous Galerkin schemes for the nonlinear shallow-water equations

Ali Haidar[a,*], Fabien Marche[a], Francois Vilar[a]

[a] *IMAG, Univ Montpellier, CNRS, Montpellier, France*

Abstract

We design and analyze a new discretization method for the nonlinear shallow water equations, which is based on an equivalent representation of arbitrary high-order Discontinuous Galerkin (DG) schemes through piecewise constant modes on a sub-grid, together with a selective *a posteriori* local correction of the sub-interface reconstructed flux. This new approach, based on [F. Vilar, J. Comput. Phys., 387:245-279, 2019], allows to combine at the subcell scale the excellent robustness properties of the Finite-Volume (FV) lowest-order method and the high-order accuracy of the DG method. For any order of polynomial approximation, the resulting algorithm is shown to: (i) accurately handle strong shocks with no robustness issues; (ii) ensure the preservation of the water height positivity at the subcell level; (iii) preserve the class of motionless steady states (well-balancing); (iv) retain the highly accurate subcell resolution of DG schemes. These assets are numerically illustrated through an extensive set of test-cases, with a particular emphasize put on the use of very-high order polynomial approximations on coarse grids.

Keywords: Non-linear shallow water equations, discontinuous Galerkin scheme, local *a posteriori* subcell FV correction, well-balancing, positivity preservation, arbitrary high-order

*Corresponding author

Email addresses: ali.haidar@umontpellier.fr (Ali Haidar), fabien.marche@umontpellier.fr (Fabien Marche), francois.vilar@umontpellier.fr (Francois Vilar)

1. Introduction

The nonlinear shallow water (NSW) equations (31) are one of the most widely used set of equations for simulating long wave hydrodynamics, under the assumption that the vertical acceleration of the fluid can be neglected. Given a smooth parametrization of the topography variations $b : \mathbb{R} \to \mathbb{R}$, and denoting by H the water height, u the horizontal (depth-averaged) velocity and $q = Hu$ the horizontal discharge (see Fig.1), the NSW equations may be written as follows:

$$\partial_t H + \partial_x q = 0, \tag{1a}$$

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

Considering their hydrostatic nature, in comparison to the dispersive nature of more sophisticated models such as the Boussinesq-type models, hyperbolic integral forms of the NSW equations generally provide an accurate representation of steep-fronted flows, such as dam breaks, flood waves or bores propagation in the surf zone. This model is also extensively used in coastal engineering, for the study of nearshore flows involving run-up and run-down on sloping beaches or coastal structures and to forecast coastal inundations.

To allow a proper description of such phenomena, accurate and robust numerical methods have to be considered. Great efforts have been made since the sixties in order to produce accurate approximations of weak solutions of the NSW equations and a large variety of numerical methods have been developed, including Finite-Volumes (FV) (2; 43; 5; 10; 104; 39), Finite-Elements (FE) (77; 96; 83; 9), spectral methods (55; 71; 80) or residual distribution methods (88; 87; 6). Among these numerical strategies, the Godunov-type FV methods are particularly praised, thanks to their low computational cost and their shock-capturing ability, which allows to preserve the discontinuous or steeply varying gradients that may occur in sharp-fronted and trans-critical shallow water flows, see for instance (89; 5; 95; 32; 79; 13; 69) among others and also some references herein. Many of them particularly focus on the issue of balancing the flux gradient and the topography source source term (7; 78; 44; 70; 69; 20; 60; 23; 74). However, FV methods usually offer low accuracy and one generally needs to use some reconstruction methods to offset the low order of convergence and the diffusive losses, see for instance (58; 64; 84).

In recent years, high-order discontinuous Galerkin (DG) methods have become very popular to approximate the solutions of various linear and nonlinear partial differential equations. Considering the approximation of hyperbolic conservation laws, DG methods combine the background of FE methods, FV methods and Riemann solvers, allowing to take into account the physic of the problem, and they have been successfully validated in many domain of applications. An arbitrary order of accuracy in space can be obtained with the use of high-order polynomials within elements, allowing to keep the stencil compact, along with being able to handle complex geometries through the use of unstructured general meshes and h/p-adaptivity. Moreover, they are highly parallelizable, and exhibit local conservation and strong stability properties. We refer the reader to (28; 29) for a general background. Several DG methods have been designed for the NSW equations since the early 2000s, see for instance (65; 106; 1; 4; 59; 12; 76; 90; 46; 41; 108; 107; 61) and some references hereafter.

However, while DG methods may be mature enough to accurately handle some realistic problems in various applications, they still suffer from the lack of nonlinear stability. In particular, high-order DG methods may produce spurious oscillations in the presence of discontinuities or steeply varying gradients, *i.e.* Gibbs phenomenon. These nonphysical overshoots may lead to nonphysical

solutions. Another challenging issue encountered in the approximation of the NSW equations is the preservation of the water height positivity, closely related to the issue of the occurrence and propagation of wet/dry fronts that may occur in dam breaks, flood waves or run-up over coastal shores. Considering the convex set of physical admissible states Θ, defined as follows:

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

a minimal nonlinear stability requirement is to ensure that this set is preserved at the discrete levels. But the use of high-order polynomials is generally not straightforwardly compatible with such a requirement and standard numerical methods may produce negative values for the water height H.

Generally speaking, robustness issues may be among the main remaining challenges for the use of high-order methods in realistic problems for many domains of applications, and in recent years, several approaches have been proposed to stabilize high-order approximations. These techniques mainly rely on two different paradigms that we referred to as *a priori* and *a posteriori*. In the so-called *a priori* framework, the correction procedure is applied before advancing the numerical piecewise polynomial solution further in time. So first, a troubled zone indicator is used to find where a correction is required (see (85) for a review of such *troubled elements* sensors). Then, sufficient efforts are made on the numerical solution or on the numerical scheme to be sure that one will be able to carry the computation out to the next time step. Among others *a priori* correction techniques, we could mention *artificial viscosity* methods (81; 97; 42; 49; 62), where some dissipative mechanism is added in shock regions, borrowing ideas from the streamline upwind Petrov Galerkin (SUPG) and Galerkin least-squares methods. Some other very popular limiting techniques can be gathered and referred to as slope and moment limiters (26; 14; 17; 63; 111; 57; 35; 66). In the former ones, as in (27; 26), the polynomial approximated solution is flattened around its mean value to control the solution jumps at cell interfaces. A smooth extrema detector is then generally used to prevent the limitation technique to spoil the accuracy in regions where no limiting is required. Moments limiters, mainly based on (14) and further developed in (17), can be seen as the extension of the aforementioned slope limiters to the case of very high orders of accuracy. In those limiting strategies, the different moments of the polynomial solution are successively scaled in a decreasing sequence, from the higher degree to the lower one, allowing the preservation of the solution accuracy, as well as ensuring the solution boundedness near discontinuities. The high-order DG limiter (63), generalized moment limiter (111), hierarchical Multi-dimensional Limiting Process (MLP) (57; 56) and vertex-based hierarchical slope limiters (35; 66) all derive from (27; 14; 17), and thus fall into this category. Now, another limiting strategy that deserves to be mentioned is the (H)WENO limiting procedure (86; 8; 115; 67; 116), where the DG polynomial is substituted in troubled regions by a reconstructed (H)WENO polynomial. An alternative way to treat this spurious oscillations issue may be to use a *solution filtering* method, see for instance (98; 94; 75; 82), which aim at removing high wave-number oscillations. Those filtering procedures are generally done in an *ad hoc* fashion, filtering being applied "as little as possible, but as much as needed". Last but not least, some original subcell FV shock capturing techniques in the frame of DG schemes (51; 22; 92; 30) have recently gained in popularity. In (51), the authors use a convex combination between high-order DG schemes and first-order finite volumes on a sub-grid, allowing them to retain the very high accurate resolution of DG in smooth areas and ensuring the scheme robustness in the presence of shocks. Similarly, in (92; 30), after having detected the troubled zones, cells are then subdivided into subcells, and a robust first-order finite volume scheme is performed on the sub-grid in troubled cells.

The *a priori* paradigm has already and extensively proved in the past its high capability and feasibility, as in the aforementioned articles. Those techniques are *a priori* in the sense that only the data at time t^n are needed to perform the limitation procedure. Then, the limited solution is used to advance the numerical scheme in time to t^{n+1}. The "worst case scenario" has to be generally considered as a precautionary principle. Furthermore, let us emphasize that most of the *a priori* correction procedures previously quoted do not ensure a maximum principle or the positivity-preservation of the solution. Generally, additional effort has to been made specifically on that matter, as for example by means of positivity-preserving limiters (112; 114). Specifically concerning the issue of positivity preservation in DG methods for the NSW equations, various *a priori* strategies have been introduced recently: a free-boundary treatment in mixed Eulerian-Lagrangian elements is introduced in (15) to locate the wet/dry interface, a fixed mesh method with a local conservative slope modification technique based on a redistribution of the fluid and cut-off in discharge is presented in (40), local first moment limitations without mass adding in (16; 60; 59; 99), high-order accuracy *a priori* polynomial reconstruction and limitation to enforce a strict maximum principle on mean values in (110; 109), in (38) for the so-called *pre-balanced* formulation of the NSW equations, or in (72) for a formulation with implicit time stepping. Let us finally mention (73) where a Finite-Volume subcell approaches has been adopted.

Now, the paradigm of *a posteriori* correction is different in the way that first an uncorrected candidate solution is computed at the new time step. The candidate solution is then checked according to some criteria (for instance positivity, discrete maximum principle, ...). If the solution is considered admissible, we go further in time. Otherwise, we return to the previous time step and correct locally the numerical solution by making use of a more robust scheme. Because the troubled zone detection is performed *a posteriori*, the correction can be done only where it is absolutely necessary. Furthermore, let us emphasize that in *a posteriori* correction procedures, the maximum principle preservation or positivity preservation is included without any additional effort. Indeed, the whole procedure will be positivity-preserving as soon as the numerical scheme used as a correction procedure is. Consequently, all the *a posteriori* techniques that make use of FV scheme as correction method will then be positivity-preserving. Recently, some new *a posteriori* limitations have arisen. Let us mention the so-called MOOD technique, (25; 33; 34). Through this procedure, the order of approximation of the numerical scheme is locally reduced in an *a posteriori* sequence until the solution becomes admissible. In (37; 36; 54), a subcell FV technique similar to the one presented in (92) has been applied to the *a posteriori* paradigm. Practically, if the numerical solution in a cell is detected as bad, the cell is then subdivided into subcells and a first-order FV, or alternatively other robust scheme (second-order TVD FV scheme, WENO scheme, ...), is applied on each subcell. Then, through these new subcell mean values, a high-order polynomial is reconstructed on the primal cell. Related strategies applied to dispersive and turbulent shallow water flows have been introduced in (18; 19).

Nonetheless, in all the aforementioned limitation techniques, *a priori* and *a posteriori*, in the troubled cells the high-order DG polynomial is either globally modified in the cell, or even discard as it is in the (H)WENO limiter or any *a posteriori* correction procedure. One of the main advantage of high-order scheme is to be able to use coarse grids while still being very precise. But even in the case where the troubled zone, as the vicinity a shock for instance, is very small regarding the characteristic length of a cell, the DG polynomial will be globally modified. In (101), we have introduced a new conservative technique to overcome this issue, by modifying the DG numerical

solution only locally at the subcell scale. This correction procedure has been designed first to avoid the occurrence of non-admissible solution, to be maximum principle preserving, or in the context of systems positivity-preserving, and to prevent the code from crashing (for instance avoiding NaN in the code). The corrected scheme is also conservative, at the subcell level. Secondly, the correction permits to essentially avoid the appearance of spurious oscillations. Thirdly, it allows us to retain as much as possible the high accuracy and subcell resolution of DG schemes, by minimizing the number of subcells in which the solution has to be recomputed. Practically, the correction procedure presented in (101) only modifies the DG solution in troubled subcell regions without impacting the solution elsewhere in the cell. It is also worth mentioning that the whole procedure is totally parameter free, and behaves properly from 2nd order to any order of accuracy.

Making use of the *a posteriori* local subcell correction introduced (101), the main objective of this paper is to develop a novel shock-capturing, positivity-preserving and well-balanced DG method for the NSW equations equations with topography source term by using specific local flux correction at the subcell level, with *a posteriori* numerical admissibility detectors. The well-balanced property for NSW equations, first introduced in (11), has been widely studied in recent years. Following the ideas of (38), we use the so-called *pre-balanced* formulation of the NSW equations. Indeed, the alternative formulation of the NSW equations in deviatoric form, obtained by subtracting an equilibrium solution, and introduced in (89) is interesting as it leads to a balanced set of hyperbolic equations that does not require specific numerical algorithms to obtain a well-balanced property. However, such a formulation is given in terms of free surface elevation above the still water level, and is therefore not suitable to model cases involving dry areas (the still water depth is undefined in dry areas). In (68; 69), a new formulation given in terms of the total free surface elevation $\eta = H + b$ (see Fig. 1) allows to alleviate this drawback.

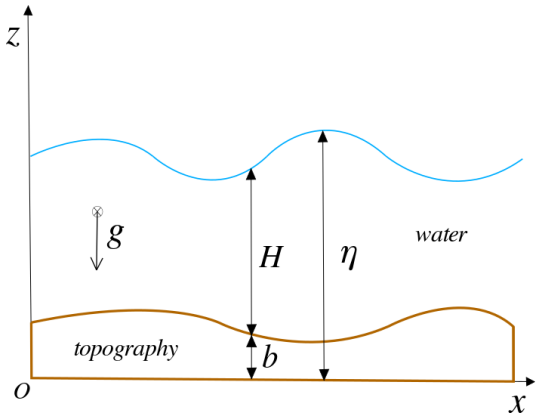

Figure 1: Free surface flow: main notations

Indeed, observing that

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

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

$$\partial_t \mathbf{v} + \partial_x \mathbf{F}(\mathbf{v}, b) = \mathbf{B}(\mathbf{v}, \partial_x b), \tag{3}$$

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

$$\mathbf{v} = \begin{pmatrix} \eta \\ q \end{pmatrix}, \ \mathbf{F}(\mathbf{v}, b) = \begin{pmatrix} q \\ uq + \frac{1}{2} g(\eta^2 - 2\,\eta\, b) \end{pmatrix}, \quad \mathbf{B}(\mathbf{v}, \partial_x b) = \begin{pmatrix} 0 \\ -g\,\eta\,\partial_x b \end{pmatrix}. \tag{4}$$

In the next section, we introduce a new discrete formulation with an arbitrary order of accuracy for equations (4), following a classical DG formalism. Then, exploiting the fact that such DG formulation can be regarded as a FV-like scheme on a sub-partition with particular (reconstructed) high-order interface fluxes, we show that it is possible to slightly and locally modify these interface fluxes in order to enforce some nonlinear stability property through the use of some *a posteriori* admissibility sensors. We also show that such an approach is fully compatible with the preservation of motionless steady states at the subcell level, provided that local mean-values of the solution on the sub-partition are carefully reconstructed, adapting the ideas of (69; 38). In the third section, we assess the ability of this new method, called *a posteriori* Local Subcell Correction (LSC) in the following, to remove nonphysical instabilities in the vicinity of discontinuities, and ensure that the mean values of the water height on the sub-partition remain positive. We highlight, on several test cases, the particularly interesting subcell resolution ability of the resulting scheme.

2. Discrete formulations

2.1. Discrete settings

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

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

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

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

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

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

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

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

Note that we have:

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

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

$$(v,w)_{\mathscr{T}_h} := \sum_{\omega\in\mathscr{T}_h} (v,w)_\omega, \qquad (v,w)_\omega := \int_\omega v(x)w(x)dx, \quad \forall \omega \in \mathscr{T}_h,$$
$$[v]_{\partial\omega_i} := v(x_{i+\frac{1}{2}}) - v(x_{i-\frac{1}{2}}), \quad \forall \omega_i \in \mathscr{T}_h.$$

For $\omega \in \mathscr{T}_h$, we denote p_ω^k the L^2-orthogonal projector onto $\mathbb{P}^k(\omega)$ and $p_{\mathscr{T}_h}^k$ the L^2-orthogonal projector onto $\mathbb{P}^k(\mathscr{T}_h)$. Similarly, we denote I_ω^k the element nodal interpolator into $\mathbb{P}^k(\omega)$. The corresponding nodal distributions in elements are chosen to be the approximate optimal nodes introduced in (24), which have better approximation properties than equidistant distributions, and include, for each element, the elements boundaries into the interpolation nodes. The global $I_{\mathscr{T}_h}^k$ interpolator into $\mathbb{P}^k(\mathscr{T}_h)$ is obtained by gathering the local interpolating polynomials defined on each elements.

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

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

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

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

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

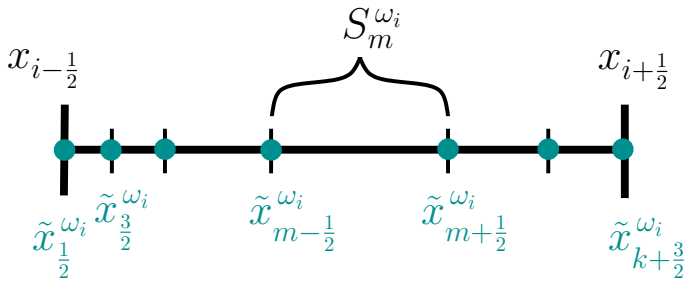

Figure 2: Partition of a mesh element ω_i in $k+1$ subcells

To define the *sub-resolution* basis functions, required in § 2.3 and initially introduced in (101), we introduce for a given mesh element $\omega \in \mathscr{T}_h$ the following set of *subcell indicator* functions $\left\{\mathbb{1}_m^\omega, \quad m \in [\![1,\, k+1]\!]\right\}$, with:

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

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

$$\phi_m^\omega = p_\omega^k(\mathbb{1}_m^\omega), \quad \forall m \in [\![1,\, k+1]\!]. \tag{5}$$

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

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

Concerning time discretization, for a given final computational time $t_{\max} > 0$, we consider a partition $(t^n)_{0\le n\le N}$ of the time interval $[0, t_{\max}]$ with $t^0 = 0$, $t^N = t_{\max}$ and $t^{n+1} - t^n =: \Delta t^n$. More details on the computation of the time step Δt^n and on the time marching algorithms are given in § 2.4. For any sufficiently regular scalar-valued function of time w, we let $w^n := w(t^n)$.

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

$$\underline{v}_j^\omega \quad \text{with} \quad j \in [\![1,\, k+1]\!] \ \text{ and } \ \omega \in \mathscr{T}_h, \tag{6}$$

such that

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

With a little abuse in terminology, we refer hereafter to the real numbers (6) as the *degrees of freedom* associated with v_h and we note $\underline{v}_\omega \in \mathbb{R}^{k+1}$ the vector gathering the degrees of freedom associated with v_ω.

Remark 2. For any $\omega \in \mathscr{T}_h$, and any $v_\omega \in \mathbb{P}^k(\omega)$, let denote

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

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

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

$$\pi_{m,p}^\omega = \frac{1}{|S_m^\omega|} \int_{S_m^\omega} \psi_p^\omega \mathrm{d}\mathrm{x}, \qquad \forall\, (m,p) \in [\![1,\, k+1]\!]^2, \tag{7}$$

we have the following relation:

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

As a consequence, any polynomial function $v_h \in \mathbb{P}^k(\omega)$ can be expressed equivalently either in terms of the degrees of freedom $\underline{v}_\omega$, or the sub-means values $\overline{v}_\omega$. Finally, let introduce the (one-to-one) following projector onto the piecewise constant sub-grid space:

$$\pi^k_{\mathscr{T}_\omega} : \mathbb{P}^k(\omega) \longrightarrow \mathbb{P}^0(\mathscr{T}_\omega) \tag{8}$$

$$v_\omega \longmapsto \pi^k_{\mathscr{T}_\omega}(v_\omega) := \overline{v}_\omega. \tag{9}$$

In practice, once the transformation matrices $\mathbf{\Pi}_\omega$ are initialized in a preprocessing step, it is straightforward and computationally inexpensive to switch from one representation to another, see Fig. 3.

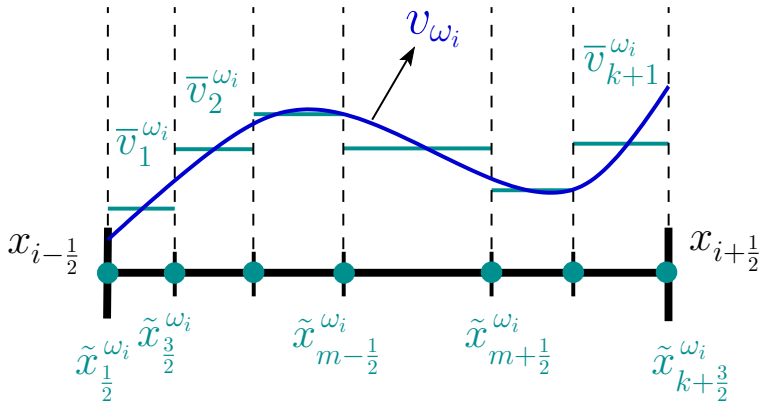

Figure 3: Piecewise polynomial function and associated sub-mean values

Remark 3. The local projection matrices (7) are obviously non-singular. This property would be straightforwardly extended to the multi-dimensional case with Cartesian grids.

2.2. Discontinuous Galerkin formulation

Let $b_h = I^k_{\mathscr{T}_h}(b)$ denote a globally continuous piecewise polynomial approximation of the topography parametrization and let denote $\nabla b_h = \nabla^k_h b_h$, for the sake of simplicity. A straightforward semi-discrete in space DG approximation of (3) reads: find $\mathbf{v}_h = (\eta_h, q_h) \in (\mathbb{P}^k(\mathscr{T}_h))^2$ such that, for all $\varphi \in \mathbb{P}^k(\mathscr{T}_h)$,

$$\big(\partial_t \mathbf{v}_h, \varphi\big)_{\mathscr{T}_h} + \big(\mathcal{A}_h(\mathbf{v}_h), \varphi\big)_{\mathscr{T}_h} = 0, \tag{10}$$

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

$$\big(\mathcal{A}_h(\mathbf{v}_h), \varphi\big)_{\mathscr{T}_h} := -\sum_{\omega\in\mathscr{T}_h} \big(\mathbf{F}(\mathbf{v}_h, b_h), \partial_x\varphi\big)_\omega + \sum_{\omega\in\mathscr{T}_h} \big[\varphi\,\mathcal{F}\big]_{\partial\omega} - \big(\mathbf{B}(\mathbf{v}_h, \nabla b_h), \varphi\big)_{\mathscr{T}_h}, \quad \forall\varphi \in \mathbb{P}^k(\mathscr{T}_h). \tag{11}$$

In (11), $\mathcal{F}$ stands for the interface numerical flux function. Denoting by $\mathbf{v}^-_{i+\frac{1}{2}}$ and $\mathbf{v}^+_{i+\frac{1}{2}}$, respectively the left and right traces of $\mathbf{v}_h$ on interface $x_{i+\frac{1}{2}}$, and by $b_{i+\frac{1}{2}} = b^-_{i+\frac{1}{2}} = b^+_{i+\frac{1}{2}}$ the trace of b_h, we define the numerical flux function $\mathcal{F}_{i+\frac{1}{2}}$ on interface $x_{i+\frac{1}{2}}$ as follows:

$$\mathcal{F}_{i+\frac{1}{2}} := \mathcal{F}(\mathbf{v}^-_{i+\frac{1}{2}}, \mathbf{v}^+_{i+\frac{1}{2}}, b_{i+\frac{1}{2}}), \tag{12}$$

where the numerical flux function chosen here is the simple global Lax-Friedrichs flux:

$$\mathcal{F}(\mathbf{v}^-, \mathbf{v}^+, b) := \frac{1}{2}\left(\mathbf{F}(\mathbf{v}^-, b) + \mathbf{F}(\mathbf{v}^+, b) - \sigma(\mathbf{v}^+ - \mathbf{v}^-)\right), \tag{13}$$

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

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

Remark 4. We require that the volume integral and source term in formula (11) are exactly computed at motionless steady states. This can be achieved, for the pre-balanced formulation (3)-(4), by using any quadrature rule exact for polynomials of degree up to $2k$, thanks to the pre-balanced formulation (3)-(4). On the other hand, for the classical NSW formulation (1), a quadrature rule exact for polynomials of degree up to $3k$ is needed.

Remark 5. The topography approximation b_h is interpolated by a piecewise polynomial but globally continuous function over the mesh. To achieve this, one can simply choose the elements boundaries among the interpolation points with any corresponding interpolation method. To ensure that the scheme is indeed well-balanced, and particularly in wet/dry context, see § 3.4, we initialize the surface elevation η_h in dry areas by setting $\eta_h = b_h$. Thus, water height positivity is also assured in dry areas since $h_h = \eta_h - b_h = 0$, by construction. We emphasize that as long as $h = \eta - b$ is non-negative, its subcell mean-values are also non-negative. However, nothing ensures that after performing a L^2 projection of the initial water height, the associated submean values of the L^2 projection h_ω are positive. This is the reason why, in wet regions, for the initialization, we start by computing the positive h submean values using (14) and then reconstruct the associated polynomial using $\mathbf{\Pi}_\omega^{-1}$.

$$\overline{h}_m^\omega = \frac{1}{|S_m^\omega|} \int_{S_m^\omega} h(x)dx. \tag{14}$$

In the following, similarly to what has been done in (101), we demonstrate an equivalence relation between (10) and a FV-like method on a sub-mesh.

2.3. DG formulation as a FV-like scheme on a sub-grid

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

$$\left(\mathbf{F}(\mathbf{v}_h, b_h), \varphi\right)_{\mathscr{T}_h} = \left(\mathbf{F}_h, \varphi\right)_{\mathscr{T}_h}, \quad \forall \varphi \in \mathbb{P}^k(\mathscr{T}_h), \tag{15a}$$

$$\left(\mathbf{B}(\mathbf{v}_h, \nabla b_h), \varphi\right)_{\mathscr{T}_h} = \left(\mathbf{B}_h, \varphi\right)_{\mathscr{T}_h}, \quad \forall \varphi \in \mathbb{P}^k(\mathscr{T}_h). \tag{15b}$$

Remark 6. As we mentioned, in DG schemes, volume integral and source term contribution are computed using quadrature rule. This quadrature rule should be used to compute the left hand side of the L^2 projections (15a) and (15b).

From (10), we now have:

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

or equivalently, using an integration by parts:

$$\left(\partial_t \mathbf{v}_h, \varphi\right)_{\mathscr{T}_h} + \sum_{\omega\in\mathscr{T}_h} \left(\partial_x \mathbf{F}_h, \varphi\right)_\omega - \sum_{\omega\in\mathscr{T}_h} \left[\varphi\left(\mathbf{F}_h - \mathcal{F}\right)\right]_{\partial\omega} - \left(\mathbf{B}_h, \varphi\right)_{\mathscr{T}_h} = 0, \quad \forall \varphi \in \mathbb{P}^k(\mathscr{T}_h). \tag{16}$$

Substituting ϕ_m^ω, defined in (5), into (16) gives the local equations on mesh element $\omega \in \mathscr{T}_h$:

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

Since $\partial_t \mathbf{v}_\omega$, $\partial_x \mathbf{F}_\omega$ and $\mathbf{B}_\omega$ belong to $\left(\mathbb{P}^k(\omega)\right)^2$ and considering (5), it follows that

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

where $\overline{\mathbf{v}}_m^\omega$ and $\overline{\mathbf{B}}_m^\omega$ are respectively the mean values of $\mathbf{v}_\omega$ and $\mathbf{B}_\omega$ on the subcell S_m^ω. Let introduce the $k+2$ subcells interfaces fluxes $\left\{\widehat{\mathbf{F}}_{m+\frac{1}{2}}^{\,\omega}\right\}_{m\in[\![0,\,k+1]\!]}$ such that:

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

so that we have

$$\partial_t \overline{\mathbf{v}}_m^\omega = -\frac{1}{|S_m^\omega|}\left(\widehat{\mathbf{F}}_{m+\frac{1}{2}}^{\,\omega} - \widehat{\mathbf{F}}_{m-\frac{1}{2}}^{\,\omega}\right) + \overline{\mathbf{B}}_m^\omega. \tag{17}$$

Formulation (17) can be seen as a FV-like scheme on subcell S_m^ω. The values $\left\{\widehat{\mathbf{F}}_{m+\frac{1}{2}}^{\,\omega}\right\}_{m\in[\![0,\,k+1]\!]}$ will be thereafter referred to as *reconstructed fluxes*. Considering the mesh element $\omega_i = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right] \in \mathscr{T}_h$, and setting the first and last reconstructed fluxes to the DG numerical flux values at cell boundaries such as:

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

the m interior reconstructed fluxes expression is then given by:

$$\widehat{\mathbf{F}}_{m+\frac{1}{2}}^{\,\omega_i} = \mathbf{F}_{\omega_i}\left(\widetilde{x}_{m+\frac{1}{2}}^{\,\omega_i}\right) - C_{m+\frac{1}{2}}^{i-\frac{1}{2}}\left(\mathbf{F}_{\omega_i}\left(x_{i-\frac{1}{2}}\right) - \mathcal{F}_{i-\frac{1}{2}}\right) - C_{m+\frac{1}{2}}^{i+\frac{1}{2}}\left(\mathbf{F}_{\omega_i}\left(x_{i+\frac{1}{2}}\right) - \mathcal{F}_{i+\frac{1}{2}}\right), \quad \forall m \in [\![1,\, k+1]\!], \tag{18}$$

where the $C_{m+\frac{1}{2}}^{i\pm\frac{1}{2}}$ are explicitly computed in (101). Their expression is recalled here for the sake of completeness:

$$C_{m+\frac{1}{2}}^{i-\frac{1}{2}} = \sum_{p=m+1}^{k+1} \phi_p^{\omega_i}\left(x_{i-\frac{1}{2}}\right) \qquad \text{and} \qquad C_{m+\frac{1}{2}}^{i+\frac{1}{2}} = \sum_{p=1}^{m} \phi_p^{\omega_i}\left(x_{i+\frac{1}{2}}\right). \tag{19}$$

Remark 7. One can see that the reconstructed flux is nothing but the polynomial interior flux $\mathbf{F}_\omega$, plus some correction terms taking into account the jump in fluxes at cell boundary $\partial\omega$. A simple explicit expression of the correction coefficients (19) only depending on k and the subcells interfaces coordinates $\widetilde{x}_{m+\frac{1}{2}}^{\,\omega}$ is given in (101).

Remark 8. Let us note that if this particular definition of $\left\{C^{i\pm\frac{1}{2}}_{m+\frac{1}{2}}\right\}_{m\in[\![0,\,k+1]\!]}$, (19), gives the equivalence with DG schemes, other choices obviously lead to other schemes. For instance, if one set these constants to zero, except for the first and last to be one, one would then recover the spectral volume method, (105; 50). Let us also mention Flux Reconstruction schemes (FR), also referred to as Correction Procedure via Reconstruction (CPR), with which we share this reconstructed fluxes framework, see for instance (52; 103; 3; 45; 53) or the dedicated paragraph in (101) for more insight on the analogy between the present theory and Flux Reconstruction schemes.

Remark 9. The choice of the sub-partition points, $\left\{\widetilde{x}^{\,\omega}_{m+\frac{1}{2}}\right\}_{m\in[\![0,\,k+1]\!]}$, has already been discussed in (101). It appeared that, regarding the reformulation of DG schemes into subcell Finite-Volume methods, the cell decomposition into subcells does not come into account, as any choice would lead to the same piecewise polynomial solution. However, for the correction procedure introduced in § 3, the sub-division does has a slight impact. Indeed, the use of a non-uniform sub-partition, for instance by means of the Gauss-Lobatto points, leads to better results compared to a uniform sub-division. This is more likely the manifestation of the Runge phenomenon in the context of histopolation, as the histopolation basis functions underlying the sub-mean value representation, are more oscillatory for a uniform cell sub-partition. Consequently, in the remainder, we make use of Gauss-Lobatto points to define the sub-partition points $\left\{\widetilde{x}^{\,\omega}_{m+\frac{1}{2}}\right\}_{m\in[\![0,\,k+1]\!]}$.

2.4. Time marching algorithm

Supplementing (10) with an initial datum $\mathbf{v}(0,\cdot)=\mathbf{v}_0=(\eta_0,q_0)^{\mathrm{t}}$, the time stepping may be carried out using explicit SSP-RK schemes, (47; 91). For instance, writing the semi-discrete equation (10) in the operator form

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

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

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

where ${\mathbf{v}_h}^{n,i}$, $1\leq i\leq 2$, are the solutions obtained at intermediate stages, Δt^n is obtained from the CFL condition (20), and the discrete initial data ${\mathbf{v}_h}^{0}$ is defined as the L^2 projection of the initial datum (see remark 5 for details). As the correction described in the following section make use of both DG scheme on the primal cells $\omega\in\mathscr{T}_h$ and FV scheme on the subcells $S^\omega_m\in\mathscr{T}_\omega$, the time step Δt^n is computed adaptively using the following CFL condition:

$$\Delta t^n = \frac{\min\limits_{\omega\in\mathscr{T}_h}\left(\dfrac{h_\omega}{2k+1},\ \min\limits_{S^\omega_m\in\mathscr{T}_\omega}|S^\omega_m|\right)}{\sigma}, \tag{20}$$

where σ is the constant previously introduced in the global Lax-Friedrichs numerical flux definition, (13).

3. *A posteriori* local subcell correction

In this section, we show how it is possible to modify the reconstructed fluxes $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ in a robust way in subcells where the *uncorrected* DG scheme (10) has failed, either by obtaining negative value for the water height or by generating nonphysical oscillations due to the Gibbs phenomenon in the vicinity of discontinuities. For sake of conciseness in notations, the superscript ω_i may be avoided in the following when no confusion is possible.

As high-order Runge-Kutta SSP time marching algorithms may be regarded as convex combinations of first-order forward Euler schemes, we consider in the following, for sake of simplicity, a fully discrete formulation obtained from (10) and a first-order forward Euler scheme. We assume that at time level n the numerical solution $\mathbf{v}_h^n$ is *admissible* in a sense to be clarified later. We then compute an updated candidate solution $\mathbf{v}_h^{n+1}$ through the *uncorrected* DG scheme (10). If the candidate $\mathbf{v}_h^{n+1}$ is admissible, no correction is needed. Otherwise, the *uncorrected* DG scheme has produced an updated solution $\mathbf{v}_h^{n+1}$ which is not admissible on at least one particular mesh element cell ω_i. Looking at the subcell level, and assuming that $\mathbf{v}_\omega^{n+1}$ is not admissible in the particular subcell $S_m \in \mathscr{T}_{\omega_i}$, which is thus called a *troubled subcell* in the following, the main idea of our *a posteriori* Local Subcell Correction (LSC) is to replace the incriminated subcell mean value $\overline{\mathbf{v}}_m^{n+1}$ by a new one, denoted with a $\star$ as follows $\overline{\mathbf{v}}_m^{\star,n+1}$, which is computed using a subcell first-order FV scheme of the form:

$$\overline{\mathbf{v}}_m^{\star,n+1} = \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|}\left(\mathcal{F}^l_{m+\frac{1}{2}} - \mathcal{F}^r_{m-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_m, \tag{21}$$

with some new subcell *lowest-order corrected* numerical fluxes $\mathcal{F}^l_{m+\frac{1}{2}}$, $\mathcal{F}^r_{m-\frac{1}{2}}$ which are defined hereafter. Indeed, because the *uncorrected* DG scheme (10) is equivalent to the subcell FV-like scheme (17) with high-order reconstructed fluxes (18), we propose to substitute, at the boundaries of S_m, the high-order reconstructed fluxes with first-order FV numerical fluxes. Finally, new degrees of freedom at discrete time t^{n+1} are computed from the modified set of sub-mean values, now given as a blend of uncorrected values $\overline{\mathbf{v}}_m^{n+1}$ and corrected values $\overline{\mathbf{v}}_m^{\star,n+1}$. This strategy is illustrated in Fig. 4, where the marked subcell is identified with red color.

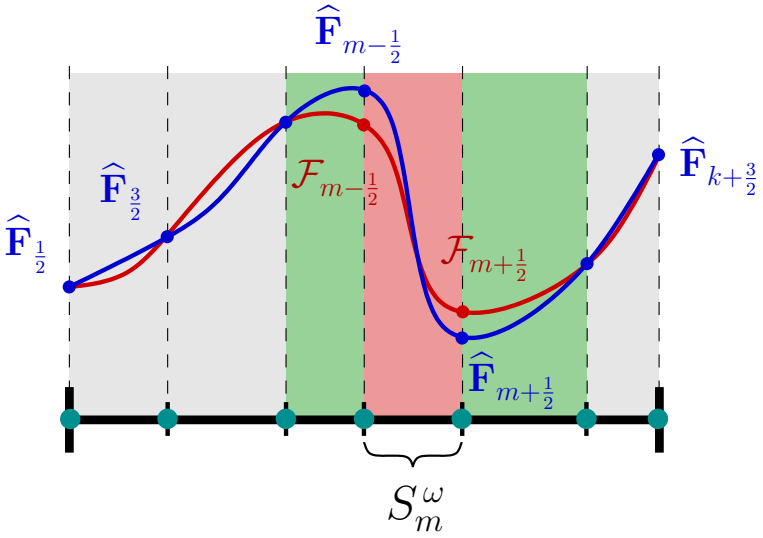

Figure 4: Sketch of the correction of the reconstructed fluxes at subcell boundaries

Additionally, to preserve the local conservation property of the resulting scheme, the left and right neighboring subcells, colored in green in Fig. 4, have to be updated too, even if they are flagged as admissible subcells, since we have substituted the reconstructed fluxes $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ with corrected ones. In the particular case depicted in Fig. 4 where S_{m-2} and S_{m+2} are also flagged as admissible, the sub-mean values $\overline{\mathbf{v}}_{m+1}^{n+1}$ and $\overline{\mathbf{v}}_{m-1}^{n+1}$ are thus replaced respectively by $\overline{\mathbf{v}}_{m-1}^{\star,n+1}$ and $\overline{\mathbf{v}}_{m+1}^{\star,n+1}$ computed through a high-order reconstructed flux on one end and a first-order FV numerical flux on the other end, as follows:

$$\overline{\mathbf{v}}_{m-1}^{\star,n+1} = \overline{\mathbf{v}}_{m-1}^{n} - \frac{\Delta t^n}{|S_{m-1}|}\left(\mathcal{F}^{l}_{m-\frac{1}{2}} - \widehat{\mathbf{F}}_{m-3/2}\right) + \Delta t^n \overline{\mathbf{B}}_{m-1}, \tag{22}$$

$$\overline{\mathbf{v}}_{m+1}^{\star,n+1} = \overline{\mathbf{v}}_{m+1}^{n} - \frac{\Delta t^n}{|S_{m+1}|}\left(\widehat{\mathbf{F}}_{m+3/2} - \mathcal{F}^{r}_{m+\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_{m+1}. \tag{23}$$

For all the remaining admissible subcells (left in grey on Fig. 4), because the associated reconstructed fluxes are not corrected, they do not require any further computation, and the corresponding sub-mean values are the values obtained through the *uncorrected* DG scheme, see Fig. 5.

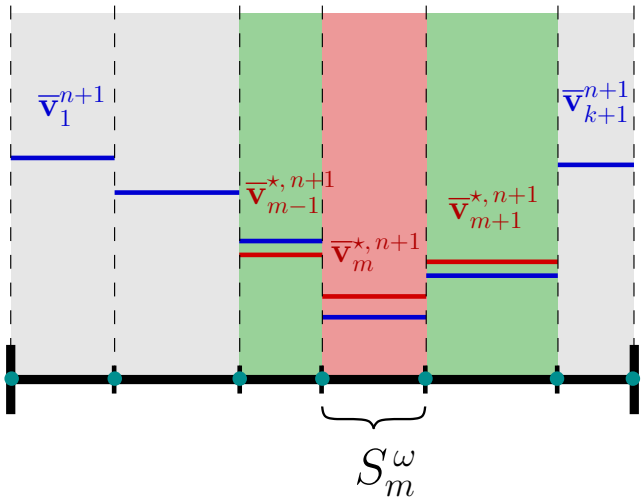

Figure 5: Sketch of sub-mean values before and after correction

3.1. Subcell low-order corrected FV fluxes

In this section, we define the *corrected FV fluxes* $\mathcal{F}^{l/r}_{m\pm\frac{1}{2}}$. Such corrected fluxes are designed in order to: (i) ensure the desired robustness properties, in particular we aim at preserving the set of admissible states (2), see § 3.3 for the details, (ii) obtain a global discrete formulation which is well-balanced. To achieve this, we adapt the ideas introduced in (69; 38) to the framework of the current FV subcells method. For any $\omega_i \in \mathscr{T}_h$ and any marked subcell $S_m \in \mathscr{T}_{\omega_i}$, let define the sub-mesh reconstructed interface values for the topography:

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

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

$$\overline{b}_m^{\pm} := \overline{b}_{m\pm\frac{1}{2}} - \max\left(0, \overline{b}_{m\pm\frac{1}{2}} - \overline{\eta}_m\right), \tag{24}$$

$$\overline{b}_{m+1}^{-} := \overline{b}_{m+\frac{1}{2}} - \max\left(0, \overline{b}_{m+\frac{1}{2}} - \overline{\eta}_m\right), \qquad \overline{b}_{m-1}^{+} := \overline{b}_{m-\frac{1}{2}} - \max\left(0, \overline{b}_{m-\frac{1}{2}} - \overline{\eta}_m\right). \tag{25}$$

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

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

and for the surface elevation and discharge:

$$\overline{\eta}_m^{\pm} := \overline{H}_m^{\pm} + \overline{b}_m^{\pm}, \qquad \overline{q}_m^{\pm} := \overline{H}_m^{\pm}\frac{\overline{q}_m}{\overline{H}_m}, \tag{26}$$

leading to the new subcell's interfaces values:

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

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

$$\mathcal{F}^l_{m+\frac{1}{2}} := \mathcal{F}\left(\overline{\mathbf{v}}_m^{+}, \overline{\mathbf{v}}_{m+1}^{-}, \overline{b}_m^{+}\right) + \begin{pmatrix} 0 \\ g\overline{\eta}_m^{+}\left(\overline{b}_m^{+} - b_{\widetilde{x}_{m+\frac{1}{2}}}\right) \end{pmatrix}, \tag{27}$$

$$\mathcal{F}^r_{m-\frac{1}{2}} := \mathcal{F}\left(\overline{\mathbf{v}}_{m-1}^{+}, \overline{\mathbf{v}}_m^{-}, \overline{b}_m^{-}\right) + \begin{pmatrix} 0 \\ g\overline{\eta}_m^{-}\left(\overline{b}_m^{-} - b_{\widetilde{x}_{m-\frac{1}{2}}}\right) \end{pmatrix}, \tag{28}$$

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

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

Remark 11. For the *uncorrected* DG scheme, nothing changes with respect to numerical fluxes and intermediate state variables. The DG fluxes $\mathcal{F}_{i-1/2}$ and $\mathcal{F}_{i+1/2}$ on cell ω_i left and right interfaces $x_{i-1/2}$ and $x_{i+1/2}$ will always be in pure DG context

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

as already defined in (12).

3.2. Flowchart

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

1. starting from an admissible piecewise polynomial approximate solution $\mathbf{v}_h^n \in (\mathbb{P}^k(\mathscr{T}_h))^2$, compute the candidate solution $\mathbf{v}_h^{n+1} \in (\mathbb{P}^k(\mathscr{T}_h))^2$ using the *uncorrected* DG scheme (10),

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

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

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

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

   Otherwise:

   (a) for all $\omega \in \mathscr{T}_h$ such that $\mathscr{T}_\omega^{\mathrm{f}} \neq \emptyset$, and all $S_m \in \mathscr{T}_\omega^{\mathrm{f}}$, substitute the corresponding *reconstructed fluxes* with some *corrected fluxes* defined in (27)-(A.3), as follows:

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

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

$$\overline{\mathbf{v}}_p^{\star,n+1} = \overline{\mathbf{v}}_p^n - \frac{\Delta t^n}{|S_p|}\left(\widetilde{\mathbf{F}}^l_{p+\frac{1}{2}} - \widetilde{\mathbf{F}}^r_{p-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_p, \tag{29}$$

   for $p \in [\![m-1,\, m+1]\!]$. This subcell corrected method (29) will either falls in one of the previously introduced cases (21), (22) or (23),

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

   (d) go to step 3,

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

3.3. Admissibility criteria

A large number of sensors or detectors have been introduced in the literature, to identify the marked subcells, where some kind of stabilization is required to avoid a loss of robustness. Following (101), we use two admissibility criteria: one for the *Physical Admissibility Detection* (PAD), another addressing the occurrence of spurious oscillations, namely the *Subcell Numerical Admissibility Detection* (SubNAD). This last criterion is supplemented with a relaxation procedure to exclude the smooth extrema from the troubled cells.

Physical Admissibility Detection (PAD)

Here, we define a sensor function that :

- Check if the sub-mean values $\overline{\mathbf{v}}_m^{n+1}$ belongs to Θ, see (2).
- Check if there is any NaN values.

Those are the minimum requirements to enforce the code robustness.

Subcell Numerical Admissibility Detection (SubNAD)

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

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

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

The SubNAD criterion relies on a DMP based on subcell mean values, and not the whole polynomial set of values. Furthermore, as the neighboring subcells set used in the SubNAD is reduced to the first left and right subcells, and not all the subcells contained in the DG cell as well as in the the left and right first neighboring DG cells, see (101), one has to introduce a relaxation mechanism in order to preserve the scheme accuracy in the vicinity of smooth extrema.

Detection of smooth extrema.

In the relaxation procedure proposed in (101), it is assumed that the numerical solution exhibits a smooth extremon if at least the following linearized version of the surface elevation spatial derivative:

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

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

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

Introducing the symmetric values $\partial_x \eta_{\min\setminus\max}^R = \min\setminus\max\left(\overline{\partial_x \eta_{\omega_i}^{n+1}}, \overline{\partial_x \eta_{\omega_{i+1}}^{n+1}}\right)$ and $\partial_x \eta_R := \overline{\partial_x \eta_{\omega_i}^{n+1}} + \frac{h_{\omega_i}}{2}\,\overline{\partial_{xx}\eta_{\omega_i}^{n+1}}$, the right detection factor α_R is obtained in a similar manner. Finally, introducing $\alpha := \min(\alpha_L, \alpha_R)$, we consider that the numerical solution presents a smooth profile on the cell ω_i

if $\alpha = 1$. In this particular case, the SubNAD criterion is relaxed, allowing the high-order accuracy preservation of smooth extrema.

Remark 12. One can apply the subcell numerical admissibility detection SubNAD and relaxation method detailed above on the Riemann invariants $I^{\pm} = u \pm 2\sqrt{gH}$ instead of the surface elevation η. Actually the simplest choice that also leads to the best results, is to make the detection on the surface elevation η. The detection on the Riemann invariants variables produces a more diffused solution.

3.4. Well-balancing property

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

Remark 13. Let us note that under the motionless steady-state assumption $\eta_h = \eta^e$ and $q_h = 0$, the following relation holds:

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

Moreover, as we have

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

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

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

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

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

thus,

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

We have then the following result:

Proposition 1. The discrete formulation obtained by gathering (10) and the local corrected FV schemes on subcells (21), (22) and (23), together with a first-order Euler time-marching algorithm, preserves the motionless steady states, providing that the integrals of (11) are exactly computed for the motionless steady states. Specifically, for all $n \geq 0$ and all $\eta^e \in \mathbb{R}$,

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

Proof. We consider the scheme (17) on uncorrected subcells, and schemes (21), (22) and (23) on corrected subcells. We have to distinguish three different situations: (i) uncorrected subcell, (ii) neighbor of a marked subcell, (iii) marked subcell. We show in what follows that in all those situations, the corrected DG scheme preserves the motionless steady-states at the subcell level:

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

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

$$\overline{\mathbf{v}}_m^{n+1} = \overline{\mathbf{v}}_m^n - \frac{\Delta t^n}{|S_m|}\left(\widehat{\mathbf{F}}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}}\right) + \Delta t^n \overline{\mathbf{B}}_m, \tag{30}$$

with $\widehat{\mathbf{F}}_{m+\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ defined in (18). We have, at steady state:

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

and therefore

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

so that

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

As a consequence, we have

$$\widehat{\mathbf{F}}_{m+\frac{1}{2}} - \widehat{\mathbf{F}}_{m-\frac{1}{2}} = \int_{S_m} \partial_x \mathbf{F}(\mathbf{v}_h, b_h)\,\mathrm{d}x, \tag{31}$$

and injecting (31) into (30) gives

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

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

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

with $\mathcal{F}^l_{m+\frac{1}{2}}$ and $\widehat{\mathbf{F}}_{m-\frac{1}{2}}$ respectively defined in (27) and (18). At steady state, the reconstruction (26) yields $\overline{\eta}^+_m = \overline{\eta}^-_{m+1} = \eta^c$ and $\overline{q}^+_m = \overline{q}^-_{m+1} = 0$. It leads to:

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

and then to:

$$\mathcal{F}^l_{m+\frac{1}{2}} = \frac{1}{2}\begin{pmatrix} 0 \\ g\left((\eta^e)^2 - 2\eta^e b_{\widetilde{x}_{m+\frac{1}{2}}}\right)\end{pmatrix} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m+\frac{1}{2}}), b_h(\widetilde{x}_{m+\frac{1}{2}})\right). \tag{32}$$

Moreover, as in the previous case:

$$\widehat{\mathbf{F}}_{m-\frac{1}{2}} = \mathbf{F}\left(\mathbf{v}_h(\widetilde{x}_{m-\frac{1}{2}}), b_h(\widetilde{x}_{m-\frac{1}{2}})\right). \tag{33}$$

Gathering (32) and (33), we then have

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

so that

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

3. **Corrected subcell**: S_m is marked.
In this case, the corresponding scheme reduces to (21). Following the lines of the previous cases, we have:

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

and therefore

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

so that

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

□

We have just shown that schemes (17)-(21)-(22)-(23) do ensure the well-balanced property in wet subcells for all contexts, wet/wet and wet/dry. As for dry subcells, we can also simply show well-balancing property. Considering a dry zone at time level n, under the assumptions $\eta^n = b^n$ and $q^n = 0$, one can easily show that the dry zone stays a dry zone a the next time level $n+1$, *i.e.* $\eta^{n+1} = b^n$ and $q^{n+1} = 0$, by following a very similar procedure as in the previous proofs.

Remark 14. Note that the use of a non-smooth topography parameterization may be allowed, while still ensuring the well-balancing property, at the price of considering interface reconstructions also for the cells interfaces, for the DG scheme, in the spirit of (38; 69).

3.5. Preservation of the water height positivity

After computing the candidate solution $\mathbf{v}_h^{n+1}$ through the *uncorrected* DG scheme (10), if we detect a negative sub-mean value on an arbitrary subcell, this subcell is then marked and a new (corrected) sub-mean value is evaluated by means of the first-order subcell FV scheme (21). As a consequence, scheme (21) with reconstruction (26) should preserve positivity.

Proposition 2. Under the CFL condition (20), if $\forall \omega \in \mathscr{T}_h$, $\forall S_m \in \mathscr{T}_\omega$, $\overline{v}_m^{\omega,\,n} \in \Theta$, then $\forall \omega \in \mathscr{T}_h$, $\forall S_m \in \mathscr{T}_\omega$, $\overline{v}_m^{\omega,\,n+1} \in \Theta$.

Proof. As the positivity-preserving property of our *a posteriori* LSC of DG schemes relies on the positivity of the first-order FV scheme used as the correction method, let us prove that if $\overline{H}^n_m$ and $\overline{H}^n_{m\pm1}$ are non-negative, then scheme (21) does produce a water height $\overline{H}^{n+1}_m$ also non-negative. Let us first recall the equation corresponding to the time evolution of the discrete surface elevation:

$$\overline{\eta}^{n+1}_m = \overline{\eta}^{n}_m - \frac{\Delta t^n}{|S_m|}\left(\mathcal{F}_1\left(\overline{\mathbf{v}}^{n,+}_m, \overline{\mathbf{v}}^{n,-}_{m+1}, \overline{b}^{+}_m\right) - \mathcal{F}_1\left(\overline{\mathbf{v}}^{n,+}_{m-1}, \overline{\mathbf{v}}^{n,-}_{m}, \overline{b}^{-}_m\right)\right), \tag{34}$$

where $\mathcal{F}_1$ represents the first component of the numerical flux $\mathcal{F}$ and $\overline{\mathbf{v}}_m^{n,\pm}$, $\overline{b}_m^{n,\pm}$ are defined in (26) and (24). For sake of simplicity, we drop in the following the superscript n. Equation (34) rewrites explicitly as:

$$\begin{aligned}\overline{\eta}_m^{n+1} = \overline{\eta}_m \quad &- \quad \frac{\Delta t^n}{2\,|S_m|}\left(\overline{H}_m^+\frac{\overline{q}_m}{\overline{H}_m} + \overline{H}_{m+1}^-\frac{\overline{q}_{m+1}}{\overline{H}_{m+1}} - \sigma\left(\overline{\eta}_{m+1}^- - \overline{\eta}_m^+\right)\right)\\ &- \quad \frac{\Delta t^n}{2\,|S_m|}\left(\overline{H}_m^-\frac{\overline{q}_m}{\overline{H}_m} + \overline{H}_{m-1}^+\frac{\overline{q}_{m-1}}{\overline{H}_{m-1}} - \sigma\left(\overline{\eta}_m^- - \overline{\eta}_{m-1}^+\right)\right).\end{aligned} \tag{35}$$

Noticing that $\overline{\eta}_{m+1}^- - \overline{\eta}_m^+ = \overline{H}_{m+1}^- - \overline{H}_m^+$ as well as $\overline{\eta}_m^- - \overline{\eta}_{m-1}^+ = \overline{H}_m^- - \overline{H}_{m-1}^+$, and subtracting $\overline{b}_m$ on both sides of this last expression, equation (35) can be reformulated as:

$$\begin{aligned}\overline{H}_m^{n+1} \quad &= \quad \left[1 - \frac{1}{2}\lambda\left(\sigma - \overline{u}_m\right)\frac{\overline{H}_m^-}{\overline{H}_m} - \frac{1}{2}\lambda\left(\sigma + \overline{u}_m\right)\frac{\overline{H}_m^+}{\overline{H}_m}\right]\overline{H}_m\\ &+ \quad \left[\frac{1}{2}\lambda\left(\sigma + \overline{u}_{m-1}\right)\frac{\overline{H}_{m-1}^+}{\overline{H}_{m-1}}\right]\overline{H}_{m-1} + \left[\frac{1}{2}\lambda\left(\sigma - \overline{u}_{m+1}\right)\frac{\overline{H}_{m+1}^-}{\overline{H}_{m+1}}\right]\overline{H}_{m+1},\end{aligned} \tag{36}$$

with $\overline{u}_m = \dfrac{\overline{q}_m}{\overline{H}_m}$ and $\lambda = \dfrac{\Delta t^n}{|S_m|}$. Therefore, $\overline{H}_m^{n+1}$ reads as a convex combination of $\overline{H}_{m-1}, \overline{H}_m$ and $\overline{H}_{m+1}$. Furthermore, since by construction $0 \leq \overline{H}_p^{\pm} \leq \overline{H}_p$, $\forall p \in [\![1,\, k+1]\!]$ and by respect of the CFL condition (20), $\lambda\alpha \leq 1$, and then all the coefficients involved in the convex combination (36) are non-negative. It follows that $\overline{H}_m^{n+1} \geq 0$. □

Remark 15. The proposed positivity criteria are based on subcell values, and indeed, our goal is to show that our scheme preserves the positivity at the subcell level. Hence, the chosen strategy, as it is, does not ensure the pointwise positivity of h at some specific nodes: such a property is not needed. If one requires such pointwise positivity, for some specific reasons, we emphasize that an additional "positivity limiter", as the one provided in (110) for instance, can be combined with our approach, to ensure the positivity of the polynomial solution at any chosen points.

4. Numerical validations

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

4.1. A new analytical solution for the NSW equations

This first test case aims at numerically evaluating the rates of convergence of the present *a posteriori* LSC of DG schemes. To do so, following the methodology introduced in (102) in the context of compressible gas dynamics, we make use here of a new manufactured smooth solution of the NSW equations. Details on the design of such solution can be found in Appendix A. This solution has the very interesting features to achieve any arbitrary regularity, *i.e.* $\mathbf{v}(.,\,t) \in \mathcal{C}^{N_s}(\Omega)$, $\forall\, t < t_c(N_s)$ and any $N_s \in \mathbb{N}^*$, allowing the study of convergence up to any order of accuracy, while involving almost vanishing water depth, together with a loss of regularity and the occurrence of discontinuous

profiles for $t \geq t_c$.

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

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

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

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

While the *uncorrected* DG scheme (10) allows to compute the solution without any robustness issue for small enough values of time, nonphysical oscillations may be generated for larger values of time, leading to the activation of the *a posteriori* LSC method . A comparison between our fourth-order numerical solution computed on a mesh made of 60 cells, and the analytical solution at $t = 0.1\,s$ is shown on Fig. 6.

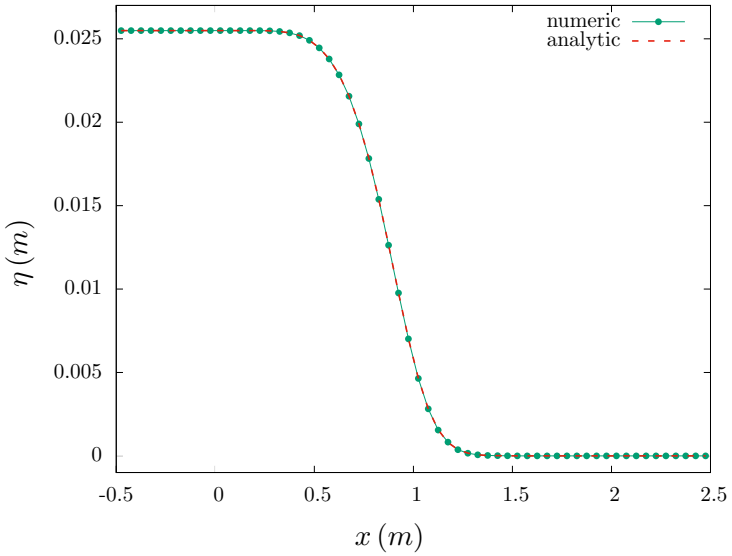

Figure 6: Test 1 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.1\,s$ with the *a posteriori* LSC method for $k = 3$ and $n_e = 60$.

One can see in Fig. 6 that only the cell mean values are displayed. We can also observe how the numerical scheme has very accurately captured to exact solution. In Table 1, we gather the global L^2-errors as well as the rates of convergence for different order of approximation, computed on the surface elevation at $t = 0.1s$. As expected, the computed rates of convergence scale as $O(k+1)$. A similar behavior can be observed for the horizontal discharge q.

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

Table 1: Test 1 - A new analytical solution for the NSW equations: L^2-errors between numerical and analytical solutions and convergence rates for η at time $t = 0.1s$

In a second time, we consider a larger final computational time $t > t_c$, so that a right-going discontinuity has developed from the initially regular profile, allowing to check the ability of the proposed *a posteriori* LSC method to stabilize the computation, namely to get rid of the spurious oscillations as well as enforcing the positivity of the water height. We run the previous case until $t = 0.55$, with $k = 3$ and $n_e = 100$ mesh elements. Note that the standard DG method crashes in this case, since nonphysical undershoots would be rapidly amplified. In Fig. 7, a comparison between the *a posteriori* corrected DG solution and a reference solution obtained with a robust first-order FV method and $n_e = 10000$ mesh elements.

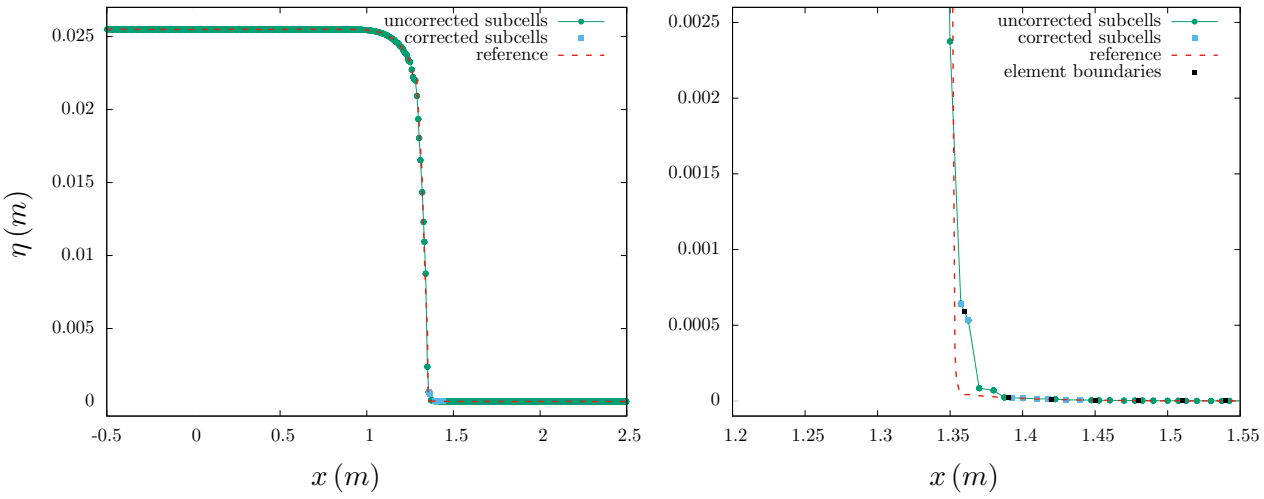

Figure 7: Test 1 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.55\,s$ with the *a posteriori* LSC method (left) for $k = 3$ and $n_e = 100$, with a zoom on the discontinuity and wet/dry interface (right).

This is a challenging computation for high-order methods since small values of water height occur and thus small undershoots generally quickly lead to larger undershoots and possibly loss of positivity. In practice, the sensor starts to be activated when the strong gradient appears, slightly before the apparition of the discontinuity. A particular emphasize is put in Fig. 7 on the location of marked subcells, where uncorrected subcells are plotted by green dots while corrected subcells are plotted with blue squares. We observe that the particular combination of admissibility criteria introduced in §3.3 works quite well in practice, as the detection has been able to accurately track the moving front, and doing so removed the spurious oscillations without impacting smooth areas.

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

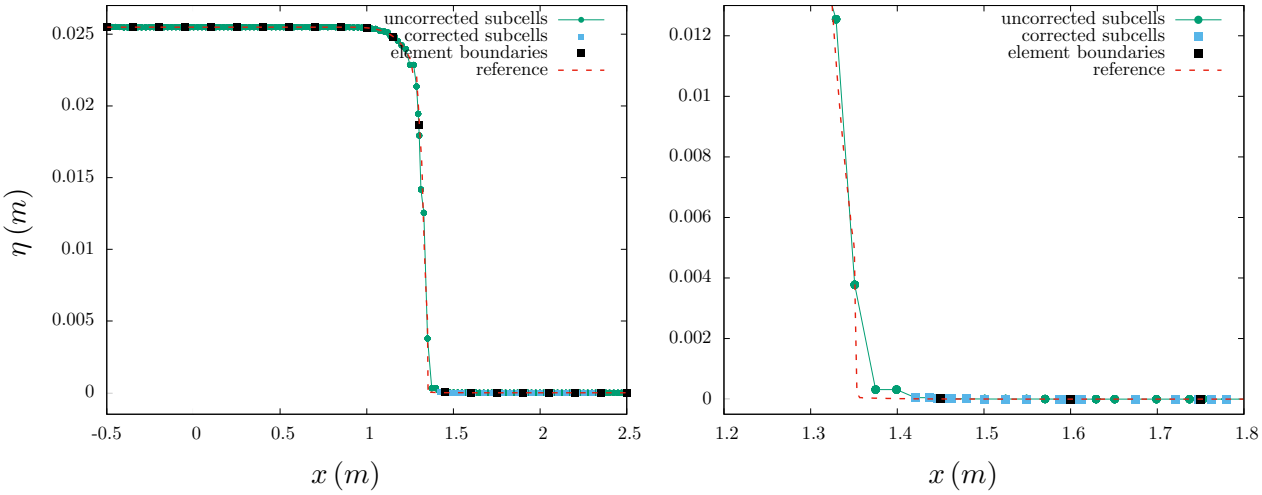

Figure 8: Test 1 - A new analytical solution for the NSW equations - Free surface elevation computed at $t = 0.55\,s$ with the *a posteriori* LSC method (left) for $k = 8$ and $n_e = 20$, with a zoom on the discontinuity and wet/dry interface (right).

4.2. Dam-break

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

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

The final time is set to $t = 0.075\,s$. In Fig. 9, on a 50 cells mesh, fourth-order *uncorrected* DG solution is displayed on the left figure, while the *corrected* solution is plotted on the right one.

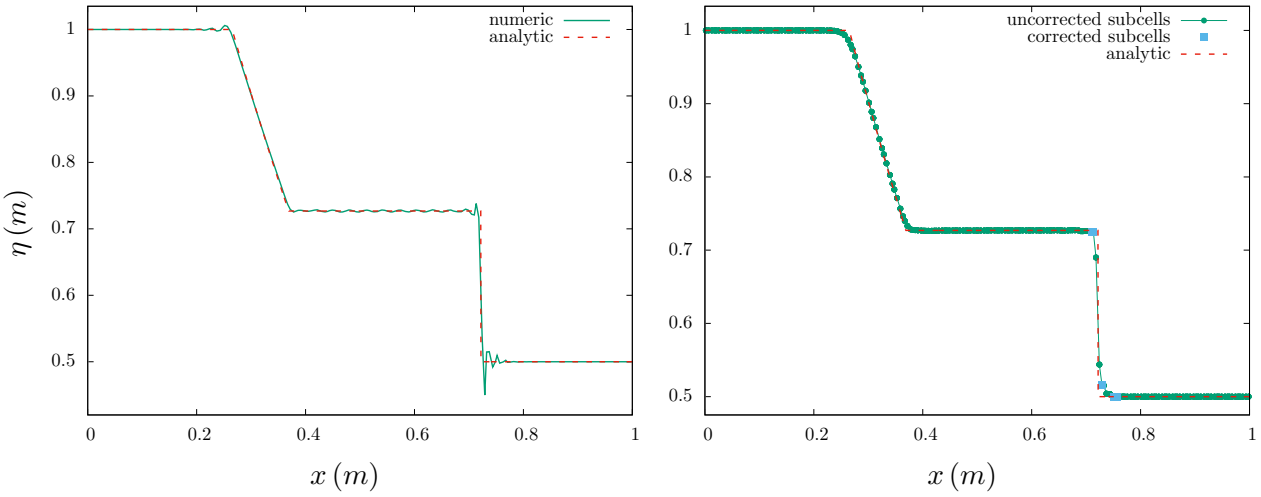

Figure 9: Test 2 - Dam break on a wet bottom- Free surface elevation computed at $t = 0.075\,s$ with the *uncorrected* DG method (left) and the *a posteriori* LSC method (right), with $k = 3$ and $n_e = 50$ mesh elements.

This illustrates very clearly that even if the correction has been activated on in a very sharp area in the vicinity of the discontinuity, the solution has still been cleansed from its spurious oscillations. Now, we compare our *a posteriori* LSC method with the limitation process introduced in (110) (refereed to as PL/TVB method in what follows), which combines the positivity-preserving limiter (113) with a standard TVB limiter (26). Following (110), the constant M involved in the TVB limiter is set to $M = 0$. The results are plotted in Fig. 10 and Fig. 11.

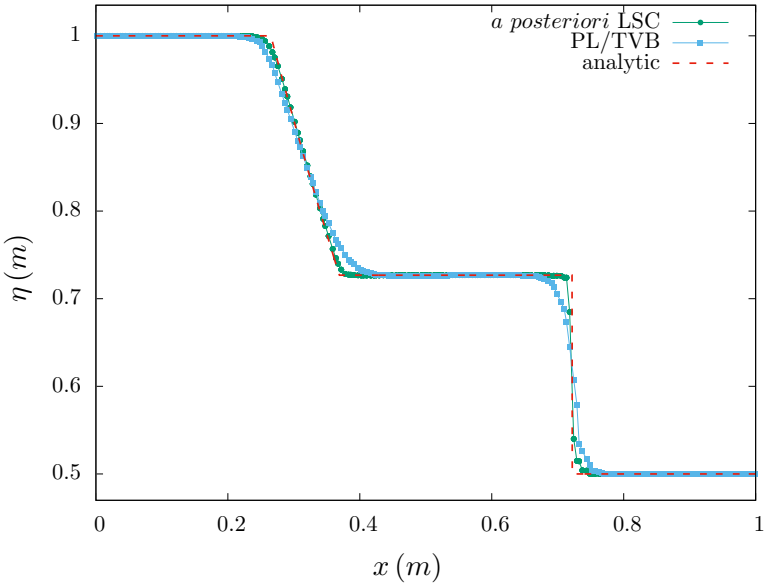

Figure 10: Test 2 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method and PL/TVB method for $k = 3$ and $n_e = 50$.

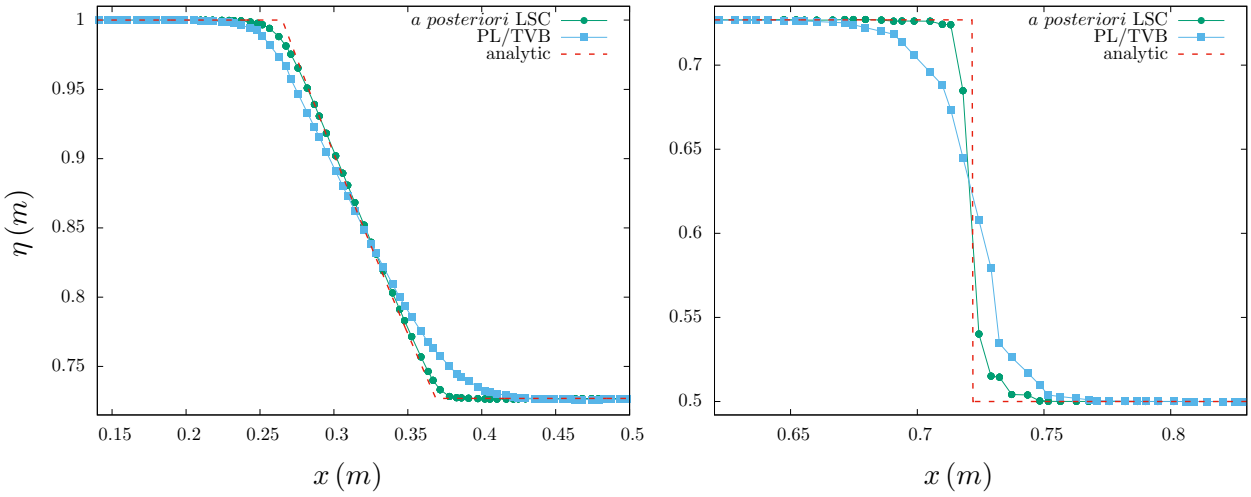

Figure 11: Test 2 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method and PL/TVB method for $k = 3$ and $n_e = 50$, with a zoom on the rarefaction wave (left) and the shock wave (right)

In Fig. 10 and 11, one can observe that the present correction technique outperforms the positivity-preserving + TVB limiter, both in the rarefaction and shock resolution. Finally, to demonstrate how this *a posteriori* LSC method scales going to very high-orders of accuracy and very coarse meshes, we run the same test with $k = 9$ and a 10 mesh elements. The corresponding numerical result is shown on Fig. 12.

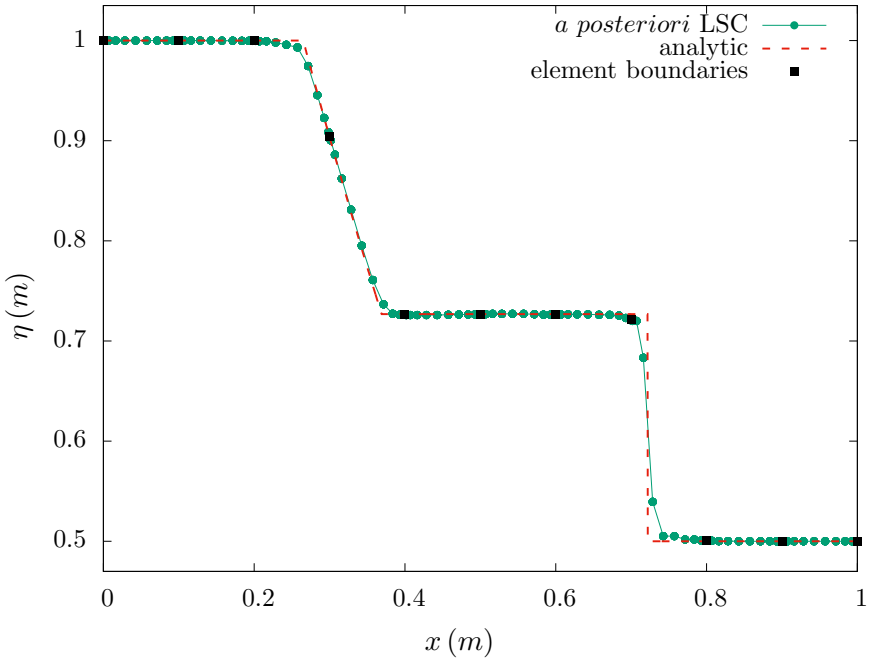

Figure 12: Test 2 - Dam break on a wet bottom - Free surface elevation computed at $t = 0.075\,s$ - Comparison between *a posteriori* LSC method (right) and PL/TVB method (left) for $k = 9$ and $n_e = 10$.

Fig. 12 illustrates the high capability of this *a posteriori* LSC method to retain the precise subcell resolution of discontinuous Galerkin schemes, allowing the use of very coarse meshes, along with being able to avoid the appearance of spurious oscillations or any unfortunate crash of the code. This figure also displays how the present correction affects the solution only at the subcell level, allowing the resolution of the shock in only one mesh element.

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

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

We compute the evolution up to $t = 0.05\,s$, with $k = 3$ and $n_e = 50$, in order to show the ability of the proposed method to compute the propagation of a wet/dry front. A comparison between the numerical results obtained with the *a posteriori* LSC method and the analytical solution is shown on Fig. 13 (left). Additionally, we compare these results with those obtained with the PL/TVB limitation process at the same times on Fig. 13 (right), together with zoomed profiles on Fig. 14.

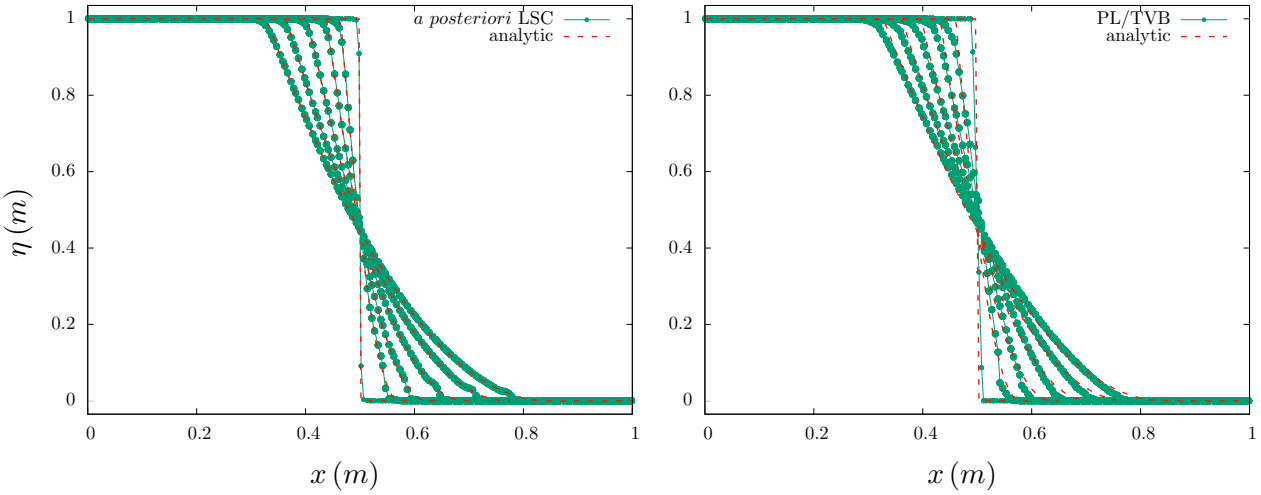

Figure 13: Test 2 - Dam break on a dry bottom - Free surface elevation computed at different times between $0.002s$ and $0.05\,s$ - Comparison between *a posteriori* LSC method (left) and PL/TVB method (right) for $k = 3$ and $n_e = 50$.

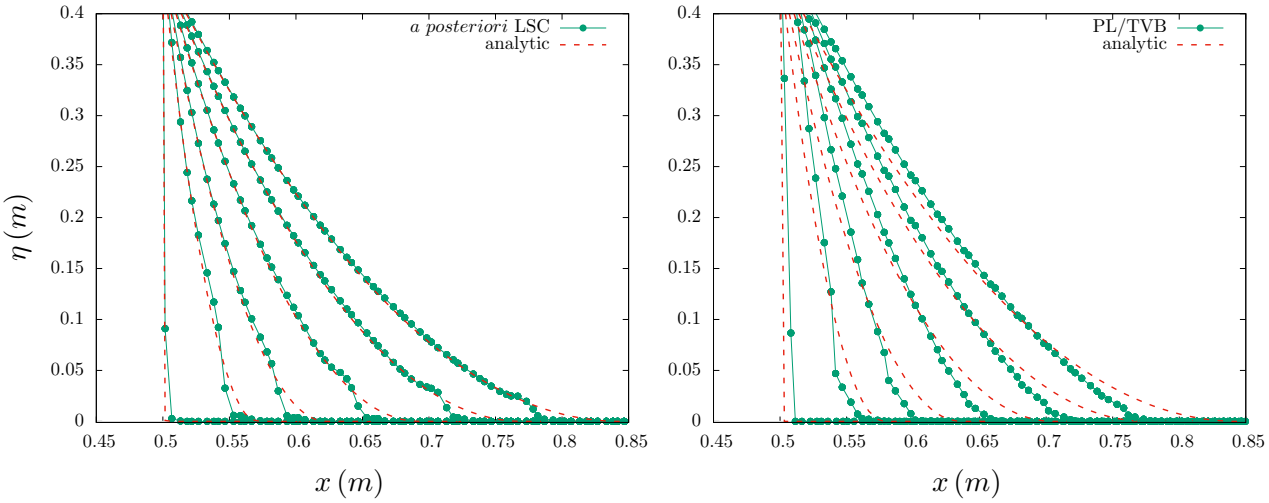

Figure 14: Test 2 - Dam break on a dry bottom - Free surface elevation computed at different times between $0.002s$ and $0.05\,s$ - Comparison between *a posteriori* LSC method (left) and PL/TVB method (right) for $k = 3$ and $n_e = 50$, with a zoom on the wet/dry interface.

Those results show how our subcell correction technique behaves in comparison to the PL/TVB limiter, in the context of the propagation of a wet/dry front. Finally, to exhibit once more the high scalability of the present *a posteriori* LSC method to very high-order of accuracy, we set $k = 8$ and $n_e = 10$. The corresponding numerical result is shown on Fig. 15.

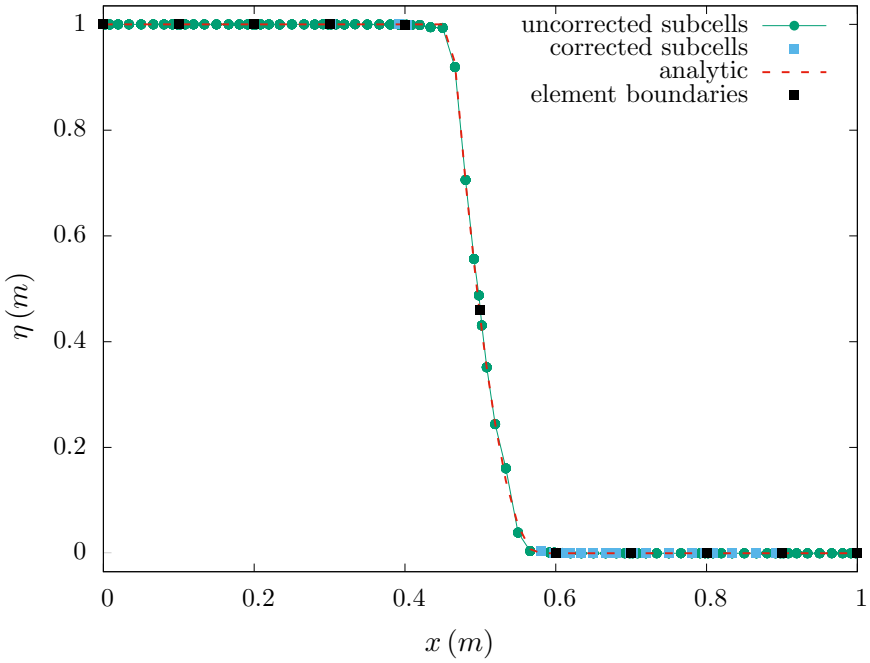

Figure 15: est 2 - Dam break on a dry bottom - with the *a posteriori* LSC method for $k = 8$ and $n_e = 10$ at t = 0.01s.

4.3. Well-balancing property

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

$$b(x) = \begin{cases} A\left(\sin\left(\dfrac{(x - x_1)\cdot\pi}{75}\right)\right)^2 & \text{if} \quad x_1 \leq x \leq x_2, \\ 0 & \text{elsewhere,} \end{cases} \tag{37}$$

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

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

We evolve this initial configuration in time up to 100000 time iterations, with a fourth-order approximation and 120 mesh elements. The numerical results obtained with the *a posteriori* LSC method method are shown on Fig. 16.

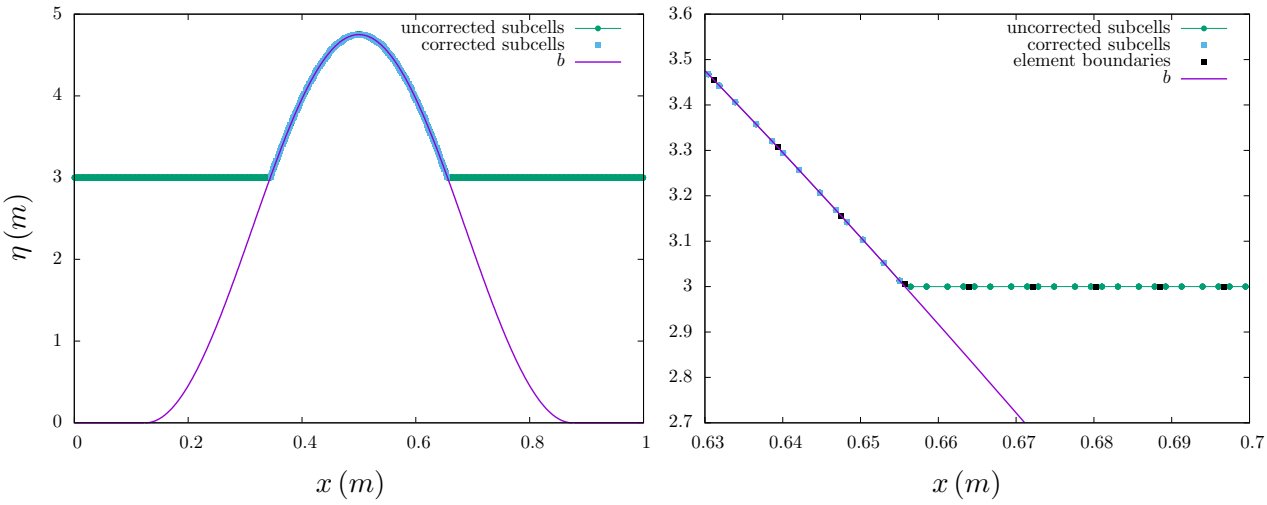

Figure 16: Test 3 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$ (left), with a zoom on the wet/dry interface (right).

We highlight in Fig. 16 the particular marked cells, in which the correction has been performed. We emphasize that the steady state is effectively preserved up to the machine accuracy, validating numerically the compatibility of the *a posteriori* LSC method with the well-balancing property. A similar behavior is reported for other values of k and n_e.

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

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

We evolve this initial configuration in time up to $100,000$ time iterations, with a fourth-order approximation and 120 mesh elements. The numerical results obtained with the *a posteriori* LSC method are shown on Fig. 17.

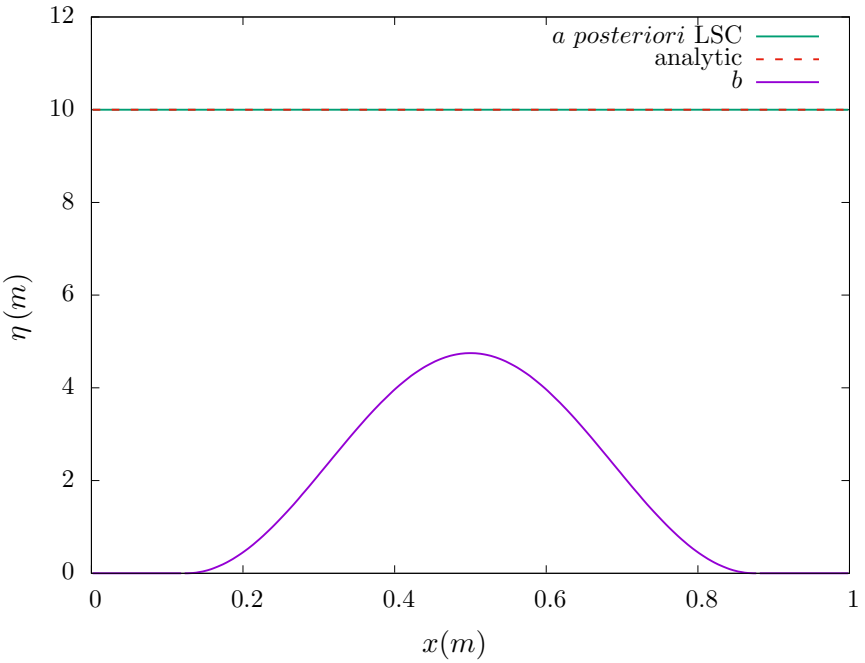

Figure 17: Test 3 - Preservation of a motionless steady state - Free surface elevation at $t = 50s$.

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

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

Table 2: Test 3 - Preservation of a motionless steady state: L^2-errors between numerical and exact steady state solutions for η at time $t = 50s$.

4.4. Transcritical flow over a bump

We focus in this test on a classical transcritical flow without shock, see for instance (48) for a complete description. The computational domain is $\Omega = [0, 25]\,(m)$. The topography profile is defined as follows:

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

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

$$\eta_0(x) = 0.66 \;\; m \quad \text{and} \quad q_0(x) = 0 \;\; m^2.s^{-1},$$

and we prescribe the following boundary conditions:

$$\begin{cases} \text{upstream: } q = 1.53 \;\; m^2.s^{-1}, \\ \text{downstream: } h = 0.66 \;\; m \text{ while the flow is subcritical.} \end{cases}$$

We run this test case with $k = 3$, $n_e = 100$ and $t = 200s$. We show on Fig. 18 the free surface elevation and the discharge obtained with the *a posteriori* LSC method , at several moments during the transient part of the flow ($3.55\,s$ and $20.3\,s$) and when the steady state is reached ($200\,s$), showing a very good agreement with the analytical solution.

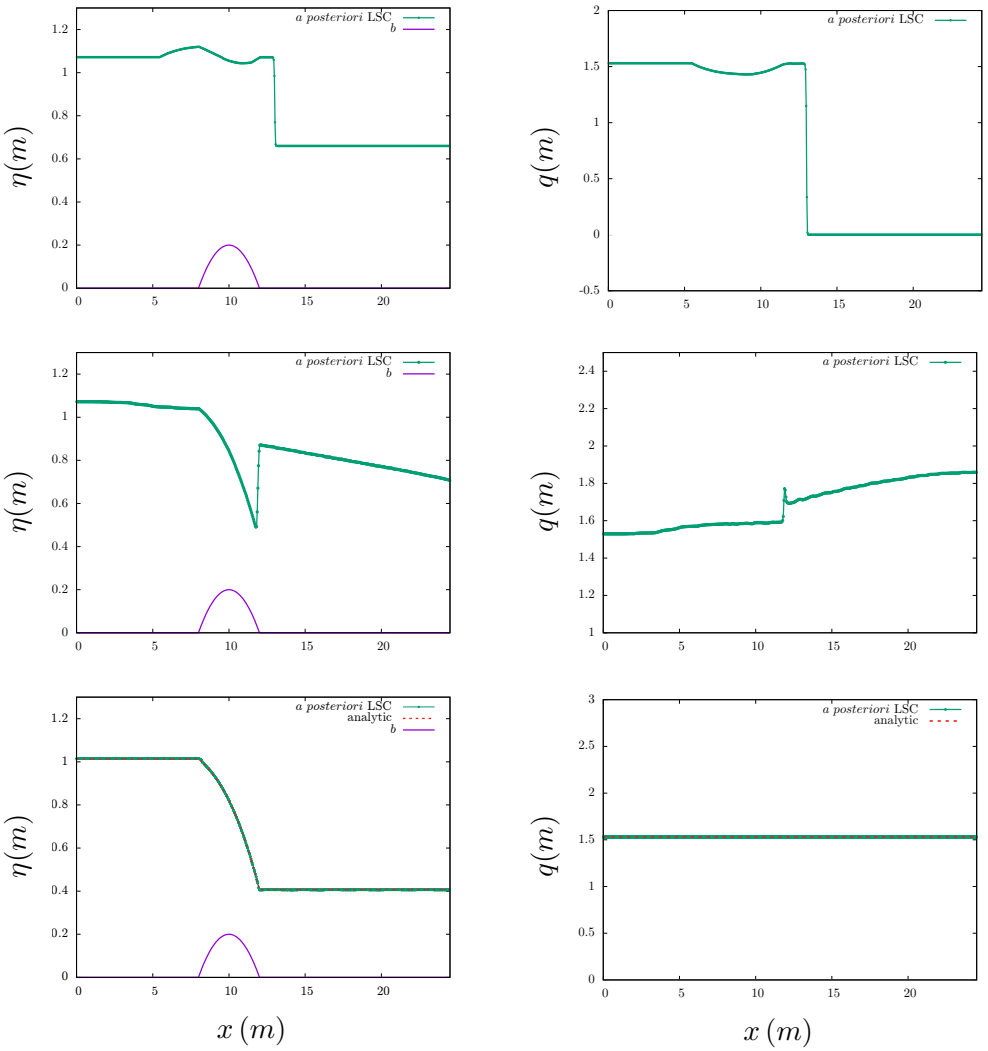

Figure 18: Test 4 - Transcritical flow without shock - Free surface elevation and discharge computed at several moments, $3.55s$, $20.3s$ and $200s$, with the *a posteriori* LSC method , for $k = 3$ and $n_e = 100$.

4.5. Carrier and Greenspan's transient solution

This test case, introduced in (21), describes the physical process in which the water level near the shoreline of a sloping beach is initially depressed, the fluid held motionless and then released at $t = 0$.

A transient wave is generated which runs up the beach, before returning to equilibrium state in a slow convergence process, reproducing some interesting conditions for assessing the robustness of the *a posteriori* LSC method in computing long waves run-up. In (21), a hodograph transformation is used to solve the NSW equations and obtain an analytical solution. The transformation makes use of two dimensionless variables (in the following, starred variables denote dimensionless quantities) σ^* and λ^* which are, respectively, a space-like and a time-like coordinate given by

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

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

$$x^* = x/l, \quad \eta^* = \eta/(\alpha l), \quad u^* = u/\sqrt{g\alpha l}, \quad t^* = t/\sqrt{l/\alpha g}, \tag{38}$$

and the non-dimensional phase speed is given by:

$$c^* = \sqrt{\eta^* - x^*}. \tag{39}$$

The initial solution is specified by the following initial conditions:

$$\eta_0^*(\sigma^*) = e\Big(1 - \frac{5}{2}\frac{a^3}{(a^2+\sigma^{*2})^{\frac{3}{2}}} + \frac{3}{2}\frac{a^5}{(a^2+\sigma^{*2})^{\frac{5}{2}}}\Big), \qquad q_0^*(\sigma^*) = 0 \qquad \text{and} \qquad x^* = -\frac{\sigma^{*2}}{16} + \eta_0^*, \tag{40}$$

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

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

where we have set $\sigma^* = a\sigma, \lambda^* = a\lambda$. This set of equations may be solved by some iterative process. In what follows, we set $e = 0.1$, $\alpha = 1/50$, the initial surface profile (40) is provided in the dimensional case with the length scale $l = 20$ m and we define $\beta := \alpha\, e\, l$. We run this test case with $k = 3$ and 50 mesh elements, for different values of discrete time t in the range $[0.5\,s, 23\,s]$, see Fig. 19 (left) and at $t = 200\,s$ on Fig. 19 (right).

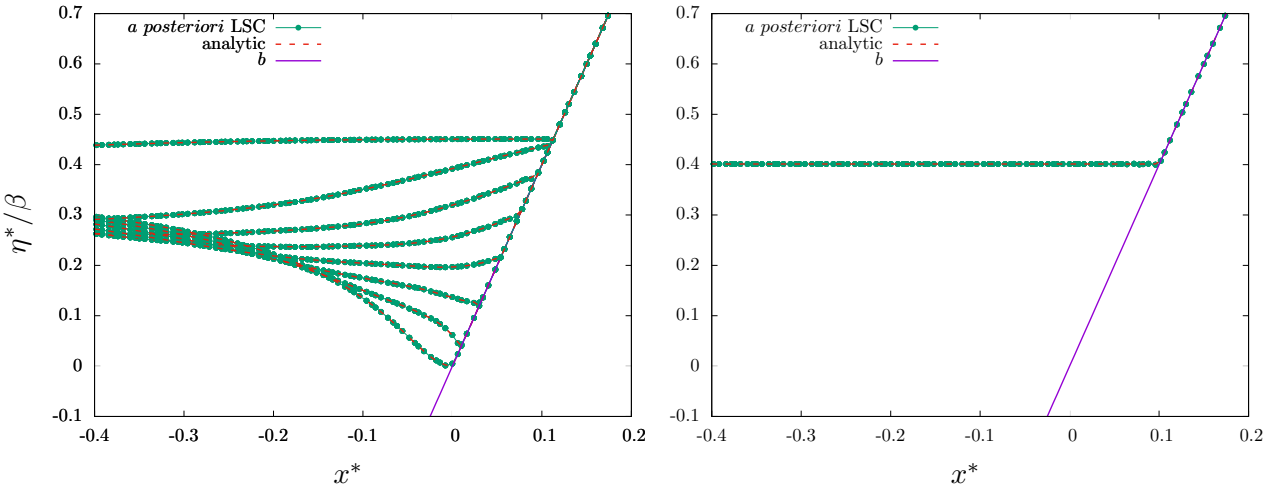

Figure 19: Test 4 - Carrier and Greenspan's transient solution - Free surface elevation η^*/β plotted versus the onshore coordinate x^* - Free surface elevation for different values of time in the range $[0.5\,s,\ 23\,s]$ (left) and at $t = 200\,s$ (right) for $k = 3$ and $n_e = 50$.

In view of result displayed in Fig.19, one can see how accurate DG scheme along with our *a posteriori* LSC method is, as the numerical solution is extremely close to the exact solution and is able to simulate the return to the equilibrium state. This is due to the ability of our correction method to surgically modified the numerical solution only in the very few concerned subcells, as illustrated on Fig. 20.

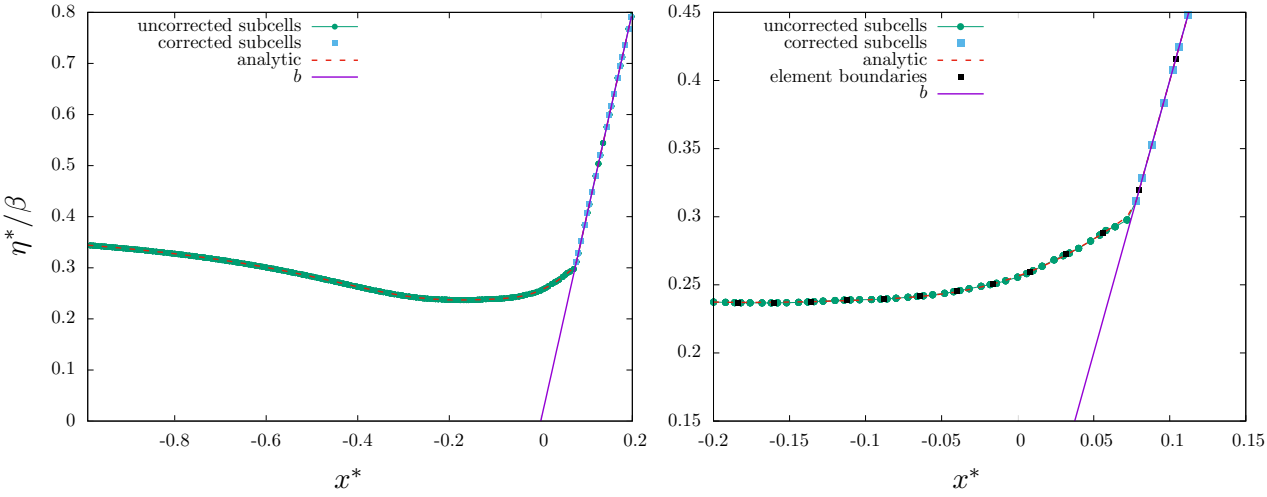

Figure 20: Test 4 - Carrier and Greenspan's transient solution - Free surface elevation computed at $t = 7\,s$ with the *a posteriori* LSC method for $k = 3$ and $n_e = 50$ (left): corrected and uncorrected subcells are respectively plotted with blue squares and green dots, with a zoom on the shoreline (right)

We finally assess the use of a high-order polynomial approximation ($k = 8$) on a very coarse mesh ($n_e = 10$) to emphasize the very accurate and interesting subcell resolution ability of the proposed approach. The results obtained at $t = 7\,s$ are plotted on Fig. 21.

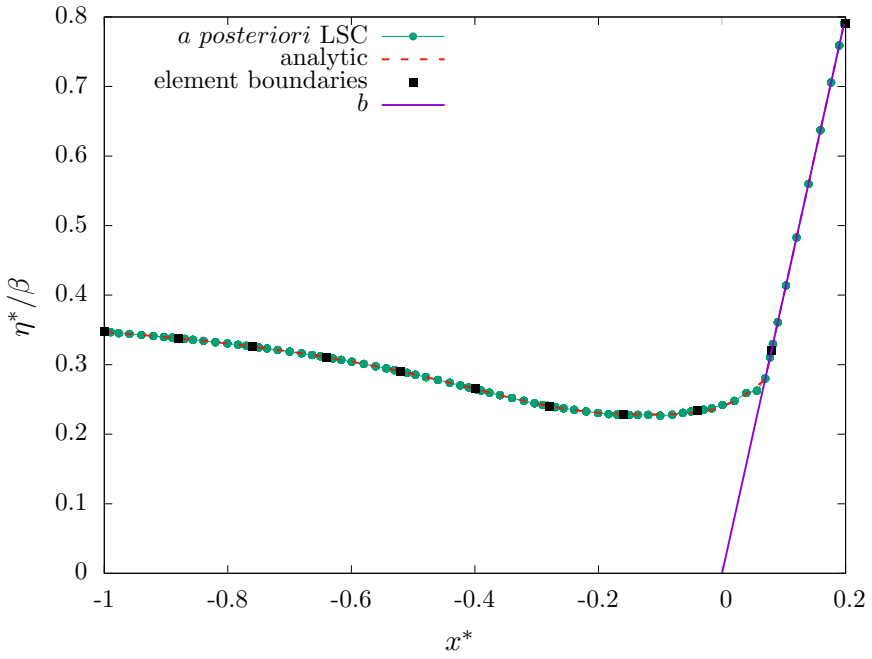

Figure 21: Test 4 - Carrier and Greenspan's transient solution - Free surface elevation computed at $t = 7\,s$ with the *a posteriori* LSC method for $k = 8$ and $n_e = 10$.

4.6. Carrier and Greenspan's periodic solution

In this test case, a monochromatic wave is let run-up and run-down on a plane beach. This solution represents the motion of a periodic wave of dimensionless amplitude A^* and frequency ω^* traveling shoreward and being reflected out to sea generating a standing wave on a plane beach. Recalling the dimensionless quantities (38) and (39), the analytical solution is formulated as follows:

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

where J_0 and J_1 stand for the Bessel functions of zero and first order. We consider the solution obtained for $A^* = 0.6$ and $\omega^* = 1$ (non-breaking wave), together with the length scale $l = 20\,m$ and a bottom slope $\alpha = 1/30$. The value of this solution at $t = 0$ is supplied as initial condition, and similarly to the previous transient case, the analytical variations of the surface elevation at the left boundary is used as an offshore inlet boundary condition, generating the motion. We refer the reader to (21) for a complete description. We set $k = 3$ and $n_e = 50$ and we compute the time evolution up to $t = 1.5\,T$, where T is the time period of the periodic forcing. We show on Fig. 22 some snapshots of the free surface elevation plotted at several discrete time in the range $[1.25T, 1.5T]$ with the *a posteriori* LSC method , showing a very good agreement between the numerical solution and the analytical one. Additionally, we compare these results with those obtained with the PL/TVB method on Fig. 23 with $M = 0$ (left) and with $M = 32$ (right). Let us note that in (110), the authors make use of $M = 0$ in every situations, except for the convergence rate analysis where

$M = 32$ is used. As this test case is for the most part smooth (except at the wet/dry transition point), a non-zero value of M can be used in order to improve the quality of the results, as depicted by Fig. 23. However, even for higher value of M, the PL/TVB limiter is outperformed by the present *a posteriori* LSC method .

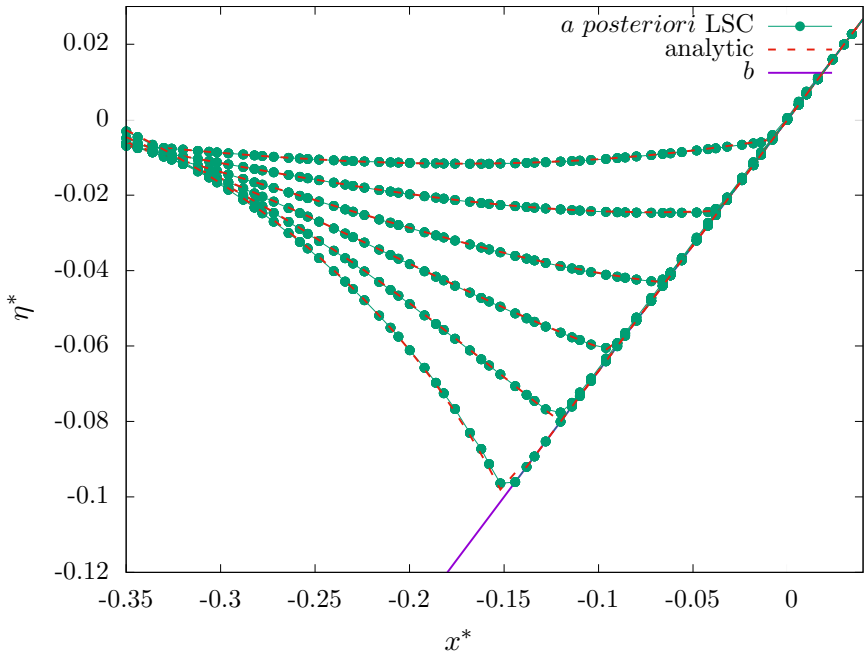

Figure 22: Test 5 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[1.25T, 1.5T]$ with the *a posteriori* LSC method for $k = 3$ and $n_e = 50$

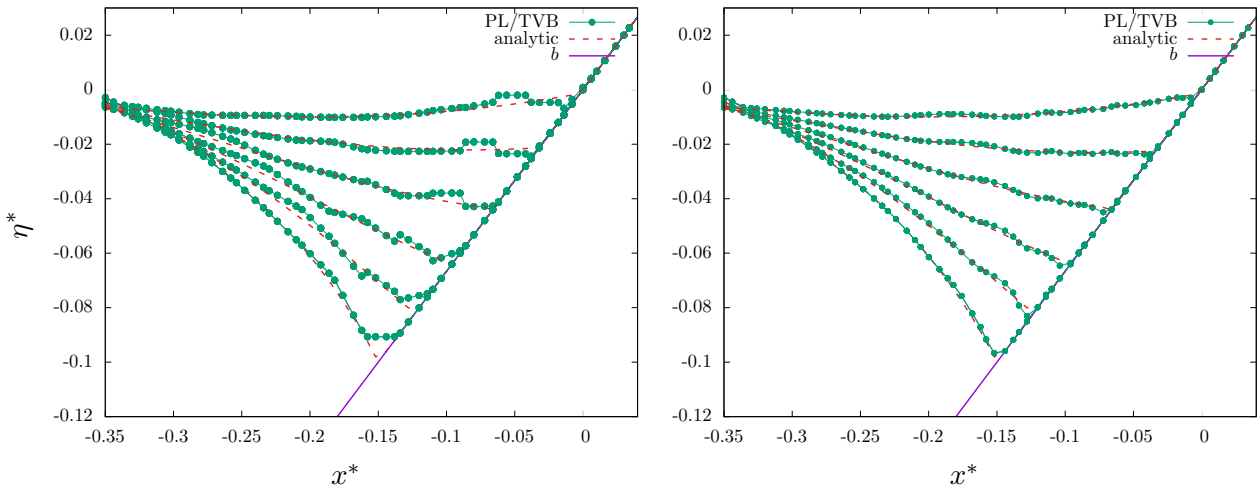

Figure 23: Test 5 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[1.25T, 1.5T]$ with the PL/TVB method for $k = 3$ and $n_e = 50$, with $M = 0$ (left) and $M = 32$ (right).

In order to emphasize the accuracy of the proposed approach for long time integration, we set

$t = 15T$ and show on Fig. 24 the free surface elevation obtained at times $t = 14.5T$ (left) and $t = 15T$ (right), for $k = 3$ and $n_e = 50$. We observe that such a long time-integration has a negligible impact on the accuracy of the predictions of the shoreline location. Such a result can be reproduced with a high-order approximation $k = 8$ and a very coarse mesh $n_e = 10$, showing again the ability of our approach to provide a high-order accurate subcell description of the motion, see Fig. 26. In Fig. 25, we show time-series of the shoreline elevation "η_s" in the range $[0, 6T]$ (left) and $[0, 15T]$ (right). We can see that the minimum and maximum water elevations are accurately computed, even after a large number of periods.

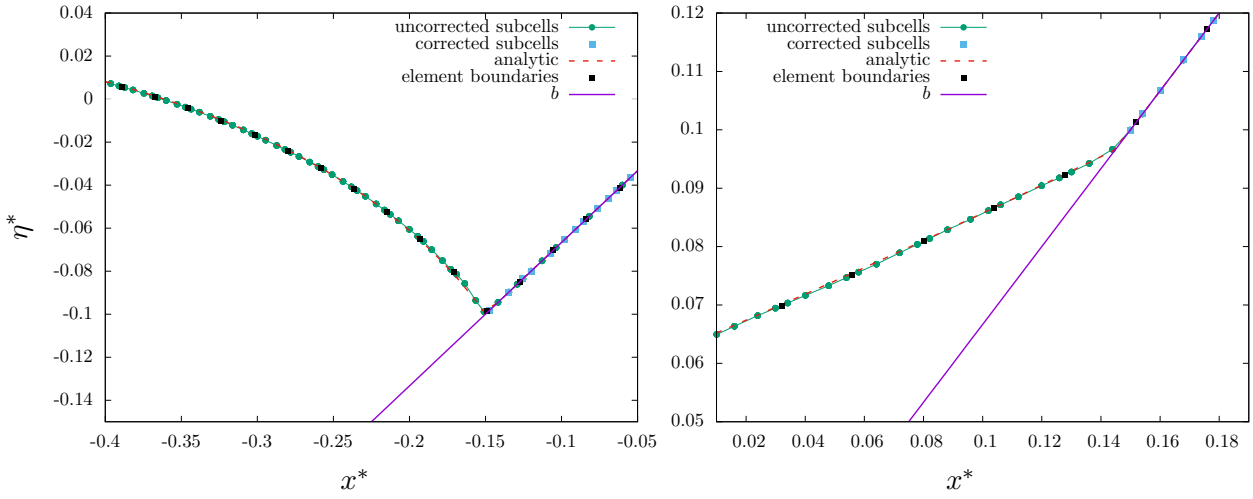

Figure 24: Test 5 - Carrier and Greenspan's periodic solution - Free surface elevation computed at $t = 14.5T$ (left) and $t = 15T$ (right) with the *a posteriori* LSC method for $k = 3$ and $n_e = 50$.

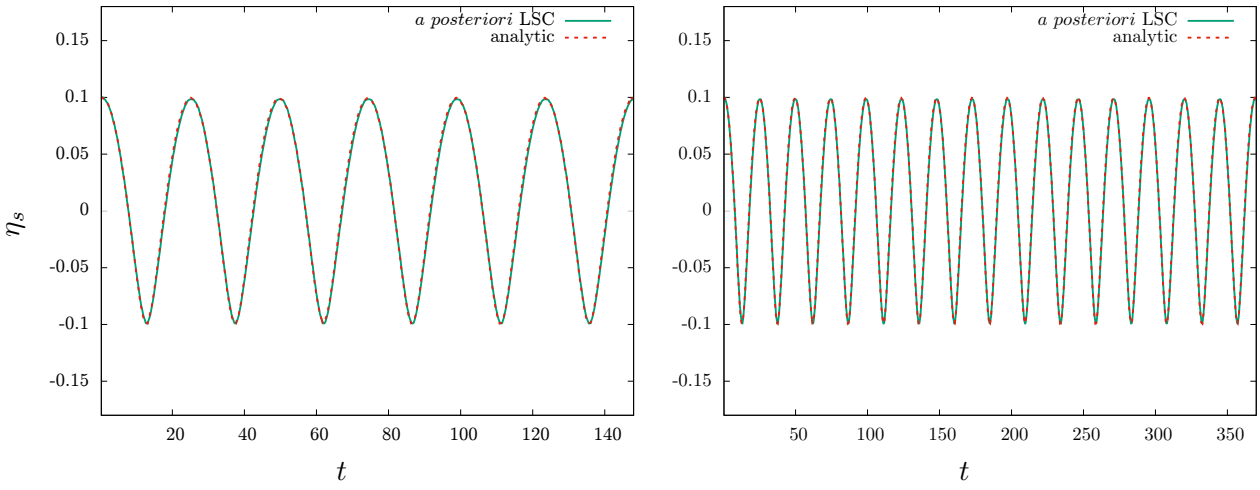

Figure 25: Test 5 - Carrier and Greenspan's periodic solution - Time-series of the shoreline elevation in the range $[0, 6T]$ (left) and $[0, 15T]$ (right), with the *a posteriori* LSC method for $k = 3$ and $n_e = 60$.

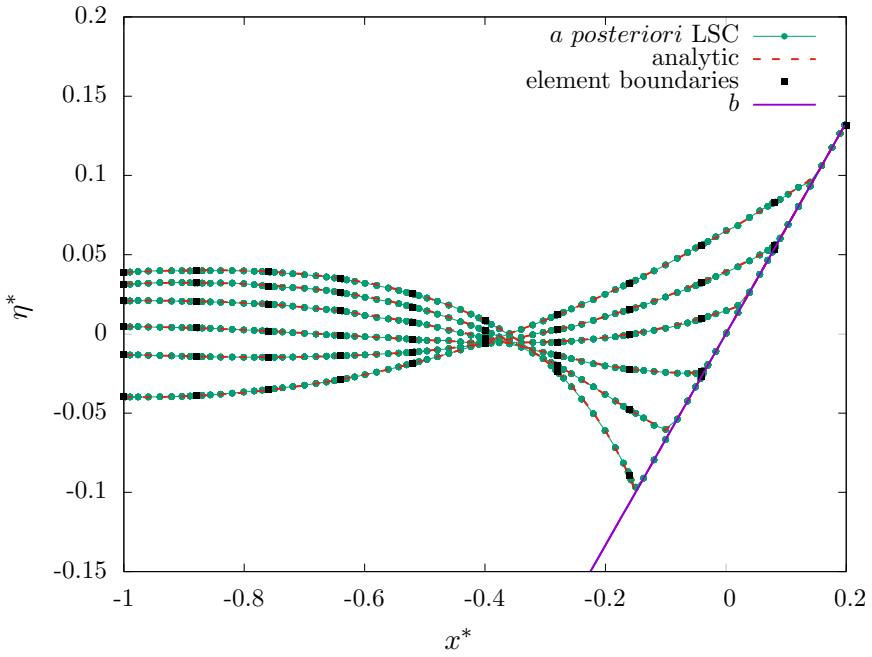

Figure 26: Test 5 - Carrier and Greenspan's periodic solution - Free surface elevation computed for different values of time in the range $[14.5T, 15T]$ for $k = 8$ and $n_e = 10$.

4.7. Run-up of a solitary wave on a plane beach

The last test case is devoted to the computation of the run-up of a solitary wave on a constant slope. Such run-up phenomena are investigated experimentally and numerically in (93). In this test, a solitary wave traveling from the shoreward is let run-up and run-down on a plane beach, before being fully reflected and evacuated from the computational domain. The topography is made of a constant depth area juxtaposed with a plane sloping beach of constant slope α such that $\cot(\alpha) = 19.85$. The right boundary condition is transmissive. The initial condition is defined as follows:

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

where $\gamma = \sqrt{\dfrac{3A}{4H_0}}$ and $x_1 = \sqrt{\frac{4H_0}{3A}}\,\mathrm{arcosh}\left(\sqrt{\frac{1}{0.05}}\right)$ is nothing but the initial position of the center of the solitary wave. This test is run with $A = 0.019\,m$, $H_0 = 1.0\,m$, $k = 8$, $n_e = 20$ and $t = 40\,s$. We show on Fig. 27 the free surface obtained with the *a posteriori* LSC method at several times in the range $[1\,s,\ 40\,s]$, showing once more a very good agreement with the reference solution obtained with a robust FV method on a very fine mesh $n_e = 10000$.

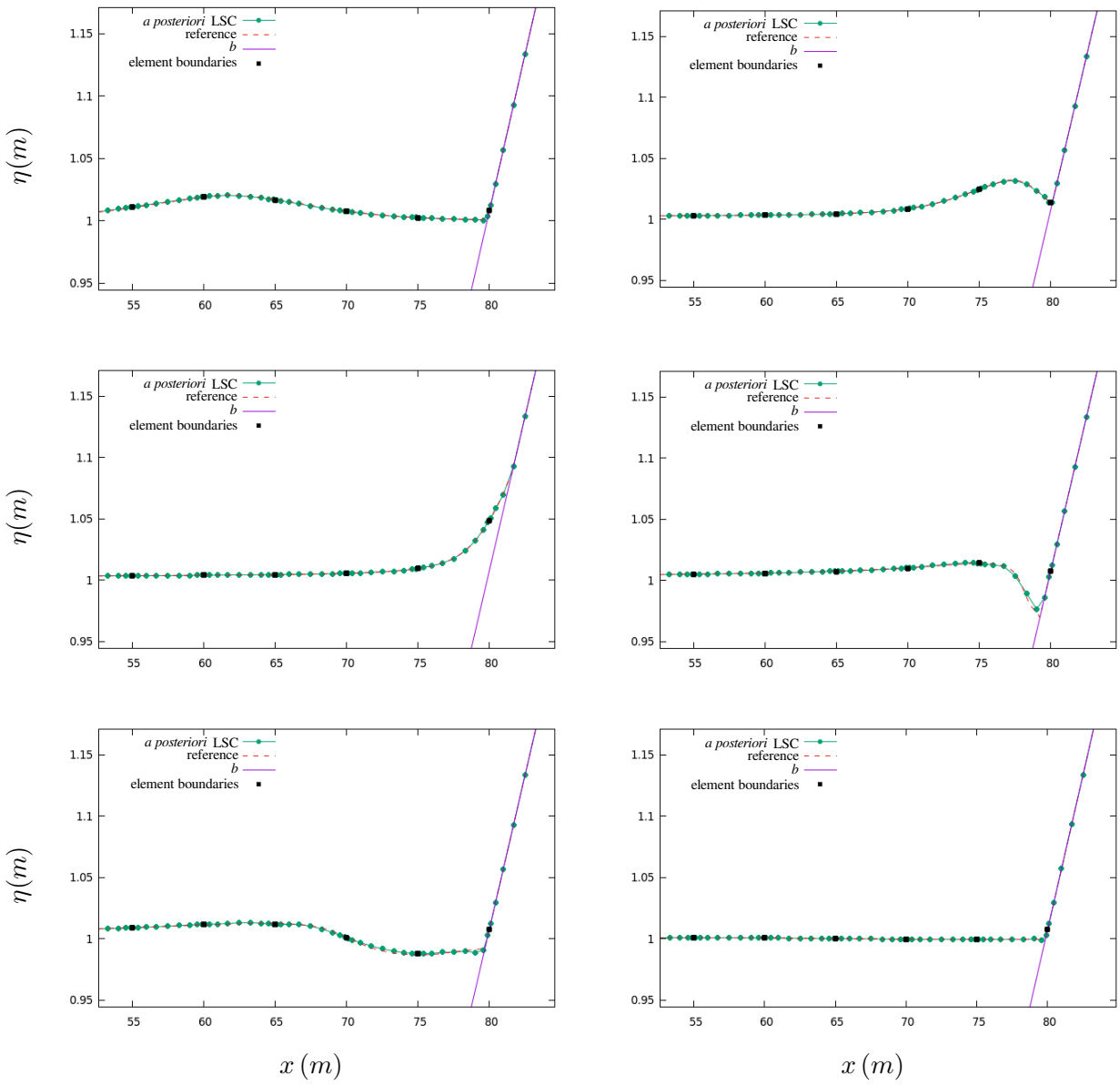

Figure 27: Test 6 - Run-up of a solitary wave on a plane beach - Free surface elevation computed for different values of time in the range $[1\,s, t = 40\,s]$ with the *a posteriori* LSC method obtained for $k = 8$ and $n_e = 20$.

5. Conclusion

In this paper, we have introduced a new well-balanced high-order discontinuous Galerkin discrete formulation with a Finite-Volume subcell correction patch designed for the NSW equations. This formulation, based on (101), combines the very high accuracy of DG schemes along with a robust correction procedure ensuring the water height positivity as well as addressing the issue of spurious oscillations in the vicinity of discontinuities. This robustness is enforced by means of an *a posteriori* local subcell correction of the conservative variables. This procedure relies on an advantageous reformulation of DG schemes as a FV-like method on a sub-grid, which makes the correction strategy surgical and flexible, as well as conservative at the subcell level. Indeed, only the non-admissible

subcells are marked and subject to correction, retaining as much as possible the very accurate subcell resolution of high-order DG formulations. The proposed strategy is investigated through an extensive set of benchmarks, including a brand new smooth solution for the computation of convergence rates, stabilization of flows with discontinuities, the preservation of motionless steady states, or moving shorelines over varying bottoms. We observe in particular that this approach provides a very accurate description of wet/dry interfaces even with the use of very high-order schemes on coarse meshes.

Regarding potential advantages of this *a posteriori* limiting strategy compared to *a priori* limiters, because the troubled zone detection is performed *a posteriori*, the correction can be done only where it is absolutely necessary. Furthermore, positivity preservation of the water height is included without any additional effort, while it is generally not the case of *a priori* limitations of high-order schemes. Let us further emphasize that this *a posteriori* LSC method scalability to any order of accuracy is also perfectly natural. Finally, it is important to note that this new correction procedure is totally parameter free.

In the future, we have the desire to extend this *a posteriori* correction technique to a general 2D case on unstructured grids for the Shallow-water system with source term. We also plan to investigate the moving mesh case, based on an arbitrary-Lagrangian-Eulerian (ALE) formalism, in the context of a coupling with a floating object.

Appendix A. New particular smooth solution for the NSW equations

This appendix aims at giving further details on the construction of a new smooth solution, of any arbitrary regularity, of the NSW equations. Following the methodology introduced in (100), we consider a smooth solution $\mathbf{v}$ in the context of flat bottom ($b = 0$), so that the NSW equations rewrite as:

$$\partial_t \mathbf{v} + \mathbf{A}(\mathbf{v})\partial_x \mathbf{v} = 0,$$

where the Jacobian matrix writes as:

$$\mathbf{A}(\mathbf{v}) = \nabla_{\mathbf{v}} \mathbf{F}(\mathbf{v}) = \begin{pmatrix} 0 & 1 \\ gH - u^2 & 2u \end{pmatrix}.$$

The eigen-analysis of the matrix $\mathbf{A}(\mathbf{v})$ leads to the following pair of eigenvalues $\lambda^{\pm} = u \pm \sqrt{gH}$ and eigenvectors:

$$E^{\pm} = \begin{pmatrix} 1 \\ u \pm \sqrt{gH} \end{pmatrix}.$$

By diagonalizing the NSW system of equations, one finally gets the following Riemann invariants $\alpha^{\pm} = u \pm 2\sqrt{gH}$, governed by the following conservation laws:

$$\partial_t \alpha^{\pm} + \lambda^{\pm} \partial_x \alpha^{\pm} = 0. \tag{A.1}$$

In light of the definition of the Riemann invariants, the system eigenvalues can be reformulated in terms of $\alpha^{\pm}$ as follows:

$$\lambda^{\pm} = \frac{\alpha^{+}(2 \pm 1) + \alpha^{-}(2 \mp 1)}{4}.$$

To uncouple the two conservation laws (A.1), we consider a particular flow regime corresponding to the trans-critical particular situation where $\alpha^{-} = 0$, *i.e.* $u = 2\sqrt{gH}$. The NSW equations then finally reduce to the following very simple Burgers equation:

$$\partial_t u + \frac{3}{2} u\, \partial_x u = 0.$$

To design a $\mathcal{C}^{N_s}$ smooth solution, we initialize the problem with the following initial data:

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

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

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

The method of characteristics provides us with the expression of the analytical solution, for any given $t \in [0, t_c[$:

$$u(x,t) = \begin{cases} 1 & \text{if } \; x \leq \frac{3}{2}\, t, \\ e^{-X^{N_s+1}} & \text{elsewhere,} \end{cases} \tag{A.2}$$

where the characteristic lines read $x(X,t) = \frac{3}{2}\, e^{-X^{N_s+1}}\, t + X$. For practical applications, to assess the position of the characteristic line origin point X given x and t, one may use an iterative root-finding process, as Newton's method, to solve the non-linear problem $g(X) = 0$, where for given x and t function $g(X) = \frac{3}{2}\, e^{-X^{N_s+1}}\, t + X - x$.

The analytical solution, (A.2), is defined $\forall\, t \in [0,\, t_c[$, where the critical time at which the characteristic lines cross is defined as follows:

$$t_c = \frac{2\; e^{\frac{N_s}{N_s+1}}}{3\; (N_s+1)^{\frac{1}{N_s+1}}\; N_s^{\frac{N_s}{N_s+1}}}.$$

$$\mathcal{F}^r_{m-\frac{1}{2}} := \mathcal{F}\left(\overline{\mathbf{v}}^+_{m-1}, \overline{\mathbf{v}}^-_m, \overline{b}^-_m\right) + \begin{pmatrix} 0 \\ g\overline{\eta}^-_m \left(\overline{b}^-_m - b_{\widetilde{x}_{m-\frac{1}{2}}}\right) \end{pmatrix}, \tag{A.3}$$

Acknowledgments. F. Marche gratefully acknowledges the support provided under Agence Nationale de la Recherche (ANR) project NABUCO (ANR-17-CE40-0025) and CNRS-LEFE-MANU project D-Wave.

[1] V. Aizinger and C. Dawson. A discontinuous Galerkin method for two-dimensional flow and transport in shallow water. *Advances in Water Resources*, 25:67–84, 2002.

[2] F. Alcrudo and P. Garcia-Navarro. A high-resolution godunov-type scheme in finite volumes for the 2d shallow-water equations. *Internat. J. Numer. Methods Fluids*, 1993.

[3] Y. Allaneau and A. Jameson. Connections between the filtered discontinuous Galerkin method and the flux reconstruction approach to high order discretizations. *Comput. Meth. Appl. Mech. Engrg.*, 200:3628–3636, 2011.

[4] V.R. Ambati and O. Bokhove. Space–time discontinuous galerkin finite element method for shallow water flows. *Journal of Computational and Applied Mathematics*, 204:452–462, 2007.

[5] K. Anastasiou and C.T. Chan. Solution of the 2d shallow water equations using the finite volume method on unstructured triangular meshes. *Int J Numer Methods Fluids*, 24:1225–1245, 1997.

[6] L. Arpia and M. Ricchiuto. Well balanced residual distribution for the ALE spherical shallow water equations on moving adaptive meshes. *J. Comput. Phys.*, 405:109–173, 2019.

[7] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comput.*, 25(6):2050–2065, 2004.

[8] D. Balsara, C. Altmann, C.D. Munz, and M. Dumbser. A sub-cell based indicator for troubled zones in RKDG schemes and a novel class of hybrid RKDG+HWENO schemes. *J. Comp. Phys.*, 226:586–620, 2007.

[9] S.R.M. Barros and J.W. Cardenas. A nonlinear galerkin method for the shallow-water equations on periodic domains. *J. Comput. Phys.*, 172:592–608, 2001.

[10] A. Bermudez, A. Dervieux, J.-A. Desideri, and M.E. Vazquez. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Comput. Methods Appl. Mech. Engrg.*, 155:49–72, 1998.

[11] Alfredo Bermudez and Ma Elena Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids*, 23(8):1049–1071, 1994.

[12] P.-E. Bernard, J.-F. Remacle, R. Comblen, V. Legat, and K. Hillewaert. High-order discontinuous galerkin schemes on general 2d manifolds applied to the shallow water equations. *J. Comput. Phys.*, 228:6514–6535, 2009.

[13] C. Berthon and F. Marche. A positive preserving high order VFRoe scheme for shallow water equations: a class of relaxation schemes. *SIAM J. Sci. Comput.*, 30(5):2587–2612, 2008.

[14] Rupak Biswas, Karen D. Devine, and Joseph E. Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14:255 – 283, 1994.

[15] O. Bokhove. Flooding and drying in discontinuous galerkin finite-element discretizations of shallow-water equations. part 1: One dimension. *J. Sci. Comput.*, 22-23:47–82, 2005.

[16] S. Bunya, E. J. Kubatko, J. J. Westerink, and C. Dawson. A wetting and drying treatment for the runge-kutta discontinuous galerkin solution to the shallow water equations. *Comput. Methods Appl. Mech. Engrg*, 198:1548–1562, 2009.

[17] A. Burbeau, P. Sagaut, and C.-H. Bruneau. A problem-independent limiter for high-order Runge-Kutta discontinuous Galerkin methods. *J. Comput. Phys.*, 169(1):111 – 150, 2001.

[18] Saray Busto, Michael Dumbser, Cipriano Escalante, Nicolas Favrie, and Sergey Gavrilyuk. On high order ader discontinuous galerkin schemes for first order hyperbolic reformulations of nonlinear dispersive systems. *Journal of Scientific Computing*, 87(2):1–47, 2021.

[19] Saray Busto, Michael Dumbser, Sergey Gavrilyuk, and Kseniya Ivanova. On thermodynamically compatible finite volume methods and path-conservative ader discontinuous galerkin schemes for turbulent shallow water flows. *Journal of Scientific Computing*, 88(1):1–45, 2021.

[20] A. Canestrelli, A. Siviglia, M. Dumbser, and E.F. Toro. Well-balanced high-order centred schemes for non-conservative hyperbolic systems. applications to shallow water equations with fixed and mobile bed. *Advances in Water Resources*, 32(6):634–644, 2009.

[21] G. Carrier and H. Greenspan. Water waves of finite amplitude on a sloping beach. *Journal of Fluid Mechanics*, 2:97—109, 1958.

[22] E. Casoni, J. Peraire, and A. Huerta. One-dimensional shock-capturing for high-order discontinuous Galerkin methods. *Int. J. Numer. Meth. Fluids*, 71:737–755, 2013.

[23] M.J. Castro Diaz, J.A. Lopez-Garcia, and C. Parès. High order exactly well-balanced numerical methods for shallow water systems. *J. Comput. Phys.*, 246:242–264, 2013.

[24] 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.

[25] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws—multi-dimensional optimal order detection (mood). *J. Comput. Phys.*, 230:4028–4050, 2011.

[26] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta Discontinuous Galerkin Method for Conservation Laws V: Multidimensional Systems. *J. Comp. Phys.*, 141:199–224, 1998.

[27] B. Cockburn and C.-W. Shu. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework. *Mathematics of Computation*, 52:411–435, 1989.

[28] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems. *J. Comput. Phys.*, 141(2):199 – 224, 1998.

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

[30] J. N. de la Rosa and C. D. Munz. Hybrid DG/FV schemes for magnetohydrodynamics and relativistic hydrodynamics. *Comp. Phys. Commun.*, 222:113–135, 2018.

[31] 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.

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

[33] S. Diot, S. Clain, and R. Loubère. Improved detection criteria for the multi-dimensional optimal order detection (MOOD) on unstructured meshes with very high-order polynomials. *Computers and Fluids*, 64:43–63, 2012.

[34] S. Diot, R. Loubère, and S. Clain. The MOOD method in the three-dimensional case: very-high-order finite volume method for hyperbolic systems. *Int. J. Numer. Meth. Fluids*, 73:362–392, 2013.

[35] D.Kuzmin. Slope limiting for discontinuous galerkin approximations with a possibly non-orthogonal taylor basis. *Int J Numer Methods Fluids*, 71(9):1178–1190, 2013.

[36] M. Dumbser and R. Loubère. A simple robust and accurate a posteriorisub-cell finite volume limiter for the discontinuousGalerkin method on unstructured meshes. *J. Comp. Phys.*, 319:163–199, 2016.

[37] M. Dumbser, O. Zanotti, R. Loubère, and S. Diot. A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *J. Comput. Phys.*, 278:47–75, 2014.

[38] 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.

[39] K.S. Erduran, V. Kutija, and C.J.M. Hewett. Performance of finite volume solutions to the shallow water equations with shock-capturing schemes. *Int J Numer Methods Fluids*, 40:1237–1273, 2002.

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

[41] C. Eskilsson and S.J.Sherwin. Discontinuous galerkin spectral/hp element modelling of dispersive shallow water systems. *J. Sci. Comput.*, 22:269–288, 2005.

[42] M. Feistauer, V. Dolejší, and V. Kučera. On the discontinuous Galerkin method for the simulation of compressible flow with wide range of Mach numbers. *Computing and Visualization in Science*, 10(1):17–27, 2007.

[43] L. Fraccarollo and E.F. Toro. Experimental and numerical assessment of the shallow water model for two-dimensional dam-break type problems. *J. Hydraulic Res.*, 33(6):843–863, 1995.

[44] J.M. Gallardo, C. Parés, and M Castro. On a well-balanced high-order finite volume scheme for shallow water equations with topography and dry areas. *J. Comput. Phys.*, 227(1):574–601, 2007.

[45] H. Gao and Z. J. Wang. A conservative correction procedure via reconstruction formulation with the Chain-Rule divergence evaluation. *J. Comp. Phys.*, 232:7–13, 2013.

[46] F.X. Giraldo, J.S. Hesthaven, and T. Warburton. Nodal high-order discontinuous galerkin methods for the spherical shallow water equations. *J. Comput. Phys.*, 181(2):499 – 525, 2002.

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

[48] N Goutal. *Proceedings of the 2nd workshop on dam-break wave simulation.* Department Laboratoire National d'Hydraulique, Groupe Hydraulique Fluviale, 1997.

[49] J.-L. Guermond, R. Pasquetti, and B. Popov. Entropy viscosity method for nonlinear conservation laws. *J. Comput. Phys.*, 230:4248–4267, 2011.

[50] R. Harris, Z.J. Wang, and Y. Liu. Efficient Quadrature-Free High-Order Spectral Volume Method on Unstructured Grids: Theory and 2D Implementation. *J. Comp. Phys.*, 227:1620–1642, 2008.

[51] A. Huerta, E. Casoni, and J. Peraire. A simple shock-capturing technique for high-order discontinuous galerkin methods. *Int. J. Numer. Meth. Fluids*, 69:1614–1632, 2012.

[52] H. T. Huynh. A Flux Reconstruction Approach to High-Order Schemes Including Discontinuous Galerkin Methods. In *18th AIAA Computational Fluid Dynamics Conference, Miami*, 2007.

[53] H. T. Huynh, Z. J. Wang, and P. E. Vincent. High-order methods for computational fluid dynamics: A brief review of compact differential formulations on unstructured grids. *J. Comp. Phys.*, 98:209–220, 2014.

[54] M. Ioriatti and M. Dumbser. A posteriori sub-cell finite volume limiting of staggered semi-implicit discontinuous Galerkin schemes for the shallow water equations. *Applied Numerical Mathematics*, 135:443–480, 2019.

[55] M. Iskandarani, D. B. Haidvogel, and J. P. Boyd. A staggered spectral element model with application to the oceanic shallow water equations. *Int J Numer Methods Fluids*, 20(5):393–414, 1995.

[56] J. S. Park and C. Kim. Hierarchical multi-dimensional limiting strategy for correction procedure via reconstruction. *J. Comp. Phys.*, 308:57–80, 2016.

[57] J. S. Park and S.-H. Yoon and C. Kim. Multi-dimensional limiting process for hyperbolic conservation laws on unstructured grids. *J. Comp. Phys.*, 229:788–812, 2010.

[58] G. Jiang and C.-W. Shu. Efficient implementation of weighted eno schemes. *J. Comput. Phys.*, 126(1):202–228, 1996.

[59] G. Kesserwani and Q. Liang. A discontinuous Galerkin algorithm for the two-dimensional shallow water equations. *Comput. Methods Appl. Mech. Engrg.*, 199(49-52):3356–3368, 2010.

[60] G. Kesserwani and Q. Liang. Well-balanced RKDG2 solutions to the shallow water equations over irregular domains with wetting and drying. *Comput. Fluids*, 39(10):2040–2050, 2010.

[61] G. Kesserwani, Q. Liang, J. Vazquez, and R. Mose. Well-balancing issues related to the RKDG2 scheme for the shallow water equations. *Int. J. Numer. Meth. Fluids*, 62:428–448, 2010.

[62] R.M. Kirby and S.J. Sherwin. Stabilisation of spectral / hp element methods through spectral vanishing viscosity: Application to fluid mechanics. *Comput. Methods Appl. Mech. Engrg.*, 195:3128–3144, 2006.

[63] L. Krivodonova. Limiters for high-order discontinuous Galerkin methods. *J. Comp. Phys.*, 226:879–896, 2007.

[64] B. Van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 135(2):227–248, 1997.

[65] H. Li and R.-X. Liu. The discontinuous Galerkin finite element method for the 2d shallow water equations. *Mathematics and Computers in Simulation*, 56:223–233, 2001.

[66] L. Li and Q. Zhang. A new vertex-based imiting approach for nodal discontinuous galerkin methods on arbitrary unstructured meshes. *Comput. Fluids*, 159:316–326, 2017.

[67] C. Liang, F. Ham, and E. Johnsen. Discontinuous galerkin method with weno limiter for flows with discontinuity. *Center for Turbulence Research 335 Annual Research Briefs 2009*, 2009.

[68] Q. Liang and A. G. L. Borthwick. Adaptive quadtree simulation of shallow flows with wet-dry fronts over complex topography. *Comput. Fluids*, 38(2):221–234, 2009.

[69] 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.

[70] M. Lukacova, S. Noelle, and M. Kraft. Well-balanced finite volume evolution galerkin methods for the shallow water equations. *J. Comput. Phys.*, 221(1):122–147, 2007.

[71] Hong Ma. A spectral element basin model for the shallow water equations. *J. Comput. Phys.*, 109(1):133 – 149, 1993.

[72] A. Meister and S. Ortleb. On unconditionally positive implicit time integration for the DG scheme applied to shallow water flows. *Int. J. Numer. Meth. Fluids*, 76::69–94, 2014.

[73] A. Meister and S. Ortleb. A positivity preserving and well-balanced DG scheme using finite volume subcells in almost dry regions. *Appl. Math. Comp.*, 272:259–273, 2016.

[74] V. Michel-Dansac. A well-balanced scheme for the shallow-water equations with topography. *Computers and Mathematics with Applications*, 72(3):568–593, 2016.

[75] H. Mirzaee, L. Ji, J.K. Ryan, and R.M Kirby. Smoothness-increasing accuracy-conserving (SIAC) post- processing for discontinuous Galerkin solutions over structured triangular meshes. *SIAM J. Numer. Anal.*, 49(5):1899–1920, 2011.

[76] R. D. Nair, S. J. Thomas, and R. D. Loft. A discontinuous galerkin global shallow water model. *Monthly Weather Review*, 133:876–888, 2004.

[77] I.M. Navon. Finite-element simulation of the shallow-water equations model on a limited-area domain. *Appl. Math. Modelling*, 3, 1979.

[78] S. Noelle, N. Pankratz, G. Puppo, and J.R. Natvig. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *J. Comput. Phys.*, 213(2):474–499, 2006.

[79] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced finite volume weno schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1):29–58, 2007.

[80] H. T. Ozkan-Haller and J.T.Kirby. A fourier-chebyshev collocation method for the shallow water equations including shoreline. *Applied Ocean Research*, 19:21–34, 1997.

[81] P.-O.Persson and J.Peraire. Sub-cell shock capturing for discontinuous galerkin methods. *AIAA Aerospace Sciences Meeting and Exhibit*, 112, 2006.

[82] K.T. Panourgiasa and J.A. Ekaterinaris. A nonlinear filter for high order discontinuous Galerkin discretizations with discontinuity resolution within the cell. *J. Comput. Phys.*, 326:234–257, 2016.

[83] J. Patera and V. Nassehi. A new two-dimensional finite element model for the shallow water equations using a lagrangian framework constructed along fluid particle trajectories. *Int. J. Numer. Meth. Fluids*, 39:4159–4182, 1996.

[84] Benoît Perthame and Youchun Qiu. A variant of Van Leer's method for multidimensional systems of conservation laws. *J. Comput. Phys.*, 112(2):370–381, 1994.

[85] J. Qiu and C.-W. Shu. A comparison of troubled-cell indicators for Runge-Kutta discontinuous Galerkin methods using weighted essentially nonoscillatory limiters. *SIAM J. Sci. Comput.*, 27:995–1013, 2005.

[86] J. Qiu and C.-W. Shu. Runge Kutta discontinuous Galerkin method using WENO limiters. *SIAM J. Sci. Comput.*, 26:907–929, 2005.

[87] M. Ricchiuto. An explicit residual based approach for shallow water flows. *J. Comput. Phys.*, 280:306–344, 2015.

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

[89] B. Rogers, M. Fujihara, and A. Borthwick. Adaptive Q-tree Godunov-type scheme for shallow water equations. *Int. J. Numer. Methods Fluids*, 35:247–280, 2001.

[90] D. Schwanenberg and M. Harms. Discontinuous galerkin finite-element method for transcritical two-dimensional shallow water flows. *J.Hydraul. Eng.*, 130(5):412–421, 2004.

[91] C.-W. Shu and S. Osher. Efficient implementation of Essentially Non-Oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77:439–471, 1988.

[92] M. Sonntag and C.D. Munz. Shock capturing for discontinuous Galerkin methods using finite volume subcells. In *Finite Volumes for Complex Applications VII-Elliptic, Parabolic and Hyperbolic Problems*, pages 945–953, 2014.

[93] 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.

[94] J. Tanner and E. Tadmor. Adaptive mollifiers - high resolution recover of piecewise smooth data from its spectral information. *Found. Comput. Math.*, 2:155–189, 2002.

[95] E.F. Toro. *Shock-capturing methods for free-surface shallow flows.* Chichester: John Wiley and Sons, 2001.

[96] T. Utnes. A finite element solution of the shallow-water wave equations. *Appl. Math. Modelling*, 14:20–29, 1990.

[97] J.J.W. Van der Vegt and H. Van der Ven. Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flows. *J. Comput. Phys.*, 182:546–585, 2002.

[98] H. Vandeven. Family of spectral filters for discontinuous problems. *J. Sci. Comput.*, 8:159–192, 1991.

[99] S. Vater, N. Beisiegel, and J. Behrens. A limiter-based well-balanced discontinuous galerkin method for shallow-water flows with wetting and drying: One-dimensional case. *Advances in Water Resources*, 85:1–13, 2015.

[100] F. Vilar. *A high-order discontinuous Galerkin discretization for solving two-dimensional Lagrangian hydrodynamics*. PhD thesis, Université Bordeaux I, 2012.

[101] F. Vilar. A posteriori correction of high-order discontinuous galerkin scheme through subcell finite volume formulation and flux reconstruction. *J. Comput. Phys.*, 387:245–279, 2019.

[102] F. Vilar, P.H. Maire, and R. Abgrall. Cell-centered discontinuous Galerkin discretizations for two-dimensional scalar conservation laws on unstructured grids and for one-dimensional Lagrangian hydrodynamics. *Comput. Fluids*, 46:498–504, 2011.

[103] P. E. Vincent, P. Castonguay, and A. Jameson. A New Class of High-Order Energy Stable Flux Reconstruction Schemes. *J. Sci. Comput.*, 47:50–72, 2011.

[104] S. Vukovic. Eno and weno schemes with the exact conservation property for one-dimensional shallow water equations. *J. Comput. Phys.*, 179(2):593–621, 2002.

[105] Z.J. Wang. High-Order Spectral Volume Method for Benchmark Aeroacoustic Problems. *AIAA Paper*, 2003-0880, 2003.

[106] D. Wirasaet, E.J. Kubatko, C.E. Michoski, S. Tanaka, and J.J. Westerink. Discontinuous Galerkin methods with nodal and hybrid modal/nodal triangular, quadrilateral, and polygonal elements for nonlinear shallow water flow. *Comput. Methods Appl. Mech. Engrg.*, 270:113–149, 2014.

[107] Y. Xing and C.-W. Shu. High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *J. Comput. Phys.*, 214:567–598, 2006.

[108] Y. Xing and C.-W. Shu. A new approach of high order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *Commun. Comput. Phys.*, 1:100–134, 2006.

[109] 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.

[110] 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.

[111] M. Yang and Z.J. Wang. A parameter-free generalized moment limiter for high-order methods on unstrucured grids. *Adv. Appl. Math. Mech.*, 4:451–480, 2009.

[112] 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.

[113] X. Zhang and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. A*, 467:2752–2776, 2011.

[114] X. Zhang, Y. Xia, and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes. *J. Sci. Comput.*, 50:29–62, 2012.

[115] J. Zhu, J. Qiu, C.-W. Shu, and M. Dumbser. Runge–Kutta discontinuous Galerkin method using WENO limiters II: Unstructured meshes. *J. Comput. Phys.*, 227:4330–4353, 2008.

[116] J. Zhu, X. Zhong, C.-W. Shu, and J. Qiu. Runge Kutta discontinuous Galerkin method using a new type of WENO type limiters on unstructured meshes. *J. Comp. Phys.*, 248:200–220, 2013.