HAL
archives-ouvertes.fr

The MOOD method for the non-conservative shallow-water system

S Clain, J Figueiredo

▶ To cite this version:

S Clain, J Figueiredo. The MOOD method for the non-conservative shallow-water system. 2014. hal-01077557

HAL Id: hal-01077557
https://hal.archives-ouvertes.fr/hal-01077557

Preprint submitted on 25 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The MOOD method for the non-conservative shallow-water system

S. Clain [a], J. Figueiredo [a]

[a] *Centre of Mathematics, Minho University, Campus de Gualtar - 4710-057 Braga, Portugal*

Abstract

We present an adaptation of the MOOD method, initially introduced in [1,2], for the two-dimensional shallow-water system with varying bathymetry, where the major novelty of the study is the non-conservative term discretization in the framework of the MOOD strategy. We derive a robust sixth-order scheme and propose a large panel of numerical tests to assess the accuracy of the method and show that numerical solutions are free of oscillations in the vicinity of discontinuities. We also demonstrate that the MOOD method guarantees the height positivity as long as the first-order scheme does.

Key words: Finite volume; High-order; Non-conservative problem; Polynomial reconstruction; Unstructured mesh; Shallow-water; MOOD; Positivity-preserving.

Contents

Email addresses: clain@math.uminho.pt (S. Clain), jmfiguei@math.uminho.pt (J. Figueiredo).

[1] Corresponding author: J. Figueiredo; Address: Centre of Mathematics, Minho University, Campus de Gualtar - 4710-057 Braga, Portugal; Telephone: +351253604367; Email address: jmfiguei@math.uminho.pt.

1 Introduction

High-order finite volume approximations for non-conservative hyperbolic systems have attracted much attention in the past three decades. More specifically, the shallow-water system with varying bathymetry is quite studied due to the large spectrum of applications such as environmental sciences or civil engineering (river, ocean, tsunami, flooding, wastewater, [3–5]), even if the Saint-Venant model is doubtful in some situations (see [6] for a rigorous derivation of the shallow-water model from the Navier-Stokes equation). A lot of numerical techniques have been proposed and tested in the shallow-water context such as the Discontinuous Galerkin method (see [7,8] for example), but the finite volume method is the usual framework to provide numerical approximations due to the conservation built-in property for both the mass and the impulsion.

A first issue concerns the accuracy and the robustness of the method. It is expected that the numerical scheme provides accurate approximations for smooth solutions, while sharp discontinuities are correctly located with very low numerical diffusion and few oscillations in the vicinity of shocks. Second-order techniques based on linear reconstructions and limiting procedures have been first developed and are still very popular due to their simplicity, efficiency, versatility and low computational cost (see [9–11] for instance). Nevertheless, numerous situations involving coarse meshes require third- or even higher-order numerical procedures to reduce the undesirable diffusion and to provide

nice steep shocks. The ENO/WENO method has received a lot of contributions in the shallow-water context [12–14] and turns out to be the most common way to achieve up to fifth-order approximations. A very high-order approximation also requires cute time discretization and the Total Variation Diminishing Runge-Kutta schemes (TVD-RK algorithm [15]) is one of the most popular techniques. The ADER method is a powerful alternative [16] and was recently used in the shallow-water context [17]. Beginning 2011, a new very high-order technique (up to sixth-order of accuracy), named the Multidimensional Optimal Order Detection (MOOD), has been proposed in [1,2] for two-dimensional geometries and has been extended in [18] for the three-dimensional case. Other applications of the MOOD strategy are proposed in [19–22]. The method is said *a posteriori* in contrast with the former *a priori* methods such as MUSCL or WENO for the following reasons. In MUSCL [23] and WENO [24], the polynomial reconstruction is altered/corrected in such a way that the reconstructed values at cell interfaces plugged into the numerical flux at time t^n provide an approximation at time t^{n+1} which satisfies some stability properties (positivity-preserving, entropy, low oscillations). As a consequence, the polynomials are modified before computing the updated numerical solution thus justifying the mention *a priori*.
In the MOOD method, we do not perform any correction of the polynomial reconstruction and first compute the fluxes with polynomials of highest degrees. We then determine a candidate solution after the time update which may, of course, present oscillations in the vicinity of shocks and discontinuities. A detecting procedure determines the cells which are problematic and we reduce the polynomial degree only for these specific cells. We iterate the degree reduction for problematic cells until the solution is considered eligible. The term *a posteriori* refers to the fact that we correct the polynomial reconstruction after evaluating the updated candidate and that we perform the limitation based on the updated approximation and not on the numerical solution at time t^n. Since the MOOD method has been experimented in the hyperbolic conservative system context, a first issue we shall address in the article is to demonstrate the capacity of the method to handle non-conservative problems as well.

The second difficulty we face concerns the non-conservative term and one has to propose adequate numerical schemes to correctly discretize the varying bathymetry term. From a theoretical point of view, non-conservative product has been defined by [25] using path integration in the space of the states and some schemes have been designed with respect to the theoretical framework [17,26–28]. Moreover, the non-conservative term is responsible for critical situations such as the resonant state: the system is no longer strictly hyperbolic and the contact discontinuity associated to the bathymetry variation may merge, for example, with a genuinely nonlinear steady shock. In 1994, Bermdez and Vsquez proposed a fundamental guideline to design correct numerical schemes introducing the C-property [29,30] that we can sum-up in the

following way: the schemes have to preserve the lake at rest (or water at rest solution). Such schemes introduce an upwinding in the non-conservative term discretization [31] (or Q-scheme [32]) to compensate the numerical flux deriving from the conservative contribution, the so-called well-balanced scheme. From that seminal paper, a large class of numerical schemes which satisfy the C-property have been proposed and analysed during the past two decades involving high-order WENO methods [12] or dry-wet situations [33,34] still preserving the C-property. We would like to mention an alternative approach using centred schemes with staggered grids (see [35–37]). Noticing that the lake at rest condition consists in preserving a specific steady-state solution of the shallow-water system, an extension of the C-property has been proposed where the authors intend to preserve all the regular steady-state solutions (moving water solution) up to a certain order [12,13,38] or exactly [14,27]. Preserving all steady-state solutions increases the scheme accuracy for transient flows close to steady-state configurations and better captures small perturbations superposed to a moving water equilibrium [39].
In a similar way, another important issue concerns the determination of accurate solutions for discontinuous topography where the challenge is to correctly solve the Riemann problems. A rich set of configurations have been found where the genuinely nonlinear simple wave merges with the 0-wave, or is split into two waves associated to the same eigenvalue. We refer to [40–44] for the determination of all the Riemann problem solutions with hydraulic jump and the numerical treatment of the resonant configurations. Notice that the non-uniqueness of the solution for that specific situation remains an open problem up to the authors knowledge since there exist multiple solutions in such a regime [43,45].
The last important issue we shall not tackle in the present study corresponds to the dry-wet situations which are of crucial importance from a practical point of view (dam-break, tsunami, coastal flooding). An important key has been proposed in Audusse *et al.* [33] based on the hydrostatic reconstruction, while a generalisation of the hydrostatic reconstruction has been proposed by Castro Días *et al.* in [27] in the context of the moving water solution preservation.

We propose a MOOD extension for the non-conservative shallow-water problem for two-dimensional geometries and unstructured meshes. The goal is to prove that the MOOD method initially developed in the conservative Euler system context [1,2,18] is an efficient alternative to the WENO technique and provides both sixth-order approximations and robust numerical solutions without spurious oscillations. We use the classical Rusanov, HLL, and HLLC numerical fluxes for the conservative contribution and we propose a discretization of the non-conservative term which preserves the lake at rest by introducing a new key notion: the physical bathymetry representative. The scheme we propose here is designed to exactly respect the C-property but no specific care is taken for the other steady-state solutions. Nevertheless, as we shall see in the numerical tests section, the MOOD method enables to preserve the

regular steady-state up to an effective sixth-order as in [12,13,38].

The paper is organised as follows. In the second section we present the key ingredients to design a very high-order scheme where we introduce the notion of physical bathymetry representative and the conservative polynomial reconstruction for two-dimensional unstructured meshes. Section 3 is dedicated to a presentation of the MOOD method where the essential notions such as the $\mathcal{A}$-eligible set and the detection criteria are detailed. The fourth section presents a large panel of numerical experiences to assess the scheme capacity to provide accurate and robust numerical approximations of the solutions.

2 Finite volume scheme

The shallow-water system equipped with the non-conservative term deriving from the varying bathymetry writes

$$\begin{aligned}
&\partial_t h + \nabla.(hU) = 0,\\
&\partial_t(hU) + \nabla.(hU \otimes U + \frac{1}{2}gh^2 I_2) = -gh\nabla b,
\end{aligned}$$

where h is the water height, $U = (u, v)^T$ the velocity, $U \otimes U$ the tensorial product, $Q = hU$ the mass flow, I_2 the $\mathbb{R}^2$ identity matrix, b the bathymetry with respect to a reference level and g the gravitational acceleration. Following [46,47] (see also [48]), we adopt the surface method gradient formulation using the total height (or free surface) $H = h + b$ as a conservative variable and we introduce the augmented system adding the bathymetry function b as an unknown function. Setting $V = (H, hu, hv, b)^T$ and $\mathbf{x} = (x, y)^T$, the model writes in a compact form $\partial_t V + \partial_x F(V) + \partial_y G(V) = S(V)$, with

$$F = \begin{pmatrix} hu \\ hu^2 + \frac{g}{2}H(H-2b) \\ huv \\ 0 \end{pmatrix}, G = \begin{pmatrix} hv \\ huv \\ hv^2 + \frac{g}{2}H(H-2b) \\ 0 \end{pmatrix}, S = -\begin{pmatrix} 0 \\ gH\partial_x b \\ gH\partial_y b \\ 0 \end{pmatrix}.$$

For a given direction characterized by the unit vector n, we denote by $Fr = U.n/\sqrt{gh}$ the associated Froude number. The system is strictly hyperbolic when $Fr \neq \pm 1$, with two genuinely nonlinear waves and two contact discontinuities, one associated to the tangential velocity $U.\tau$, with τ orthogonal to n, and one associated to the null eigenvalue deriving from the bathymetry equation. We recall that $Fr = \pm 1$ corresponds to situations where two eigenvalues are superposed. In that case, strict hyperbolicity does not hold any longer

leading to resonance situations where two simple waves merge. The boundary of the domain $\partial\Omega = \Gamma_D \cup \Gamma_N$ is composed of two non-overlapping parts where we shall prescribe Dirichlet conditions (Γ_D) and reflection/transmission conditions (Γ_N).

2.1 Discretization

We propose a new finite volume scheme to achieve very high-order approximations (up to sixth-order of convergence) based, on the one hand, on local polynomial reconstructions to provide the accuracy and, on the other hand, the MOOD methodology to guarantee the stability and avoid non-physical oscillations close to discontinuities. We introduce the following notations illustrated in Figure 1 to design the numerical scheme. The computational domain Ω is assumed to be a polygonal bounded set of $\mathbb{R}^2$ divided into polygonal cells c_i with m_i the cell centroid, $i \in \mathcal{E}_{el}$ the cell index set. For a given cell c_i, we denote by e_{ij} the edges of c_i such that

- $j \in \mathcal{E}_{el}$ if there exists an adjacent cell c_j with $e_{ij} = c_i \cap c_j$;
- $j = D$ if $e_{iD} = c_i \cap \Gamma_D$;
- $j = N$ if $e_{iN} = c_i \cap \Gamma_N$.

To avoid a specific treatment of the boundary edges we introduce $\widetilde{\mathcal{E}_{el}} = \mathcal{E}_{el} \cup \{D, N\}$, the cell index set augmented with index D for the Dirichlet condition and N for the reflection/transmission condition. We then define the set ν_i of all the indexes $j \in \widetilde{\mathcal{E}_{el}}$ such that e_{ij} is an edge of c_i.
For each edge e_{ij}, $i \in \mathcal{E}_{el}$, $j \in \nu_i$, n_{ij} stands for the unit normal vector going from c_i to c_j and τ_{ij} is the unit tangent vector such that n_{ij}, τ_{ij} is a counterclockwise oriented basis. We denote by m_{ij} the edge midpoint, while $\left(\xi_r, q_{ij,r}\right)$, $r = 1, ..., R$ stands for the quadrature rule for the numerical integration on e_{ij}, where ξ_r is the weight associated to the r^{th} quadrature point $q_{ij,r}$. If index $j = D$ (resp. $j = N$), n_{iD} and τ_{iD} represent the outward unit normal vector and unit tangent vector while m_{iD} and $q_{iD,r}$ are the edge midpoint and Gauss points.

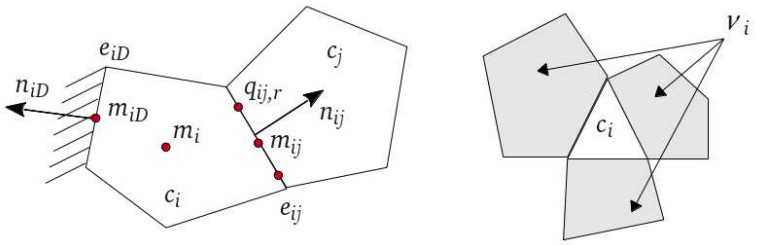

Fig. 1. Mesh and notations (left). Definition of index set ν_i (right).

The generic high-order finite volume scheme associated to the shallow-water system writes

$$V_i^{n+1} = V_i^n - \Delta t \sum_{j \in \nu_i} \frac{|e_{ij}|}{|c_i|} \sum_{r=1}^{R} \xi_r (\mathcal{F}_{ij,r}^n + \varepsilon_{ij,r}^n) + \Delta t \mathcal{S}_i^n, \tag{1}$$

where V_i^n is an approximation of the mean value of V at time t^n on cell c_i, Δt stands for the time step, $|e_{ij}|$ and $|c_i|$ are, respectively, the length of edge e_{ij} and the area of cell c_i. Vector $\mathcal{F}_{ij,r}$ represents a numerical approximation of the conservative flux at Gauss point $q_{ij,r}$ and $\mathcal{S}_i^n$ stands for an approximation of the geometrical source term over the cell c_i. At last, the term $\varepsilon_{ij,r}^n$ concerns the non-conservative contribution due to the bathymetry discontinuity across the edge e_{ij}.
To achieve high-order approximations, polynomial reconstructions are involved to produce local representations of the approximation (see [1,2,18] for the conservative case and [49] for the extension to the diffusive flux case). We recall here the fundamental lines of the reconstruction for the sake of consistency and to introduce the notations.
For a given cell c_i and a polynomial degree d, we associate the stencil $S(c_i, d)$ constituted of cells we pick-up around the reference cell c_i. For any variable $\phi = H, hu, hv, b$, we shall denote by $\boldsymbol{\phi}_i(\mathbf{x}; d)$ a local polynomial function of degree d associated to cell c_i with the following structure

$$\boldsymbol{\phi}_i(\mathbf{x}; d) = \phi_i \; + \sum_{1 \leq |\alpha| \leq d} \mathcal{R}_i^\alpha \Big((\mathbf{x} - m_i)^\alpha - M_i^\alpha \Big),$$

with ϕ_i an approximation of the ϕ mean value on cell c_i, $\alpha = (\alpha_1, \alpha_2)$ the multi-index, $|\alpha| = \alpha_1 + \alpha_2$ (see [18] for a detailed description) and

$$M_i^\alpha = \frac{1}{|c_i|} \int_{c_i} (\mathbf{x} - m_i)^\alpha \, d\mathbf{x},$$

such that the following conservativity property holds

$$\frac{1}{|c_i|} \int_{c_i} \boldsymbol{\phi}_i(\mathbf{x}; d) \, d\mathbf{x} = \phi_i.$$

To compute the reconstruction coefficients, we introduce the quadratic functional

$$E_i(\mathcal{R}_i) = \sum_{\ell \in S(c_i, d)} \left(\frac{1}{|c_\ell|} \int_{c_\ell} \boldsymbol{\phi}_i(\mathbf{x}; d) \, d\mathbf{x} - \phi_\ell \right)^2,$$

where ϕ_ℓ are approximated mean values on cells c_ℓ of the stencil and $\mathcal{R}_i = (\mathcal{R}_i^\alpha)_{1 \leq |\alpha| \leq d}$ is the vector which gathers all the components. We seek for vector $\mathcal{R}_i$ which minimises the functional and denote by $\widehat{\phi}_i(\mathbf{x}; d)$ the associated polynomial. In [49], a detailed presentation of the method is given to provide

the solution $\mathcal{R}_i$. Applying the reconstruction process to the conservative variables, namely $\phi = H, hu, hv, b$, provides a vectorial polynomial reconstruction $\widehat{V}_i(\mathbf{x}; d)$ for each cell c_i.

Let e_{ij} be a side of c_i and set $V_{ij,r} = \widehat{V}_i(q_{ij,r}; d)$ (we skip the time index for the sake of simplicity). To define $V_{ji,r}$, three situations have to be considered:

- if $e_{ij} = c_i \cap c_j$ with $j \in \mathcal{E}_{el}$ then we set $V_{ji,r} = \widehat{V}_j(q_{ij,r}; d)$;
- if the edge is on Γ_D then we set $V_{Di,r} = V_D(q_{iD,r})$ where V_D is given on the boundary;
- if the edge is on Γ_N then we define $V_{Ni,r}$ on point $q_{iD,r}$ with $H_{Ni,r} = H_{iN,r}$, $b_{Ni,r} = b_{iN,r}$, $U_{Ni,r} = U_{iN,r} - 2(U_{iN,r} \cdot n_{iN})n_{iN}$ to provide the reflection condition. When dealing with transmission conditions for a given variable ϕ we just set $\phi_{Ni,r} = \phi_{iN,r}$.

Then the numerical conservative flux across edge e_{ij} following direction n_{ij} will take the form

$$\mathcal{F}_{ij,r} = \mathbb{F}(V_{ij,r}, V_{ji,r}; n_{ij}),$$

which depends only on the values on both sides of the Gauss point $q_{ij,r}$ and on n_{ij}, while the source term takes the generic form

$$\mathcal{S}_i = \mathbb{S}(\widehat{H}_i, \widehat{b}_i).$$

Explicit expressions for S_i and $\varepsilon^n_{ij,r}$ will be provided in the next section.
Vector $V^n = (V^n_i)_{i \in \mathcal{E}_{el}}$ gathers the mean value approximations at time t^n and we introduce the vectorial operator $\mathcal{H}(V^n)$ such that relation (1) rewrites as the forward Euler time discretization

$$V^{n+1} = V^n + \Delta t \mathcal{H}(V^n). \tag{2}$$

From the original first-order discretization in time given by relation (2) we derive high-order approximation in time. For instance, we shall use the so–called Runge–Kutta TVD-RK3 method [1,2]:

$$V^{n+1} = \frac{V^n + 2V^{(3)}}{3} \quad \text{with} \quad \begin{cases} V^{(1)} = V^n + \Delta t\ \mathcal{H}(V^n)\,, \\ V^{(2)} = V^{(1)} + \Delta t\ \mathcal{H}(V^{(1)})\,, \\ V^{(3)} = V^{\flat(2)} + \Delta t\ \mathcal{H}(V^{\flat(2)})\,, \end{cases}$$

where $V^{\flat(2)}$ is the convex combination $(3V^n + V^{(2)})/4$.
This time discretization introduces a 3^{rd}-order error which makes the whole scheme to be formally 3^{rd}-order accurate. However, setting $\Delta t = \Delta x^{r/3}$, where r is the spatial order of accuracy and Δx is a characteristic length, provides the same order for both the spatial and the time errors.

Numerical schemes for non-conservative problems are usually decomposed into two parts: a numerical flux to handle the conservative contribution associated to F and G and a discretization of the non-conservative term, which deeply depends on the choice of the conservative numerical flux, to achieve the well-balanced property. To provide an approximation of the source term, a natural choice would be

$$\mathbb{S}(\widehat{H}_i^n, \widehat{b}_i) = - \begin{pmatrix} 0 \\ \dfrac{1}{|c_i|} \displaystyle\int_{c_i} g\widehat{H}_i^n(\mathbf{x}) \nabla \widehat{b}_i(\mathbf{x}) \, d\mathbf{x} \\ 0 \end{pmatrix} \tag{3}$$

to mimic the physical non-conservative $gH\nabla b$ term substituting the local polynomial representation $g\widehat{H}_i\nabla\widehat{b}_i$ on each cell c_i. Unfortunately, expression (3) is not satisfactory since steady-state situations such as the lake at rest are not preserved (the C-property). To overcome this problem, we shall introduce a corrective term $\varepsilon_{ij,t}^n$ in the expression of the non-conservative part.

We first recall the classical properties that a numerical conservative flux has to fulfil, namely the consistency and the conservativity,

$$\mathbb{F}(V, V; n) = F(V)n_x + G(V)n_y, \qquad \mathbb{F}(V, W; n) = \mathbb{F}(W, V; -n).$$

To link the numerical flux for the conservative contribution with the non-conservative term discretization, we shall introduce in the sequel a new bathymetry function $b^\star = b^\star(V_L, V_R)$, continuous with respect to the physical states, and we require that the numerical flux equipped with $b^\star$ fulfils a new property.

Definition 2.1 (physical bathymetry representative)

- *The bathymetry function $b^\star = b^\star(V_L, V_R)$ is convex if for any left and right states V_L, V_R, there exists $\theta = \theta(V_L, V_R) \in [0, 1]$ such that $b^\star = (1-\theta)b_L + \theta b_R$.*
- *Let us denote by $V_L = (\bar{H}, 0, 0, b_L)^T$ and $V_R = (\bar{H}, 0, 0, b_R)^T$ the left and right states corresponding to the lake at rest with $\bar{H} > \max(b_L, b_R)$. We say that $b^\star$ is a physical bathymetry representative for $\mathbb{F}$ if the state $V^\star = (\bar{H}, 0, 0, b^\star)^T$ satisfies*

$$\mathbb{F}(V_L, V_R; n) = F(V^\star)n_x + G(V^\star)n_y. \tag{4}$$

Based on this definition, we introduce the corrective term for the non-conservative discretization and show that the scheme (1) satisfies the C-property.

Theorem 2.2 *Let $V_{ij,r}^n$ and $V_{ji,r}^n$ be the left and right approximations at the Gauss point $q_{ij,r}$ with respect to the n_{ij} orientation at time t^n. We define the non-conservative flux as*

$$\varepsilon_{ij,r}^n = -gH_{ij,r}^n \begin{pmatrix} 0 \\ (b_{ij,r}^\star - b_{ij,r})n_{ij} \\ 0 \end{pmatrix}. \tag{5}$$

Assume that $b^\star$ is a physical bathymetry representative for $\mathbb{F}$, i.e. for $\bar{H}$ greater than $b_{ij,r}, b_{ji,r}$, and $V_{ij,r} = (\bar{H}, 0, 0, b_{ij,r})^T$, $V_{ji,r} = (\bar{H}, 0, 0, b_{ji,r})^T$, $b_{ij,r}^\star = b^\star(V_{ij,r}, V_{ji,r})$, we have

$$\mathbb{F}(V_{ij,r}, V_{ji,r}; n) = F\Big((\bar{H}, 0, 0, b_{ij,r}^\star)^T\Big)n_x + G\Big((\bar{H}, 0, 0, b_{ij,r}^\star)^T\Big)n_y.$$

Then, the scheme

$$V_i^{n+1} = V_i^n - \Delta t \sum_{j\in\nu_i} \frac{|e_{ij}|}{|c_i|} \sum_{r=1}^{R} \xi_r \Big[\mathbb{F}(V_{ij,r}^n, V_{ji,r}^n; n_{ij}) + \varepsilon_{ij,r}^n\Big] + \Delta t \mathbb{S}_i^n \tag{6}$$

with $\mathbb{S}_i^n$ given by (3), satisfies the C-property.

Proof. Assume that at time t^n we have a lake at rest, *i.e.* $H_i^n = \bar{H}$, $Q_i^n = (0,0)^T$ for all $i \in \mathcal{E}_{el}$. For the sake of consistency the polynomial reconstruction provides a constant function and we have $H_{ij,r}^n = H_{ji,r}^n = \bar{H}$ and $Q_{ij,r}^n = Q_{ji,r}^n = (0,0)^T$. From relation (4) we deduce the flux expression

$$\sum_{j\in\nu_i} \frac{|e_{ij}|}{|c_i|} \sum_{r=1}^{R} \xi_r \Big[\mathbb{F}(V_{ij,r}^n, V_{ji,r}^n, n_{ij}) + \varepsilon_{ij,r}^n\Big] =$$

$$\sum_{j\in\nu_i} \frac{|e_{ij}|}{|c_i|} \sum_{r=1}^{R} \xi_r \frac{g}{2} \begin{pmatrix} 0 \\ \Big[\bar{H}(\bar{H} - 2b_{ij,r}^\star) - \bar{H}(2b_{ij,r}^\star - 2b_{ij,r})\Big]n_{ij} \\ 0 \end{pmatrix} =$$

$$g\bar{H} \sum_{j\in\nu_i} \frac{|e_{ij}|}{|c_i|} \sum_{r=1}^{R} \xi_r \begin{pmatrix} 0 \\ b_{ij,r}n_{ij} \\ 0 \end{pmatrix}, \tag{7}$$

where we have used the properties: $\sum_{j\in\nu_i} |e_{ij}|n_{ij} = 0$ and $\sum_{r=1}^{R} \xi_r = 1$.
On the other hand, taking advantage that the numerical integration is exact for polynomial functions of degree d we have

$$
\begin{aligned}
\mathbb{S}_i^n &= -g\bar{H}\frac{1}{|c_i|}\int_{c_i}\begin{pmatrix}0\\ \nabla\widehat{b}_i(\mathbf{x})\ d\mathbf{x}\\ 0\end{pmatrix}\\
&= -g\bar{H}\sum_{j\in\nu_i}\frac{1}{|c_i|}\int_{e_{ij}}\begin{pmatrix}0\\ \widehat{b}_i(\mathbf{x})n_{ij}ds\\ 0\end{pmatrix}\\
&= -g\bar{H}\sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r\begin{pmatrix}0\\ b_{ij,r}n_{ij}\\ 0\end{pmatrix}. \qquad (8)
\end{aligned}
$$

To conclude, we observe that expressions (7) and (8) exactly compensate the flux contribution so that equation (6) yields $V_i^{n+1} = V_i^n$. Hence the C-property is satisfied. $\square$

Theorem 2.3 *Let*

$$\varepsilon_i^n = \sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r\varepsilon_{ij,r}^n \qquad (9)$$

be the non-conservative contribution. Assume that the hypotheses of theorem 2.2 are satisfied. If the bathymetry $b^\star$ is convex and function $b \in C^{d+1}(\Omega)$ then

$$|\varepsilon_i^n| \leq gH_{\max}^n\frac{|\partial c_i|}{|c_i|}CM\eta^{d+1} \qquad (10)$$

with $H_{\max}^n = \max_{i\in\mathcal{E}_{el},\, j\in\nu_i,\, r=1,\cdots,R}|H_{ij,r}^n|$, $\eta = \max_{i\in\mathcal{E}_{el}}\emptyset(c_i)$ the maximum diameter of the cells, $M = \max_{\mathbf{x}\in\Omega}\|\nabla^{d+1}b(\mathbf{x})\|$ and C a positive constant independent of the mesh, b and H.

Proof. Since $b^\star$ is convex, then there exists $\theta_{ij,r} \in [0,1]$ such that $b_{ij,r}^\star = (1-\theta_{ij,r})b_{ij,r} + \theta_{ij,r}b_{ji,r}$. Hence $b_{ij,r}^\star - b_{ij,r} = \theta_{ij,r}(b_{ji,r} - b_{ij,r})$. From relations (5) and (9) we have

$$
\begin{aligned}
\varepsilon_i^n &= -\sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r g H_{ij,r}^n\theta_{ij,r}(b_{ji,r}-b_{ij,r})\begin{pmatrix}0\\ n_{ij}\\ 0\end{pmatrix}\\
&= -g\sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r H_{ij,r}^n\theta_{ij,r}(\widehat{b}_j(q_{ij,r})-\widehat{b}_i(q_{ij,r}))\begin{pmatrix}0\\ n_{ij}\\ 0\end{pmatrix}\\
&= -g\sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r H_{ij,r}^n\theta_{ij,r}\Big([\widehat{b}_j(q_{ij,r})-b(q_{ij,r})]+[b(q_{ij,r})-\widehat{b}_i(q_{ij,r})]\Big)\begin{pmatrix}0\\ n_{ij}\\ 0\end{pmatrix}.
\end{aligned}
$$

Since the reconstruction $\widehat{b}_i$ is exact for polynomial functions belonging to $\mathbb{P}^d$, we deduce that if $b\in C^{d+1}(\Omega)$ there exists a positive constant C' such that

$$
\begin{aligned}
|\widehat{b}_j(q_{ij,r})-b(q_{ij,r})| &\le C'M|m_j-q_{ij,r}|^{d+1}\le C'M\eta^{d+1},\\
|\widehat{b}_i(q_{ij,r})-b(q_{ij,r})| &\le C'M|m_i-q_{ij,r}|^{d+1}\le C'M\eta^{d+1},
\end{aligned}
$$

and we obtain the estimate

$$
|\varepsilon_i^n|\le gH_{\max}^n\sum_{j\in\nu_i}\frac{|e_{ij}|}{|c_i|}\sum_{r=1}^{R}\xi_r\theta_{ij,r}2C'M\eta^{d+1}\le\frac{|\partial c_i|}{|c_i|}gH_{\max}^nCM\eta^{d+1},
$$

where the positive constant C is independent of H, b and η. $\square$

Remark 2.4 *Let us assume some regularity of the mesh, namely that there exists a constant $c>0$ such that $\frac{|\partial c_i|}{|c_i|}<\frac{c}{\eta}$. We deduce from (10) the new estimate*

$$
|\varepsilon_i^n|\le CMgH_{\max}^n\eta^d.
$$

This last estimate shows that the correction term is a very small perturbation when dealing with smooth approximations. In concrete simulations, we have an error of order $O(\eta^{d+1})$, which suggests that a better estimate might be obtained.

2.3 Conservative flux and the physical bathymetry representative

We now turn to the numerical flux for the conservative term to determine the bathymetry $b^\star$ for different popular fluxes (see the textbook [50]). To deal with, let (n,τ) be the direct orthogonal basis associated to a generic edge e.

Matrix

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & n_x & n_y & 0 \\ 0 & \tau_x & \tau_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

corresponds to the transformation from the canonical basis to the new basis. Setting $U_n = U.n$, $U_\tau = U.\tau$ for the normal and the tangential velocity and $U' = (U_n, U_\tau)^T$ (resp. with Q_n, Q_τ and Q'), the physical flux writes

$$F(V)n_x + G(V)n_y = T^{-1}F(TV).$$

Hence, we shall only provide a flux approximation $\mathbb{F}(V'_L, V'_R)$ (with only two entries) of F such that

$$\mathbb{F}(V_L, V_R; n) = T^{-1}\mathbb{F}(V'_L, V'_R),$$

where $V'_L = TV_L = (H_L, Q_{n,L}, Q_{\tau,L}, b_L)^T$ and $V'_R = TV_R = (H_R, Q_{n,R}, Q_{\tau,R}, b_R)^T$.

2.3.1 The Rusanov flux

The Rusanov flux for the shallow-water problem writes

$$\mathbb{F}_{Rus}(V'_L, V'_R) = \frac{F(V'_L) + F(V'_R)}{2} - \frac{\lambda}{2}\begin{pmatrix} H_R - H_L \\ Q'_R - Q'_L \\ 0 \end{pmatrix},$$

where the viscosity contribution is controlled by coefficient $\lambda = \max(\sqrt{gh_L} + |U_{n,L}|, \sqrt{gh_R} + |U_{n,R}|)$.

Let $V_L = (\bar{H}, 0, 0, b_L)^T$ and $V_R = (\bar{H}, 0, 0, b_R)^T$ be two states with null velocity such that $\bar{H} > \max(b_L, b_R)$. Computing the numerical flux for these specific states leads to

$$\begin{aligned} \mathbb{F}_{Rus}(V_L, V_R) &= \frac{g}{4}\Big((0, \bar{H}(\bar{H} - 2b_L), 0, 0)^T + (0, \bar{H}(\bar{H} - 2b_R), 0, 0)^T\Big) \\ &= \frac{g}{2}(0, \bar{H}(\bar{H} - b_L - b_R), 0, 0)^T \\ &= F\Big(\Big(\bar{H}, 0, 0, \frac{b_L + b_R}{2}\Big)^T\Big), \end{aligned}$$

Setting now $b^\star(V_L, V_R) = \frac{1}{2}(b_L + b_R)$, we deduce that $b^\star$ is a physical bathymetry representative for the Rusanov flux. Now, extending the definition of $b^\star$ for any

left and right admissible states, function $b^\star$ is clearly continuous and convex with $\theta = \frac{1}{2}$.

2.3.2 The Harten, Lax, van Leer (HLL) flux

The HLL flux is based on an approximation of the solution by three states split by two characteristic eigenvalues a_L and a_R with $a_L < a_R$ [51]. There exist several expressions (see [50] for instance) to set the values a_L, a_R, and we have adopted in the numerical experiences the simple formula

$$a_L = \min(U_{n,L} - \sqrt{gh_L}, U_{n,R} - \sqrt{gh_R}), \quad a_R = \max(U_{n,L} + \sqrt{gh_L}, U_{n,R} + \sqrt{gh_R}).$$

The numerical flux and function $b^\star$ are written following three possible situations:

- if $a_L \geq 0$, $\mathbb{F}_{HLL}(V'_L, V'_R) = F(V'_L)$, $b^\star = b_L$;
- if $a_R \leq 0$, $\mathbb{F}_{HLL}(V'_L, V'_R) = F(V'_R)$, $b^\star = b_R$;
- otherwise,

$$\mathbb{F}_{HLL}(V'_L, V'_R) = \frac{a_R F(V'_L) - a_L F(V'_R)}{a_R - a_L} + \frac{a_R a_L}{a_R - a_L} \begin{pmatrix} H_R - H_L \\ Q_{n,R} - Q_{n,L} \\ Q_{\tau,R} - Q_{\tau,L} \\ 0 \end{pmatrix} \tag{11}$$

and $b^\star = \dfrac{a_R b_L - a_L b_R}{a_R - a_L}$.

From the definition, one checks that the bathymetry $b^\star(V_L, V_R)$ is convex. Now let us consider the states $V'_L = (\bar{H}, 0, 0, b_L)^T$ and $V'_R = (\bar{H}, 0, 0, b_R)^T$ with $\bar{H} > \max(b_L, b_R)$. Since the velocity is null, the flux is given by relation (11) with $a_L < 0$, $a_R > 0$ and we have

$$\mathbb{F}_{HLL}(V'_L, V'_R) = \frac{g}{2} \left(0, \bar{H}\Big(\bar{H} - 2\frac{a_R b_L - a_L b_R}{a_R - a_L}\Big), 0, 0\right)^T = F\Big(\Big(\bar{H}, 0, 0, \frac{a_R b_L - a_L b_R}{a_R - a_L}\Big)^T\Big).$$

We conclude that $b^\star$ enjoys the physical bathymetry representative property for the HLL flux.

2.3.3 The HLLC flux

We consider a modified version of the classic HLLC solver [52], initially presented in [53], which derives from the simple HLL flux adding one more contact wave to take into account of the tangential velocity. For such, an additional

simple intermediate wave, with velocity $a^* \in [a_L, a_R]$ in the approximated Riemann problem solution, is considered and the numerical flux vector takes the following expression:

$$F_{HLLC}(V_L', V_R') = \begin{cases} F(V_L') \text{ if } & a_L \geq 0, \\ F_L^* \quad \text{if } & a_L < 0 \leq a^*, \\ F_R^* \quad \text{if } & a^* < 0 < a_R, \\ F(V_R') \text{ if } & 0 \geq a_R, \end{cases}$$

where F_L^* and F_R^* are the numerical fluxes in the left and right parts of the middle region of the Riemann solution. Introducing the notation

$$\mathbb{F}_{HLL} = ((\mathbb{F}_{HLL})^1, (\mathbb{F}_{HLL})^2, (\mathbb{F}_{HLL})^3, 0)^T,$$

then the HLLC expression writes

$$F_k^* = ((\mathbb{F}_{HLL})^1, (\mathbb{F}_{HLL})^2, U_{\tau,k}(\mathbb{F}_{HLL})^1, 0)^T, \quad k = R, L,$$

where the third component $(\mathbb{F}_{HLL})^3$ of the initial HLL flux is substituted with $U_{\tau,k}(\mathbb{F}_{HLL})^1$. Notice that the bathymetry $b^\star$ is the one corresponding to the HLL solver given in the previous paragraph.

To provide an explicit expression for the flux, one has to first define the approximated eigenvalues a_L and a_R taking

$$a_L = \min(U_{n,R} - \sqrt{gh_R}, u^* - \sqrt{gh^*}), \quad a_R = \max(U_{n,L} + \sqrt{gh_L}, u^* + \sqrt{gh^*}),$$

where

$$\begin{aligned} u^* &= \frac{U_{n,L} + U_{n,R}}{2} + \sqrt{gh_L} - \sqrt{gh_R}, \\ h^* &= \frac{1}{g}\left[\frac{U_{n,L} - U_{n,R}}{4} + \frac{\sqrt{gh_L} + \sqrt{gh_R}}{2}\right]^2, \end{aligned}$$

while the middle wave speed a^* derives from [53], namely

$$a^* = \frac{a_L h_R (U_{n,R} - a_R) - a_R h_L (U_{n,L} - a_L)}{h_R (U_{n,R} - a_R) - h_L (U_{n,L} - a_L)}.$$

The function $b^\star$ is the same as the one defined for the HLL flux. However, since the definition of a_L and a_R is different from the HLL case, we have to check the $b^\star$ properties, namely, continuity, convexity and physical bathymetry representative. Let V_L and V_R be two admissible states. If $a_L \geq 0$ (resp. $0 \geq a_R$) we have $b^\star = b_L$ (resp. $b^\star = b_R$), hence the convexity property holds. For the

two other cases, we have $a_L < 0 < a_R$ and therefore $b^\star = \dfrac{a_R b_L - a_L b_R}{a_R - a_L}$ is also a convex combination. Moreover, function $b^\star(V_L, V_R)$ is continuous since h^* and u^* are continuous with respect to the admissible states. Now let us consider the lake at rest states $V'_L = (\bar{H}, 0, 0, b_L)^T$ and $V'_R = (\bar{H}, 0, 0, b_R)^T$ with $\bar{H} > \max(b_L, b_R)$. Since the velocity is null, we have

$$a_L = \min(-\sqrt{gh_R}, u^* - \sqrt{gh^*}) < 0, \quad a_R = \max(\sqrt{gh_L}, u^* + \sqrt{gh^*}) > 0,$$

and from the previous section we deduce

$$\begin{aligned}
\mathbb{F}_{HLLC}(V'_L, V'_R) &= ((\mathbb{F}_{HLL})^1, (\mathbb{F}_{HLL})^2, 0, 0)^T \\
&= F^*_L = F^*_R \\
&= \mathbb{F}_{HLL}(V'_L, V'_R) = F\Big(\Big(\bar{H}, 0, 0, \frac{a_R b_L - a_L b_R}{a_R - a_L}\Big)^T\Big).
\end{aligned}$$

We conclude that $b^\star$ is a physical bathymetry representative for the HLLC flux and also that continuity with respect to the admissible states holds.

3 The MOOD method

High-order approximations generate spurious oscillations leading to non-physical solutions when dealing with rough functions. To overcome this problem, several strategies have been developed, such as the MUSCL method or the ENO/WENO technique, where the accuracy is locally reduced in the vicinity of discontinuities to provide robustness. The MOOD method differs from other high-order methods since the limitation strategy is performed *a posteriori*, *i*.e. after the solution update procedure. The main idea of the MOOD method is to determine, for each cell, the optimal degree that one can employ in the polynomial reconstruction that provides both the best accuracy and satisfies some stability conditions. In the following, we summarise the main ingredients of the method and refer to [2,18,20,21] for others details. The point is to compute an admissible and accurate solution V^{n+1} from V^n in a sense we shall present is the sequel.
To this end, we introduce the Cell Polynomial Degree d_i (in short CellPD) as the degree of the polynomial function associated to cell c_i, while d_{ij} stands for the Edge Polynomial Degree (in short EdgePD) associated to edge e_{ij}. We deduce the EdgePD map from the CellPD map using the simple rule $\mathsf{d}_{ij} = \min(\mathsf{d}_i, \mathsf{d}_j)$ and compute the approximations $\phi_{ij,r}$, $\phi_{ji,r}$, $r = 1, \ldots, R$ at point $q_{ij,r}$ on both sides of the edge using the polynomial reconstructions $\widehat{\phi}_i$ and $\widehat{\phi}_j$ of degree d_{ij}. The main problem is the determination of the CellPD map such that the solution V^{n+1} is admissible.

Remark 3.1 *Notice that we need several polynomial reconstructions (one per degree) on cell c_i since d_{ij} may be different from one edge to another.*

Two independent mechanisms are involved in the MOOD method: the detection procedure and the limitation procedure. The detection stage is based on the notion of $\mathcal{A}$-eligible set, where we check each cell to determine whether the numerical solution is admissible or not. The limitation procedure mainly consists in reducing the polynomial degree where it is necessary to avoid the appearance of numerical instabilities.

3.1 $\mathcal{A}$-eligible set

The detection procedure is the core of the method. We establish criteria to determine whether the approximation of the mean values on cells correspond to an admissible solution or not. We here rephrase the abstract framework proposed in [2,18] and denote by $\mathcal{A}$ the set of detection criteria (for example the positivity of the water height in the shallow-water context) that the numerical approximation has to respect on each cell. We say that a candidate solution is $\mathcal{A}$-eligible if it fulfils all the criteria of $\mathcal{A}$.
If the candidate solution is not $\mathcal{A}$-eligible on cell c_i, then we reduce the polynomial degree of the respective cell. However, the solution may not be $\mathcal{A}$-eligible regardless of the set $\mathcal{A}$ even if the polynomial degree is zero for the cell. Consequently, we shall consider the numerical solution *acceptable* on the cell if either it is $\mathcal{A}$-eligible or is a first-order approximation (*i.e.* the CellPD has been decremented to zero). Several techniques have been developed in [2,18] to reduce the computational cost and avoid re-evaluation of all the fluxes on the whole domain. On the other hand, we extend the MOOD algorithm initially designed for a one-time step Euler scheme to the TVD-RK3 scheme by applying the MOOD procedure to each substep of the TVD-RK3 procedure. Therefore, for sake of simplicity, in the following we shall present the MOOD procedure for just one Euler step, bearing in mind that the TVD-RK3 scheme is a succession of Euler steps.

3.2 Detection criteria

We define specific detection criteria for the shallow-water problem which mainly derive from the ones used for the Euler system. For that purpose it is useful to consider, for each cell c_i, a mesh parameter δ_i that depends only on the cell dimensions and on the geometry of the domain. More specifically, we define δ_i as the ratio between the maximal length of the interfaces of cell c_i and a characteristic length of the domain (in general, its maximal length along the x and the y direction). This parameter will be used in most of the

detectors described below for relaxation purposes following closely the algorithm proposed by [18], which was successfully tested for the Euler system and convection equation.

3.2.1 PAD

A first important criterion we shall require is the physical admissibility of the solution for ensuring the physical meaningfulness of the primitive variables. We introduce the Physically Admissible Detector (PAD in short) which considers that the candidate solution on a cell c_i is not valid if $h_i^\star = H_i^\star - b_i$ is negative. We underline the important property that a high-order scheme (whichever the degree of the polynomial reconstruction) equipped with the PAD and a first-order scheme preserving the water height positivity under an appropriate CFL condition is automatically positivity-preserving. This property straightforwardly follows from the *a posteriori* nature of the MOOD method and has been proved in [2] in the Euler context, but clearly holds for the shallow-water system as well.

3.2.2 DMP

The PAD allows the numerical solution computation but does not prevent spurious oscillations to appear in the vicinity of discontinuities. Thus, other criteria will be added to the $\mathcal{A}$-eligible set for that purpose, and in particular to control the local numerical smoothness of the numerical solution. We adapt the Discrete Maximum Principle (DMP) and the u2 criteria initially proposed for the Euler system by using the total water height H instead of the density as detection variable. Indeed, high-order approximations close to a discontinuity develop oscillations (Gibbs phenomenon), in particular local extrema will appear. We introduce the Discrete Maximum Principle Detector which considers that the candidate solution $V_i^\star$ on a cell c_i may be problematic if the property

$$\min\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right) \leq H_i^\star \leq \max\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right)$$

is not satisfied, where the index set $\overline{\nu}(i)$ corresponds to the cells which share a common edge or a common vertex with c_i (the first layer of cells around c_i). In concrete simulations, the above condition is relaxed to reduce overdetection of problematic cells due to floating round procedure, in particular for local constant functions. In practice, we implement the following DMP criterion

$$\min\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right) - \varepsilon_D \leq H_i^\star \leq \max\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right) + \varepsilon_D,$$

where ε_D is a positive constant usually set to 10^{-14}. Additionally, the DMP criterion is not checked for cell c_i if

$$\max\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right) - \min\left(H_i^n; H_{j\in\overline{\nu}(i)}^n\right) \leq \varepsilon_F.$$

In the numerical simulations we take $\varepsilon_F = \delta_i^3$. This criterion intends to identify flat areas where extrema detection does not make sense.

3.2.3 ED

The Extrema Detector (ED) is a reformulation of the DMP criterion where we only employ the candidate solution. We consider that the candidate solution $V_i^\star$ on a cell c_i may be problematic if the property

$$\min\left(H_{j\in\overline{\nu}(i)}^\star\right) \leq H_i^\star \leq \max\left(H_{j\in\overline{\nu}(i)}^\star\right)$$

is not satisfied. This property detects if cell c_i corresponds to a local extremum, hence a potential oscillation. As in the DMP case, the concrete implementation involves a relaxation procedure to avoid the overdetection problem, the relaxed version being

$$\min\left(H_{j\in\overline{\nu}(i)}^\star\right) - \varepsilon_E \leq H_i^\star \leq \max\left(H_{j\in\overline{\nu}(i)}^\star\right) + \varepsilon_E,$$

where we take $\varepsilon_E = \delta_i^3$ in our applications.

3.2.4 u2

The DMP/ED criterion successfully eliminates the non-physical oscillations but, unfortunately, it can also depreciate the accuracy since smooth extrema are misinterpreted as discontinuities. The purpose of the u2 criterion is, when the DMP or ED criterion is activated, to separate the discontinuous situations from the regular extrema.
For a given cell c_i, we consider the quadratic polynomial reconstruction $\widehat{V}_i^n(\mathbf{x}, 2)$ and $\widehat{V}_j^n(\mathbf{x}, 2)$, $j \in \overline{\nu}(i)$. Noticing that the second derivatives are constant, we define for a given function ϕ,

$$\mathcal{X}_i^{min}(\phi) = \min_{j\in\overline{\nu}(i)}\left(\partial_{xx}\widehat{\phi}_i^n, \partial_{xx}\widehat{\phi}_j^n\right), \quad \mathcal{X}_i^{max}(\phi) = \max_{j\in\overline{\nu}(i)}\left(\partial_{xx}\widehat{\phi}_i^n, \partial_{xx}\widehat{\phi}_j^n\right),$$

$$\mathcal{Y}_i^{min}(\phi) = \min_{j\in\overline{\nu}(i)}\left(\partial_{yy}\widehat{\phi}_i^n, \partial_{yy}\widehat{\phi}_j^n\right), \quad \mathcal{Y}_i^{max}(\phi) = \max_{j\in\overline{\nu}(i)}\left(\partial_{yy}\widehat{\phi}_i^n, \partial_{yy}\widehat{\phi}_j^n\right).$$

From the curvatures we introduce three criteria where we skip the function dependency for the sake of simplicity.

3.2.4.1 Oscillation criterion We consider that the numerical solution around cell c_i is not oscillating if

$$\mathcal{X}_i^{min}\mathcal{X}_i^{max} \geq 0 \text{ and } \mathcal{Y}_i^{min}\mathcal{Y}_i^{max} \geq 0.$$

The main motivation of the definition is that an oscillation is characterised by a change of the curvature sign in the vicinity of a discontinuity. Hence, the criterion detects the transition of the numerical curvature sign. In practice, we use a relaxed counterpart to avoid overdetection of problematic cells and we concretely implement the condition

$$\mathcal{X}_i^{min}\mathcal{X}_i^{max} \geq -\varepsilon_O \text{ and } \mathcal{Y}_i^{min}\mathcal{Y}_i^{max} \geq -\varepsilon_O, \tag{12}$$

where ε_O is a positive value that is set equal to δ_i in the simulations.

3.2.4.2 Plateau criterion The numerical solution is a plateau solution on cell c_i if

$$\max(|\mathcal{X}_i^{min}|, |\mathcal{X}_i^{max}|, |\mathcal{Y}_i^{min}|, |\mathcal{Y}_i^{max}|) \leq \varepsilon_P, \tag{13}$$

where ε_P is the plateau threshold parameter that we take equal to δ_i. Indeed, for a smooth constant solution (local plateau solution) small numerical artefacts due to the floating point truncation arise, leading to very small errors which are misinterpreted as oscillation by criterion (12).

3.2.4.3 Smooth curvature The numerical solution is smooth on cell c_i if

$$\frac{\min(|\mathcal{X}_i^{min}|, \mathcal{X}_i^{max}|)}{\max(|\mathcal{X}_i^{min}|, |\mathcal{X}_i^{max}|)} \geq 1 - \varepsilon_S, \text{ and } \frac{\min(|\mathcal{Y}_i^{min}|, \mathcal{Y}_i^{max}|)}{\max(|\mathcal{Y}_i^{min}|, |\mathcal{Y}_i^{max}|)} \geq 1 - \varepsilon_S, \tag{14}$$

where ε_S is the smoothness parameter (in practice we set $\varepsilon_S = \frac{1}{2}$).
The u2 criterion is a combination of the three previous criteria when the DMP or ED criterion is activated, which indicates the existence of a potential oscillation. We state that

- the cell is not eligible if (12) and (13) do not hold;
- the cell is not eligible if (14) does not hold.

Several choices for the u2 detection variable are available. In [2,18], the density has been chosen since it is the most representative for shocks and contact discontinuities. For that reason, the free surface variable H is employed as the detection variable to determine the CellPD map for the approximations. In [19], the authors introduce an entropy function as the detection variable and define a specific $\mathcal{A}$-eligible set with respect to the entropy.

3.2.5 u2$^{\nu}$

In order to improve the accuracy, we propose a relaxation of the smoothness detector u2 still maintaining the robustness of the scheme. The main idea is driven by the fact that both the DMP/ED and the u2 detectors involve a set of constrains that uses, for a given cell c_i, the stencil formed by the whole layer around c_i, *i.e.* the cells that share at least one vertex with cell c_i and is characterised by the index set $\overline{\nu}(i)$. To relax the constraint, we consider the stencil formed by the cells only sharing an edge with cell c_i, characterised by the index set ν_i. We substitute the index set $\overline{\nu}(i)$ in the original u2 detector with the less restrictive set ν_i, hence the name u2$^{\nu}$.

3.3 The MOOD loop

We sum-up the MOOD algorithm for the explicit discretization in time which consists in the following stages:

0. Initialise $\mathsf{d}_i = \mathsf{d}_{max}, \quad \forall i \in \mathcal{E}_{el}$.
1. Compute EdgePD d_{ij} and evaluate $V_{ij,r}$, $V_{ji,r}$ at each Gauss point using the polynomial reconstructions of degree d_{ij}.
2. Compute candidate solution mean values.
3. Apply the detection process and reduce the CellPD for the cells which are not *acceptable*.
4. Stop if the solution is acceptable everywhere, otherwise go to back to stage 1.

The loop goes on until an admissible candidate solution is reached and we set $V^{n+1} = V^{\star}$. We refer to [2,18] where the authors established the conditions such that the MOOD loop always converges.

Remark 3.2 *The CellPD map for the b function is treated in a specific manner since b is known and therefore the MOOD loop described above does not apply to b. Two cases may occur, either b is known analytically or only (estimates of) mean values on cells are known. In the first case, which happens typically when running test cases, we use the exact values in the reconstruction procedure to provide an accurate local representation of b, the degree of the polynomial reconstruction being set for each cell according to smoothness of b. In the second case, only topographical samples are available providing just an estimate of the mean value of b in each cell. We use these data to compute the reconstruction polynomials, but to provide a non-oscillating representation of b the MOOD procedure (PAD+ED+u2) is applied just once as a pre-processing procedure. In both cases, the b reconstruction obtained is used to compute the fluxes as well as non-conservative the source term (3).*

4 Numerical tests

We present several numerical tests to validate the properties of the schemes. The time step Δt is controlled by a CFL coefficient with respect to the first-order time step that we set equal to 0.65 of the maximum admissible time step. For the convergence studies on smooth solutions we use the time step $\Delta t = \Delta x^{r/3}$ to achieve a global r^{th}-order of accuracy (see Section 2.1) and compute the L^1- and L^∞-errors for a bounded L^1 function ϕ by

$$L^1\text{-error: } \sum_{i\in\mathcal{E}_{el}} |\phi_i^N - \phi_i^{ex}||c_i| \quad \text{and} \quad L^\infty\text{-error: } \max_{i\in\mathcal{E}_{el}} |\phi_i^N - \phi_i^{ex}|,$$

where $(\phi_i^{ex})_{i\in\mathcal{E}_{el}}$ and $(\phi_i^N)_{i\in\mathcal{E}_{el}}$ are respectively the exact and the approximated cell mean values at final time $t^N = t_{\text{final}}$.

Finally, we take the gravitational acceleration constant $g = 9.81\, ms^{-2}$ in all the numerical tests.

The set of figures presented in this section as well as the Delaunay meshes used were obtained using the Gmsh package [54].

4.1 Lake at rest

Simulations of the lake at rest have been performed using the three numerical fluxes presented in Section 2.3. This sanity test assesses if the numerical scheme respects the C-property. The domain is the square $[-1,1]\times[-1,1]$ where we assume that the fluid is initially at rest, *i.e.* the velocity is zero and the total height is constant ($H = 1.0$). We consider two different scenarios whether the bathymetry function b is smooth or not, and experiment different polynomial degrees for the reconstruction ($\mathbb{P}_2$, $\mathbb{P}_3$ and $\mathbb{P}_5$). The bathymetry reconstruction involves a $\mathbb{P}_5$ reconstruction for the smooth case, while a $\mathbb{P}_0$ reconstruction is used for the piecewise constant case. In all the simulations, reflection conditions are prescribed at the boundary and, given the smoothness and stationarity of the solution, only the Physical Admissible Detector (PAD) is employed within the MOOD algorithm.

4.1.1 Lake at rest with smooth bathymetry

The domain is partitioned in 10050 triangles (Delaunay mesh) and the smooth bathymetry function is given by $b(x,y) = 0.5e^{-3(4x^2+8xy+9y^2)}$ (see Figure 2).

As a first test, we carry out the simulation until $t_{\text{final}} = 15\, s$ cancelling the non-conservative flux $\varepsilon_{ij,r}$ term and report the L^1- and L^∞-errors for H in

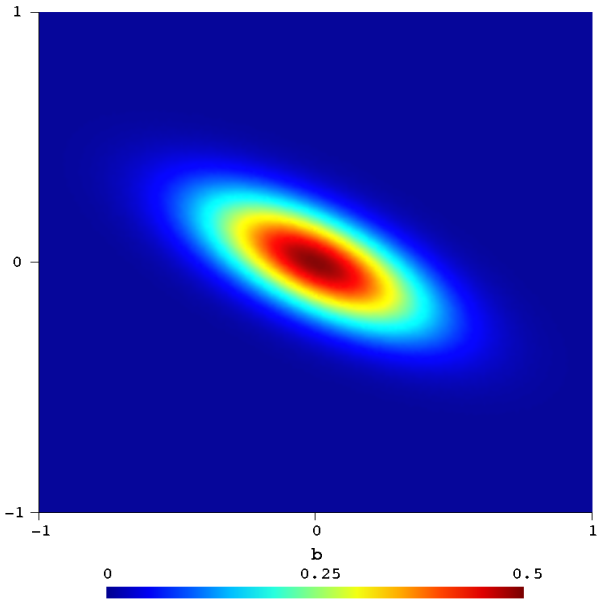

Fig. 2. Bathymetry function for the smooth case.

Table 1. Then, we perform the same simulations including the non-conservative contribution and present the corresponding errors in Table 2.

Table 1
Total height L^1- and L^∞-errors cancelling the non-conservative flux (smooth bathymetry).

Flux	$\mathbb{P}_2$		$\mathbb{P}_3$		$\mathbb{P}_5$	
scheme	err_1	err_∞	err_1	err_∞	err_1	err_∞
Rus	5.90e-07	3.82e-06	6.19e-07	4.67e-06	6.32e-07	3.43e-06
HLL	5.90e-07	3.82e-06	6.19e-07	4.67e-06	6.32e-07	3.43e-06
HLLC	2.99e-06	1.05e-04	8.81e-06	6.05e-04	5.29e-06	3.66e-04

Table 2
Total height L^1- and L^∞-errors using the non-conservative flux (smooth bathymetry).

Flux	$\mathbb{P}_2$		$\mathbb{P}_3$		$\mathbb{P}_5$	
scheme	err_1	err_∞	err_1	err_∞	err_1	err_∞
Rus	5.78e-17	7.77e-16	1.12e-16	6.66e-16	1.31e-16	1.55e-15
HLL	4.20e-17	6.66e-16	1.09e-16	6.66e-16	1.31e-16	1.67e-15
HLLC	1.89e-15	3.62e-14	5.99e-15	1.00e-12	3.95e-15	1.46e-13

Clearly, the steady-state configuration is not preserved when the non-conservative flux is cancelled, whereas we maintain the exact solution with the flux correction. In the last case, the results are mainly flux independent (we have the same truncation error using Rusanov, HLL and HLLC fluxes). Moreover,

the degree of the polynomial reconstruction does not affect the error, which proves that the scheme remains well-balanced for any degree. Notice that this last point is not so straightforward since functions b and h are not constant in space.

4.1.2 Lake at rest with non-smooth bathymetry

To check the C-property with a discontinuous bottom, we consider the bathymetry presented in Figure 3 where the height of the bump is 0.5 and the the free surface is 1.0. We mesh the domain with 10352 triangles and carry out the simulation until a final time $t_{\text{final}} = 15\,s$, the time step being controlled by the CFL condition. We reproduce in Table 3 the L^1- and L^∞-errors for the different polynomial reconstructions and fluxes. The simulations confirm the efficiency of the method to preserve this specific and important steady-state.

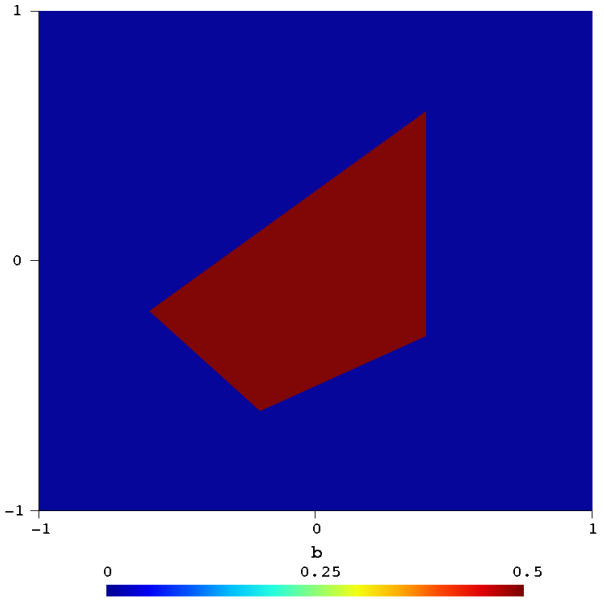

Fig. 3. Bathymetry function for the discontinuous case.

Table 3
Total height L^1- and L^∞-errors using the non-conservative flux (non-smooth bathymetry.)

Flux	$\mathbb{P}_2$		$\mathbb{P}_3$		$\mathbb{P}_5$	
scheme	err_1	err_∞	err_1	err_∞	err_1	err_∞
Rus	6.68e-17	8.88e-16	1.07e-16	6.66e-16	1.24e-16	8.88e-16
HLL	3.01e-17	6.66e-16	1.10e-16	7.77e-16	1.26e-16	1.11e-15
HLLC	6.66e-15	2.09e-12	2.23e-09	7.64e-07	1.43e-10	5.83e-08

We report that the Rusanov and HLL fluxes provide better results than the HLLC flux providing less numerical errors. It worthwhile mentioning that, as

in the smooth bathymetry case, we do not preserve the steady-state when the non-conservative flux is cancelled and the discrepancy is considerably worse in the discontinuous case.

4.2 1D steady-state solutions

We now tackle more general situations of steady-state flow, namely, subcritical, supercritical and transcritical flows without shock to assess the capacity of the scheme to preserve such regimes with very high accuracy. Notice that the scheme has not been designed to exactly preserve these non-null velocity steady-states as in [14,27], but we intend to recover an high-order of accuracy, depending on the polynomial reconstruction used.

It is worthwhile to refer that a wide set of shallow-water analytic solutions for 1D and 2D configurations can be found in [55]. The configurations used in the numerical tests presented below follow closely some of the benchmarks presented in that compilation.

We consider a unidimensional flow for which the upstream boundary is located at $x = 0$ and the downstream one at $x = L$, such that the stationary solution writes

$$hu = q_0, \quad hv = 0, \quad \frac{q_0^2}{2gh^2(x)} + h(x) + b(x) = E_0,$$

where q_0 and E_0 are given constant values.

The nonlinear system is numerically solved at any location x using a Newton-Raphson algorithm once $b(x)$ and the boundary conditions are known, *i.e.* q_0 and E_0 are fixed. To carry out the simulations, we consider a 16 meters long channel of 4 meters width, $\Omega = [0, 16] \times [-2, 2]$, while the bathymetry is given by

$$b(x) = \begin{cases} 0.2 - 0.05(x-6)^2 \ , & 4 < x < 8, \\ 0 \qquad\qquad\qquad\quad , & \text{otherwise.} \end{cases}$$

Accordingly to the configuration of the domain with respect to b, we adopt the following mesh partition: $0 \leq x < 4$ (Z_1), $4 \leq x < 8$ (Z_2), and $8 \leq x \leq 16$ (Z_3) as depicted in Figure 4. The mesh derives from a division of an initial uniform mesh, in agreement with the one-dimensional character of the problem, and numerical simulations are performed with 800, 3200, 12800 and 51200 triangles cells for the convergence studies.

Due to the domain decomposition in three zones, the stencils are adapted to the partition in the sense that for any cell c_i, the associated stencil is composed

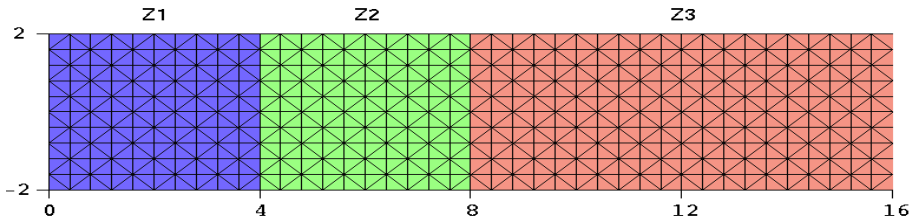

Fig. 4. Partition of the computational domain and mesh pattern.

of cells of the same zone. Notice that we only use a $\mathbb{P}_2$ reconstruction for b to recover the exact function. The water enters from the left side and flows out to the right side. We prescribe Dirichlet conditions at the Gauss points of the left and right border edges and use a ghost cell technique to enforce the values at the lower and upper boundaries of the domain. All the numerical simulations are performed using the HLLC numerical flux and we use the PAD detector to perform the MOOD method since the solution is essentially smooth. We use the steady-state solution as an initial condition and carry out the computation until the final time t_{final} and compare the numerical solution with the exact one in L^1- and L^∞-norms.

4.2.1 Subcritical case

A subcritical steady-state flow is obtained using the following boundary conditions: we set $hu = 1.53$ for the upstream side and prescribe $H = 2.0$ for the downstream side. We deduce E_0 and the maximal Froude number reaches 0.6. Numerical simulations are performed until the final time $t_{\text{final}} = 20\,s$ using three different polynomial degrees for the reconstruction: $\mathbb{P}_2$, $\mathbb{P}_3$ and $\mathbb{P}_5$.

In [49] the authors show the advantage to adapt the stencil with respect to the physical properties, *i.e.* the bathymetry in the present case. We perform a first simulation where for any cell c_i, the stencil used for functions H, hu and hv is constituted of neighbour cells that may belong to the different subdomains Z_1, Z_2 or Z_3. Table 4 gives the errors and convergence rate for the free surface H (similar results are obtained for both hu and hv). Clearly, the mixing of cells belonging to different zones provides a first-order of accuracy for the L^∞-norm and a second-order one for the L^1-norm due to the derivative discontinuities between the different areas. In Figure 5 we present the corresponding L^∞-error map for H for the 800 cells mesh and a $\mathbb{P}_2$ reconstruction. The largest errors are observed close to the interfaces between the three zones.

To overcome such a discrepancy, we adopt a new procedure where we propose a more adequate choice of the stencils to be used in the reconstruction of both b and the unknowns following [49]: build the stencil $S(c_i)$ with cells which belong to the same zone. We carry out a second simulation with the new reconstructions and the L^1- and L^∞-errors for the free surface are presented

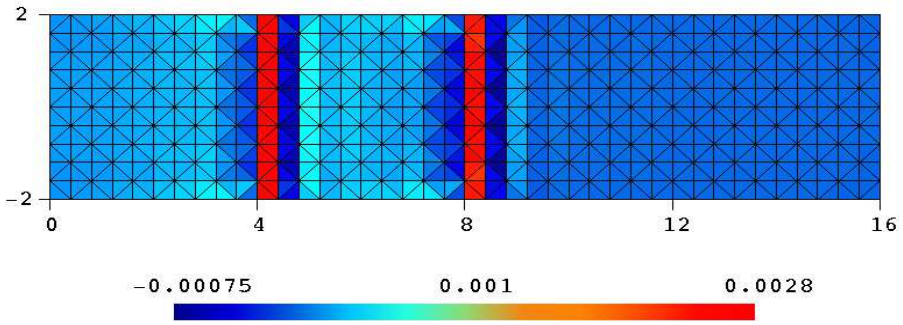

Fig. 5. Illustration of the total height L^∞-error field for the 800 cells mesh when a $\mathbb{P}_2$ polynomial reconstruction is used with stencils that are not topologically constrained.

Table 4
Total height L^1- and L^∞-errors and convergence order for the subcritical case when reconstruction stencils for H, hu and hv are not topologically constrained.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	2.73e-04	—	2.83e-03	—	6.42e-04	—	4.11e-03	—	3.67e-04	—	2.36e-03	—
3200	7.36e-05	1.9	1.32e-03	1.1	1.82e-04	1.8	1.90e-03	1.1	9.81e-05	1.9	1.10e-03	1.1
12800	1.93e-05	1.9	6.40e-04	1.0	4.86e-05	1.9	9.05e-04	1.1	2.59e-05	1.9	5.19e-04	1.1
51200	4.99e-60	2.0	3.16e-04	1.0	1.27e-05	1.9	4.40e-04	1.0	6.76e-06	1.9	2.58e-04	1.0

in Table 5 together with the order of convergence. Since the results for hu and hv are similar to the ones presented for H, the corresponding tables are not presented here. This last test confirms the adequate methodology we have proposed since we recover the optimal order in all the cases and no spurious oscillations are reported. On the other hand, the results demonstrate the capacity of the scheme to capture moving water regular situations up to the sixth-order of convergence even if the scheme was not specifically designed to exactly suit with all the steady-state situations, as also observed in [12,13,38].

4.2.2 Supercritical case

The framework is essentially the same as the one proposed for the subcritical case but now we set $q_0 = hu = 13.29$ and $H = 2.0$ at the upstream side leading to a Froude number larger than 1.28. Transmission conditions are used at the downstream boundary. Numerical simulations are carried out until the final time $t_{\text{final}} = 8\,s$ using the adapted stencil technique proposed in the previous section to provide the optimal order. Table 6 provides the corresponding L^1- and L^∞-errors and convergence order. We observe that the effective order is very close to the optimal one for both the total height and the impulsion. The

Table 5
Total height L^1- and L^∞-errors and convergence order for the subcritical case with stencils adapted with respect to the bathymetry.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	2.27e-05	—	1.40e-04	—	2.38e-06	—	1.82e-05	—	2.74e-07	—	2.00e-06	—
3200	2.81e-06	3.0	1.95e-05	2.8	1.02e-07	4.5	9.44e-07	4.3	2.35e-09	6.9	4.15e-08	5.6
12800	3.51e-07	3.0	2.63e-06	2.9	5.22e-09	4.3	6.30e-08	3.9	3.46e-11	6.1	4.72e-10	6.5
51200	4.38e-08	3.0	3.37e-07	3.0	2.91e-10	4.2	4.03e-09	4.0	5.46e-13	6.0	7.18e-12	6.0

steady-state situation is not exactly preserved since the schemes have not been designed for that purpose, but the errors are clearly controlled by the degree of the polynomial reconstructions.

Table 6
Total height L^1- and L^∞-errors and convergence order for the supercritical case.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	5.68e-05	—	4.11e-04	—	1.20e-06	—	1.10e-04	—	1.30e-06	—	1.27e-05	—
3200	6.71e-06	3.1	6.01e-05	2.8	6.08e-07	4.3	6.61e-06	4.1	3.31e-08	5.3	3.12e-07	5.3
12800	8.33e-07	3.0	7.59e-06	3.0	3.62e-08	4.1	4.33e-07	3.9	5.34e-10	6.0	5.50e-09	5.8
51200	1.04e-07	3.0	9.53e-07	3.0	2.24e-09	4.0	2.76e-08	4.0	8.27e-12	6.0	8.48e-11	6.0

4.2.3 Transcritical case without shock

To end the series of numerical simulations of unidimensional steady-state flows, the transcritical case was performed until $t_{\text{final}} = 8\,s$ with $q_0 = 1.53$ such that the flow is subcritical upstream (left side) with a Froude number around 0.48 at $x = 0$ and supercritical downstream (right side) with a Froude number of the order of 1.89 at $x = 16$. The transition occurs at the top of the bathymetry bump $x = 6$. Given the transcritical nature of the flow, we prescribe the Dirichlet condition $hu = q_0$ for the inflow side and transmission conditions for the outflow side. Errors and convergence rates for the free surface are provided in Table 7. Similar convergence results are obtained for the impulsion that we do not reproduce here.

The simulation results confirm the scheme ability to provide the optimal convergence order even for the complex transcritical case. We obtain, for instance, an effective sixth-order of convergence with the $\mathbb{P}_5$ reconstruction which enables to conclude once again that the scheme is well-balanced up to an error controlled by the degree of the polynomial reconstructions.

Table 7
Total height L^1- and L^∞-errors and convergence order for the transcritical case without shock.

Nb of Cells	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	1.15e-05	—	8.15e-05	—	2.53e-07	—	1.98e-06	—	2.82e-09	—	1.68e-08	—
3200	1.32e-06	3.1	1.24e-05	2.7	1.33e-08	4.2	1.15e-07	4.1	5.02e-11	5.8	2.26e-10	6.2
12800	1.64e-07	3.0	1.67e-06	2.9	7.19e-10	4.2	7.08e-09	4.0	7.99e-13	6.0	3.60e-12	6.0
25088	5.96e-08	3.0	6.22e-07	2.9	1.79e-10	4.1	1.82e-09	4.0	1.07e-13	6.0	4.82e-13	6.0

4.3 Steady-state vortex with varying bathymetry

We now turn to a real bidimensional test considering a steady-state vortex flow with varying bathymetry characterised by

$$H(x,y) = H_\infty - \frac{A^2}{4g} e^{2(1-r^2)}, \quad u(x,y) = A\hat{y}e^{(1-r^2)}, \quad v(x,y) = A\hat{x}e^{(1-r^2)},$$

with $\hat{x} = x - x_0$, $\hat{y} = y - y_0$, and $r^2 = \hat{x}^2 + \hat{y}^2$. We take $H_\infty = 1$, $A = 1$, and $x_0 = y_0 = 0$, while the bathymetry function is given by $b(r) = 0.2e^{(1-r^2)/2}$. Figures 6 and 7 depict the geometry of the vortex as well as the velocity field for the square domain $\Omega = [-3, 3] \times [-3, 3]$.

The simulations are carried out until the final time $t_{\text{final}} = 1\,s$ where we test the MOOD procedure performance using different detectors, namely the DMP against DMP+u2. For that purpose, we consider four Delaunay meshes of 800, 3194, 12742 and 50958 triangles and perform simulations with $\mathbb{P}_2$, $\mathbb{P}_3$ and $\mathbb{P}_5$ for the conservative variables, while the reconstruction for the b function is exact with the $\mathbb{P}_2$ polynomial. Initial conditions are prescribed using the steady-state solution and the Dirichlet boundary conditions imposed on the Gauss points of the boundary edges. The convergence results obtained for the total height are presented in Table 8 (DMP only) and Table 9 (DMP+u2).

Convergence rates clearly show the DMP drawback to catch the optimal order when dealing with smooth solutions since extrema are many times interpreted as discontinuities leading to a dramatic reduction of the local polynomial degree and providing an effective second-order scheme. This problem is overcome by introducing the u2 relaxation criterion since this specific detection procedure enables to identify the real smooth extrema detected by DMP from the oscillations deriving from the Gibbs phenomenon that DMP also detects. The CellPD map is optimal and we obtain a sixth-order convergence with the $\mathbb{P}_5$ reconstruction.

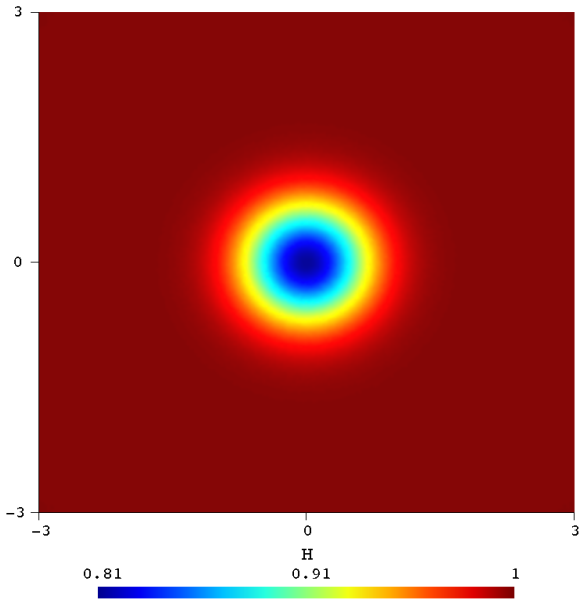

Fig. 6. Free surface for the static vortex.

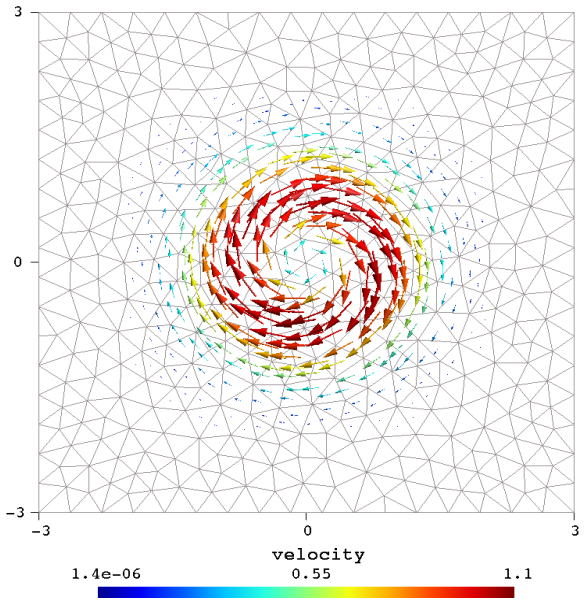

Fig. 7. Velocity field for the static vortex and the 800 triangles mesh.

4.4 Rising vortex with variable bathymetry

We aim to test the Extrema Detector (ED) in conjunction with the u2 procedure as an alternative to the DMP+u2 detector. Indeed, there exist some situations where the DMP is activated almost everywhere even if the solution does not present any extrema. For example, when the total height increases in time, the DMP is activated almost everywhere, whereas the ED only focus on real extrema. We also intend to compare the convergence results obtained

Table 8
Total height-L^1 and L^∞-errors and convergence order for the static vortex using DMP.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	7.73e-04	—	2.40e-02	—	4.43e-04	—	1.86e-02	—	4.76e-04	—	1.74e-02	—
3194	1.22e-04	2.7	7.12e-03	1.8	8.02e-05	2.5	6.86e-03	1.4	9.77e-05	2.3	6.49e-03	1.4
12742	1.98e-05	2.6	2.03e-03	1.8	1.52e-05	2.4	2.02e-03	1.8	1.84e-05	2.4	1.60e-03	2.0
50918	7.43e-06	1.4	5.39e-04	1.9	4.93e-06	1.6	4.74e-04	2.1	5.33e-06	1.8	5.58e-04	1.5

Table 9
Total height L^1- and L^∞-errors and convergence order for the static vortex using DMP+u2.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	4.85e-04	—	6.82e-03	—	8.69e-05	—	1.39e-03	—	3.89e-05	—	9.25e-04	—
3194	7.66e-05	2.7	9.99e-04	2.8	5.86e-06	3.9	8.16e-05	4.1	6.64e-07	5.9	1.41e-05	6.0
12742	1.02e-05	2.9	1.41e-04	2.8	3.67e-07	4.0	5.57e-06	3.9	1.05e-08	6.0	2.21e-07	6.0
50918	1.30e-06	3.0	1.86e-05	2.9	2.29e-08	4.0	4.26e-07	3.7	1.82e-10	5.9	3.55e-09	6.0

with standard Delaunay meshes and more regular triangular meshes where the triangles are generated with a frontal algorithm.

To this end, we slightly modify the previous example considering the following set of primitive variables

$$H(x,y,t) = e^{0.2t}(1-\frac{1}{4g}e^{2(1-x^2-y^2)}), \quad u(x,y,t) = A\hat{y}e^{(1-r^2)}, \quad v(x,y,t) = A\hat{x}e^{(1-r^2)},$$

where u and v remain constant in time, while the free surface globally rises in time, *i.e.* all the points rise with the same vertical velocity. Of course, the system is no longer conservative with respect to the mass. Therefore, an adequate source term S_m is required, manufactured from the exact solution, and the mass equation writes $\partial_t H + \nabla \cdot (hU) = S_m$. In the same way, the hydrostatic pressure increases since the free surface rises and an adequate source term for the impulsion equation is also added to the right-hand side to the non-conservative term. Numerical simulations are performed until the final time $t_{\text{final}} = 1\,s$ on domain $\Omega = [-3,3] \times [-3,3]$ with four frontal meshes of 856, 3372, 13422 and 53578 triangles. Initial conditions at $t = 0$ and Dirichlet conditions on the boundary are prescribed using the exact solution as in the steady-state case.

The first test aims to highlight the DMP detector drawbacks. Indeed, since H increases in time, the vortex rises as a block, the DMP detector reports an high percentage of problematic cells (around 70%) leading to an effective first-order scheme. Most of the cells correspond to new extrema with respect to the previous configuration and the DMP is inappropriately activated. Coupling with the u2 procedure to relax the DMP enables to increase the order but the computational effort turns out to be very important since we apply the u2 procedure for most of the cells to recover the efficient order. The u2 detector requires an important computational effort since one has to compute several polynomials of degree 2 in the vicinity of the selected cell and check a series of criteria to relax the CellPD to disable the degree reduction.

The second test consists in substituting the DMP by the ED as main detector. In this case, extrema detection is performed comparing cells of the same candidate solution and the global rising of the solution does not affect the detection. In practice, less than 0.05% of the cells are considered problematic by the ED detector. Combined with the u2 relaxation procedure, we obtain a very accurate and efficient MOOD procedure with a lower computational cost. The efficiency of the ED is very well-illustrated by the simulations using frontal meshes with 13422 and 53578 triangles, a reconstruction in space with polynomial functions of degree two and a RK3 scheme in time such that the space and time discretizations are both third-order. Note that the time step is exclusively constrained by the CFL condition. For the 13422 triangles mesh, we obtain a CPU running time of $5\,m\,14\,s$ with DMP+u2, whereas the same computation takes $2\,m\,58\,s$ with ED+u2. Notice that the time evaluation only concerns the main loop (*i.e.* discarding polynomial reconstruction and initial data reading). As for the finest mesh, we get $52\,m\,49\,s$ for DMP+u2 and $26\,m\,22\,s$ for ED+u2, which clearly demonstrates the computational time reduction when using ED as detector.

Table 10 provides the errors and convergence rates using the ED with Delaunay meshes, while Table 11 gives the results when adding the relaxation u2 detector. In the same way, Table 12 presents the errors and convergence rates with the ED technique using frontal meshes and we print in Table 13 similar results adding the u2 detector [2].

The MOOD strategy coupling ED with u2 provides the optimal order of convergence. Although the Delaunay meshes are less regular, we get numerical errors of the same order with respect to the ones obtained with the frontal meshes. The ED detector correctly identifies the potential problematic cells (only a few ones around the tip of the vortex) and does not make an overdetection of problematic cells as DMP does. In short, the ED is a good candidate as a detector to substitute the DMP procedure since it only focus on the real

[2] Symbol (*) is used when the convergence order does not make sense

Table 10
Total height L^1- and L^∞-errors and convergence order for the moving vortex using ED and Delaunay meshes.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	5.09e-04	—	1.03e-02	—	1.15e-04	—	3.73e-03	—	3.71e-05	—	8.61e-04	—
3194	1.01e-04	2.3	6.37e-03	0.7	8.28e-06	3.8	8.86e-04	2.1	4.04e-06	3.2	4.14e-04	1.1
12742	1.32e-05	2.9	1.94e-03	1.7	5.13e-06	0.7	8.21e-04	0.1	6.37e-07	2.7	2.87e-04	0.5
50918	1.83e-06	2.9	5.38e-04	1.9	2.51e-06	1.0	3.22e-04	1.4	1.56e-06	(*)	1.91e-04	0.6

Table 11
Total height L^1- and L^∞-errors and convergence order for the moving vortex using ED+u2 and Delaunay meshes.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
800	4.74e-04	—	7.07e-03	—	8.53e-05	—	1.37e-03	—	3.70e-05	—	8.58e-04	—
3194	7.27e-05	2.7	1.03e-03	2.8	5.88e-06	3.9	8.47e-05	4.0	6.29e-07	5.9	1.28e-05	6.1
12742	9.65e-06	2.9	1.45e-04	2.8	3.74e-07	4.0	5.98e-06	3.8	1.04e-08	5.9	2.46e-07	5.7
50918	1.23e-06	3.0	1.89e-05	2.9	2.33e-08	4.0	5.16e-07	3.5	1.78e-10	5.9	3.97e-09	6.0

Table 12
Total height L^1- and L^∞-errors and convergence order for the moving vortex using ED and frontal meshes.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
856	4.93e-04	—	1.17e-02	—	9.91e-05	—	1.81e-03	—	3.43e-05	—	4.70e-04	—
3372	6.92e-05	2.9	5.52e-03	1.5	1.15e-05	3.1	2.20e-03	(*)	3.18e-06	3.5	8.81e-04	(*)
13422	9.47e-05	2.9	1.59e-03	1.8	2.06e-06	2.5	1.15e-03	0.9	4.75e-07	2.7	3.82e-04	1.2
53578	1.10e-06	3.1	3.28e-04	2.3	1.91e-07	3.4	3.12e-04	1.9	1.03e-07	2.2	1.76e-04	1.1

extrema. Combined with the relaxation detector u2, we get a robust and more efficient up to sixth-order scheme.

4.5 Partial dam-break with a slope

We propose a more complex and realistic simulation test considering an extension of the classical 2D partial dam-break problem (see *e.g.* [11] and references therein). We assume that the reservoir (left part of the domain in Figure 8) is

Table 13
Total height L^1- and L^∞-errors and convergence order for the moving vortex using ED+u2 and frontal meshes.

Nb of	$\mathbb{P}_2$				$\mathbb{P}_3$				$\mathbb{P}_5$			
Cells	err_1		err_∞		err_1		err_∞		err_1		err_∞	
856	4.56e-04	—	7.14e-03	—	9.15e-05	—	1.08e-03	—	3.49e-05	—	5.60e-04	—
3372	6.39e-05	2.9	8.85e-04	3.0	6.39e-06	3.9	6.75e-05	4.0	6.35e-07	5.8	1.16e-05	5.7
13422	8.30e-06	3.0	1.09e-04	3.0	3.18e-07	4.3	3.60e-06	4.2	8.70e-09	6.2	1.58e-07	6.2
53578	1.04e-06	3.0	1.37e-05	3.0	1.99e-08	4.0	2.42e-07	3.9	1.42e-10	5.9	2.46e-09	6.0

higher than the river (right part of the domain), the two entities being relied by a ramp with constant slope. We study the outflow just after the dam rupture until a final simulation time $t_{\text{final}} = 7\,s$. Several characteristic structures will be analysed to evaluate the scheme accuracy and robustness, namely numerical diffusion of the discontinuity, the vortexes deepness as an accuracy assessment and the oscillations around shocks generated by the outflow as a robustness assessment. The domain we consider has been proposed in [11] and

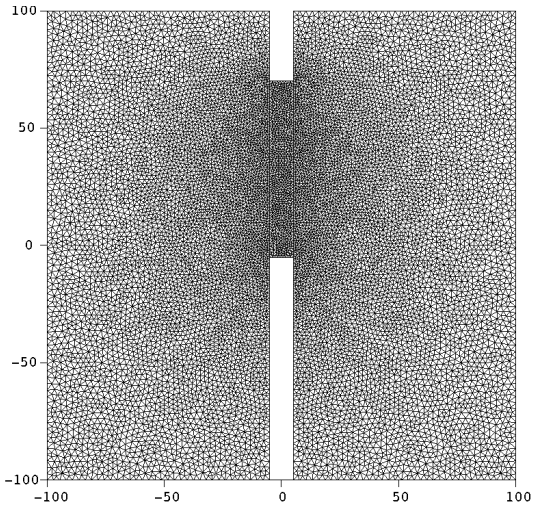

Fig. 8. Partial dam-break geometry and the Delaunay mesh (24750 triangles).

the Delaunay mesh, composed of 24750 triangles, is depicted in Figure 8. The breach corresponds to the subdomain $[-5, 5] \times [-5, 70]$ and the bathymetry function is given by

$$b(x,y) = \begin{cases} 1 & , \quad -100 \leq x < -5, \\ 0.1(5-x) & , \quad -5 \leq x < 5, \\ 0 & , \quad 5 \leq x \leq 100, \end{cases}$$

while the initial free surface is given by

$$H(x,y,0) = \begin{cases} 10\ , & -100 \leq x < 5, \\ 5\ , & 5 \leq x \leq 100. \end{cases}$$

At the initial time $t = 0$ the system is assumed to be at rest and we prescribe reflection boundary conditions on the whole boundary. The bathymetry is characterised by a $\mathbb{P}_1$ polynomial reconstruction since the domain is flat or constituted of a linear ramp. Numerical simulations have been carried out using different detectors (PAD, DMP and u2) and several polynomial reconstruction degrees ($\mathbb{P}_2$, $\mathbb{P}_3$ and $\mathbb{P}_5$).

4.5.1 Comparisons between the PAD, the DMP and the u2 detectors

We compute the solution using the $\mathbb{P}_5$ reconstruction and different detectors and display the free surface at the final time in Figure 9 using the PAD, PAD+DMP, PAD+DMP+u2 detectors respectively. We have labelled H_i, $i = 1, \ldots, 4$, the total height of the four vortexes at the final time t_{final} (see the left top panel for the corresponding location). In the same way, we denote H_{min} and H_{max} the minimum and maximum total height of the shock wave of the downstream flow front at the final time t_{final}. We recall that the PAD detector aims at preserving the non-negativity of the water height, the DMP or ED detectors reveal the extrema of the numerical solution and the u2 detector determines whether an extremum is smooth or derives from an oscillation.

We first carry out a simulation using only the PAD detector. No problematic cells are detected, hence the code basically runs without any limitation on the polynomial reconstruction used in the fluxes evaluation. Consequently, oscillations appear in the vicinity of the discontinuities as shown in Figure 9 left top panel. For instance, oscillations are clearly present in the vicinity of the flow front discontinuous shock wave leading to local overshoots and undershoots (also see the free surface values in Table 14). Nevertheless, we notice that there is almost no oscillations in the rarefaction wave travelling upstream as well as in the vortexes since the solution is locally smooth.

Introducing the DMP as detector within the MOOD scheme effectively reduces the oscillations (Figure 9 right top panel) but the number of cells corresponding to a $\mathbb{P}_0$ reconstruction significantly increases as shown in Figure 9 right bottom panel, leading to an important numerical viscosity.

To reduce such a diffusion associated to use of the DMP detector, the u2 detector is activated to separate the extrema deriving from a smooth function and the ones deriving from oscillations. Figure 9 left bottom panel shows that the discontinuities are sharper with respect to the DMP case, while no

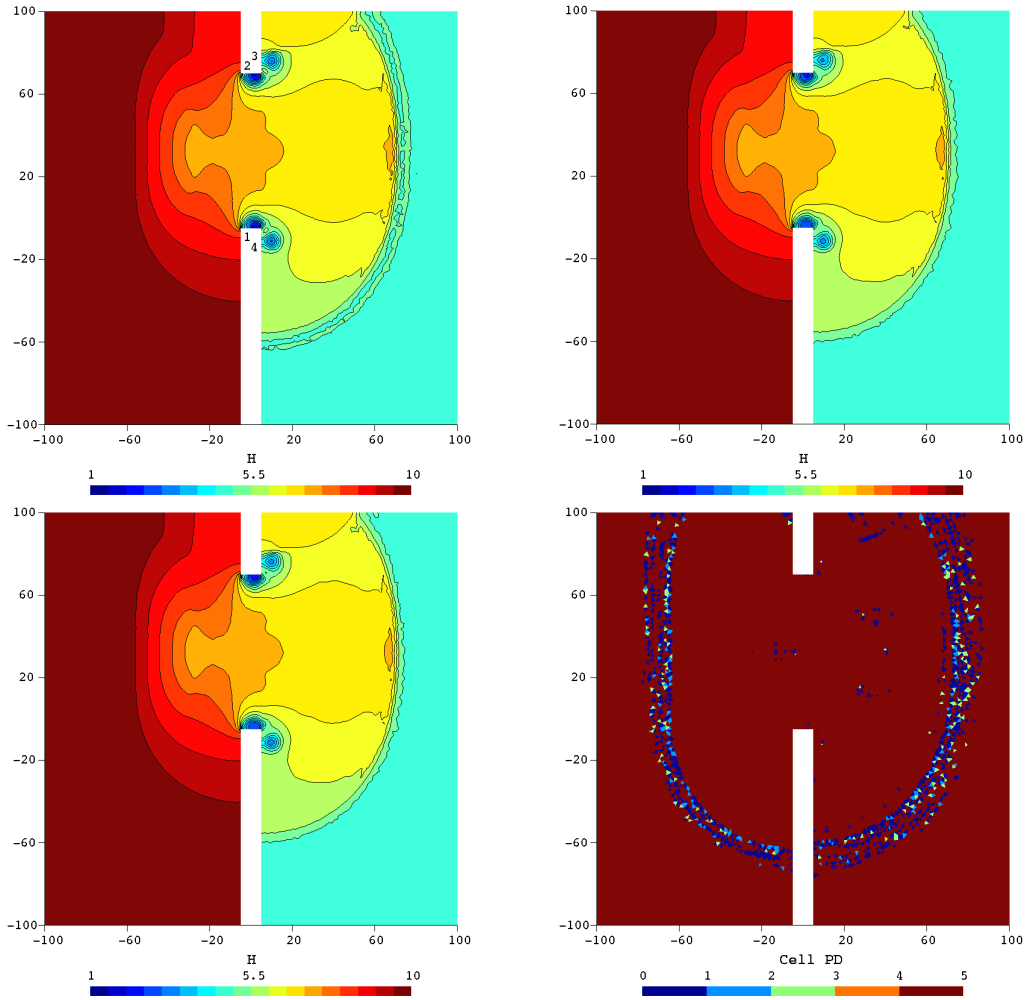

Fig. 9. Free surface at t_{final} using the $\mathbb{P}_5$ reconstruction. Left top panel: PAD. Right top panel: PAD+DMP. Left bottom panel: PAD+DMP+u2. Right bottom panel: CellPD map for the PAD+DMP at final time.

oscillations are generated along the fronts.

To assess the accuracy of the scheme, we measure the deepness of the four vortexes reported in Table 14. Indeed, the 4 vortexes show regular patterns despite their considerable deepness, in particular the deepest vortexes 1 and 2 located just above the ramp. To complete the test cases, we have carried out similar simulation using PAD, PAD+DMP and PAD+DMP+u2 but using $\mathbb{P}_2$ and $\mathbb{P}_3$ reconstructions and also report the vortex deepness and front total heights in Table 14. The situation for $\mathbb{P}_2$ and $\mathbb{P}_3$ is similar but softened with respect to $\mathbb{P}_5$ due to a higher numerical diffusion. The smoothing effect is also noticed in the front shock where the discontinuity is typically captured within 2 cells for the $\mathbb{P}_5$ reconstruction, between 2 and 3 cells with $\mathbb{P}_3$ and at least 3 or 4 cells with $\mathbb{P}_2$. We underline the important impact of the polynomial degree for the reconstruction for evaluating the small structures such as vortexes (covered by 10-15 cells) where the deepness values range between 2.46 ($\mathbb{P}_2$)

and 1.72 ($\mathbb{P}_5$).

Figure 9 right bottom panel displays the CellPD map of the solution at the final time using the $\mathbb{P}_5$ reconstruction and the DMP detector. A noticeable point is the very few problematic cells near the vortexes. However, it is enough to strongly reduce the local polynomial degree and to provide rough approximations of the deepness to highlight the strong impact of the limitation. On the other hand, the discontinuous fronts clearly appear in the CellPD map and we observe a degree reduction of the rarefaction tail despite the smoothness of the solution in this zone.

Table 14
Reference total water heights for the four vortexes and for the flow front, as well as CellPD percentages, for different reconstruction polynomial degrees and MOOD detection processes at final time.

Scheme		Vortexes Total Heights				Flow Front Total Heights		CellPD (%)			
Deg.	Detect.	H_1	H_2	H_3	H_4	H_{min}	H_{max}	$\mathbb{P}_0$	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_5$
$\mathbb{P}_2$	PAD	2.46	2.58	3.40	3.40	4.85	7.41	0	100		
	DMP	2.79	2.85	3.67	3.63	5.00	7.18	3.2	96.8		
	DMP+u2	2.78	2.83	3.54	3.51	5.00	7.18	1.0	99.0		
	DMP+u2$^\nu$	2.65	2.68	3.42	3.40	5.00	7.34	0.7	99.3		
$\mathbb{P}_3$	PAD	2.08	2.03	3.08	2.95	4.79	7.45	0	0	100	
	DMP	2.78	2.56	3.48	3.41	5.00	7.18	3.3	0.8	95.9	
	DMP+u2	2.69	2.49	3.35	3.15	4.99	7.19	1.0	0.2	98.8	
	DMP+u2$^\nu$	2.40	2.09	3.11	2.98	4.97	7.36	0.6	0.1	99.3	
$\mathbb{P}_5$	PAD	1.72	1.81	2.87	2.77	4.76	7.44	0	0	0	100
	DMP	2.50	2.31	3.36	3.33	5.00	7.16	3.9	0.5	0.6	95.0
	DMP+u2	2.51	2.16	3.17	3.21	4.99	7.16	0.9	0.2	0.1	98.8
	DMP+u2$^\nu$	1.90	1.85	2.90	2.83	4.95	7.38	0.4	0.1	0.0	99.5

The pictures analysis shows that oscillations using DMP+u2 are slightly larger than those obtained with the simple DMP, but considerable smaller than the ones appearing when only PAD is used. On the other hand, the vortexes 3 and 4 deepness are closer to the values obtained with PAD, whereas the values for vortexes 1 and 2, as well as the flow front free surface, are very similar to the ones obtained with DMP. Thus, the combination DMP+u2 improves the results obtained with PAD and DMP, combining the accuracy and robustness, but we do not recover the optimal accuracy in zones where the solution is relatively smooth as is the case of vortexes 1 and 2.

4.5.2 The u2 versus u2$^\nu$ detector

To close the section, we compare the two relaxation detectors to observe the consequences in term of accuracy and stability. We recall that the u2$^\nu$ uses a smaller index subset to detect whether the solution is locally admissible or not. Hence, using u2$^\nu$ we reduce the probability of rejecting the candidate solution.

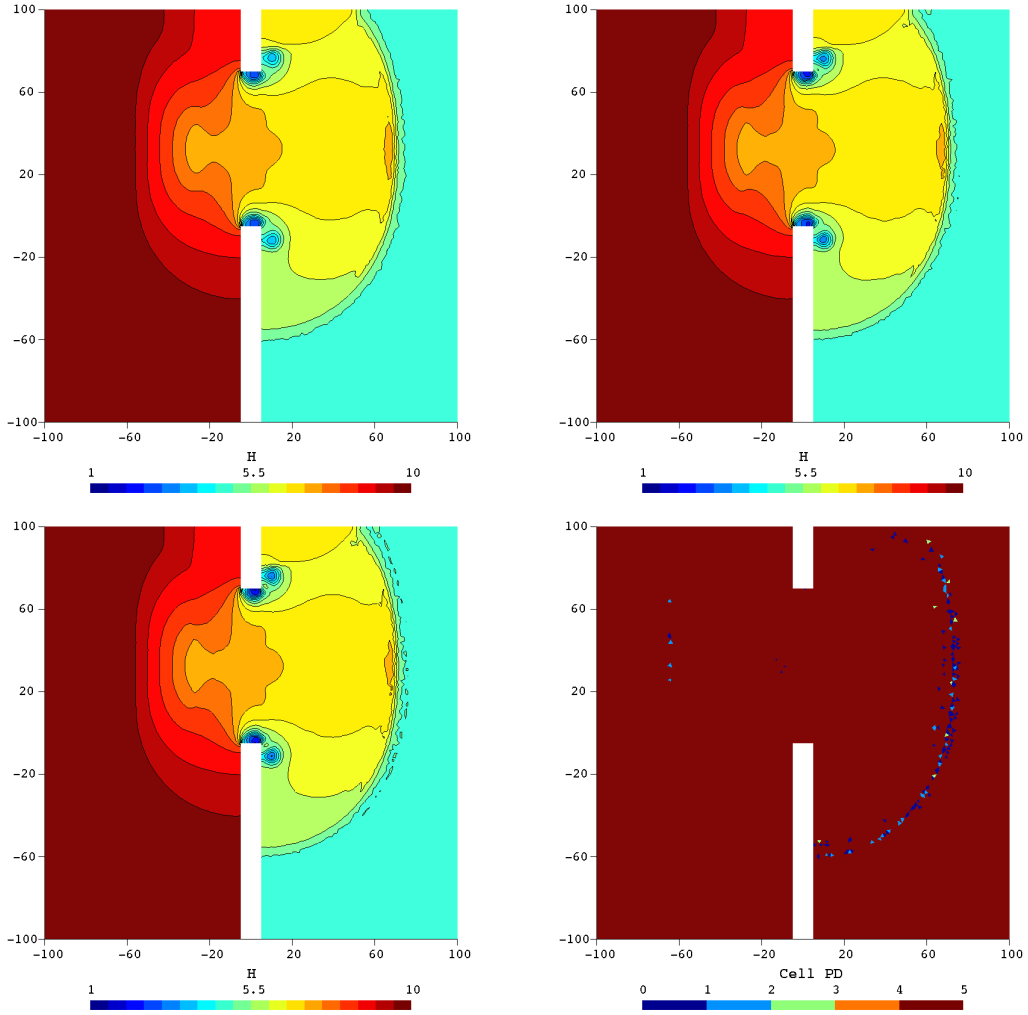

Fig. 10. Total height at t_{final} using the DMP+u2$^\nu$ detector. Left top panel: $\mathbb{P}_2$. Right top panel: $\mathbb{P}_3$. Left bottom panel: $\mathbb{P}_5$. Right bottom panel: CellPD map with the $\mathbb{P}_5$ reconstruction at final time.

Combining with the DMP detector, we display in Figure 10 the total height at the final time using the new DMP+u2$^\nu$ detector for different polynomial reconstructions $\mathbb{P}_2$, $\mathbb{P}_3$ and $\mathbb{P}_5$, while we report in Table 14 the characteristic values for the vortexes and front line. Vortexes 3 and 4 deepness are essentially those obtained with PAD, and vortexes 1 and 2 deepness are now much closer to the PAD values. We obtain a similar result with the maximum total height H_{max} of the flow front shock wave. The last column of Table 14 provides the

relative distribution (in percentage) of the polynomial degrees for the last time step. We observe that DMP+u2$^{\nu}$ strongly reduces the number of problematic cells, about one half with respect to the DMP+u2 values, providing around 99.4% of cells with the maximal order.

From the stability point of view, the oscillations nearby the shock wave (see also H_{min}) are very well-contained for the $\mathbb{P}_2$ case and are small (below 0.6%) with the $\mathbb{P}_3$ reconstruction, mainly confined near the upper boundary. As for the $\mathbb{P}_5$ situation, oscillations are spread along a large part of the shock wave and represent up to 1.0% of the total height. The CellPD map (see Figure 10 right bottom) shows that the polynomial degree is mainly maximal so the u2$^{\nu}$ detector has relaxed too much the DMP detector leading to larger overshoots and undershoots. Nevertheless, the DMP+u2$^{\nu}$ manages to reduce the oscillations far better than the PAD case and provide an acceptable solution.

5 Conclusion

The MOOD strategy, originally developed in the conservative framework of the Euler system has been extended to the non-conservative case using the classical 2D shallow-water system with varying bathymetry as a test case. Moreover, the concept of physical bathymetry representative has also been introduced to design a new class of numerical schemes. A large number of tests have been carried out to assess the performance of the scheme and we prove that we get an effective sixth-order of accuracy when dealing with smooth solutions, while no oscillations are reported in the vicinity of discontinuities. We have also checked that the scheme exactly preserves the lake at rest situation while the other steady-state situations (moving water case) are preserved up to the sixth-order. These encouraging results indicate that the MOOD strategy suits well to the non-conservative situations and future efforts will focus on the dry/wet situation where a new class of detectors has to be designed.

Acknowledgements

This research was financed by FEDER Funds through Programa Operacional Factores de Competitividade — COMPETE and by Portuguese Funds through FCT — Fundação para a Ciência e a Tecnologia, within the Projects PEst-OE/MAT/UI0013/2014 and FCT-ANR/MAT-NAN/0122/2012.

References

[1] S. Clain, S. Diot, R. Loubère, A high-order finite volume method for hyperbolic systems: Multi-dimensional Optimal Order Detection (MOOD), J. Comput. Phys. 230 Issue 10 (2011) 4028–4050.

[2] S. Diot, S. Clain, R. Loubère, Improved detection criteria for the Multi-dimensional Optimal Order Detection (MOOD) on unstructured meshes with very high-order polynomials, Comput. & Fluids Issue 64 (2012) 43–63.

[3] C. E. Castro, E. F. Toro, M.- käser, ADER scheme on unstructured meshes for shallow water: simulation of tsunami waves, Geophysical Journal International 189 (2012) 1505–1520.

[4] R. Omira, M.A. Baptista, J. M. Miranda, Evaluating tsunami impact on the Gulf of Cadiz coast (Northeast Atlantic), Pure and applied geophysics 168 (2011) 1033–1043.

[5] J. J. Wijetunge, Numerical simulation of the 2004 indian ocean tsunami: case study of effect of sand dunes on the spatial distribution of inundation in Hambantota, Sri Lanka, J. Appl. Fluid Mech. 3 (2010) 125–135.

[6] D. Brecht, P. Noble, Mathematical justification of the shallow water model, Methods and applications of analysis 14 (2007) 87–118.

[7] Y. Xing, Exactly well-balanced discontinuous Galerkin methods for the shallow water equations with moving water equilibrium, J. Comput. Phys. 257 (2014) 536–553.

[8] Y. Xing, 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 (2006) 567–598.

[9] C. Berthon, F. Fouchet, Efficient well-balanced hydrostatic upwind schemes for shallow-water equations, J. Comput. Phys. 231 (2012) 4993–5015.

[10] T. Gallouët, J. M. Hérard, N. Seguin, Some approximate Godunov scheme to compute shallow-water equations with topography, Computers and Fluids 32 (2003) 479–513.

[11] I.K. Nikolos, A.I. Delis, An unstructured node-centered finite volume scheme for shallow water flows with wet/dry fronts over complex topography, Comput. Methods Appl. Mech. Engrg. 198 (2009) 3723–3750.

[12] S. Vukovic, L. Sopta, ENO and WNO schemes with the exact conservation property for one-dimensional shallow water equations, J. Comput. Phys. 179 (2002) 593–621.

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

[14] S. Noelle, Y. Xing, C.-W. Shu, High-order well-balanced finite volume WENO schemes for shallow water equations with moving water, J. Comput. Phys. 226 (2007) 29–58.

[15] S. Gottlieb, C. W. shu, Total variation diminishing Runge-Kutta schemes, Math of Compt. 67 (1998) 73–85.

[16] V. A. Titarev, E. F. Toro, ADER schemes for three-dimensional non-linear hyperbolic systems, J. Comput. Phys. 204 (2005) 715–736.

[17] M. Dumbser, M. Castro, C. Parés, E. F. Toro, ADER schemes on unstructured meshes for nonconservative hyperbolic systems: applications to geophysical flows, Computers and Fluids 38 (2009) 1731–1748.

[18] S. Diot, R. Loubère, 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 (2013) 362–392.

[19] C. Berthon, V. Desveaux, An entropy preserving MOOD scheme for the Euler equations, Int. J. finite volumes, 11 (2014) 1–39.

[20] M. Dumber, O. Zanotti, R. Loubère, S. Diot, A posteriori subcell limiting for discontinuous Galerkin finite element method for hyperbolic system of conservation laws, J. Comput. Phys. 278 (2014) 47–75.

[21] R. Loubère, M. Dumbser, S. Diot, A new family of high order unstructured MOOD and ADER finite volume schemes for multidimensional systems of hyperbolic conservation laws, Communications in Computational Physics 16 (2014) 718–763.

[22] S. Diot1,M.M. François, E.D. Dendy, A higher-order unsplit 2D direct Eulerian finite volume method for two-material compressible flows based on the MOOD paradigms, Int. J. Numer. Meth. Fluids (2014), online version, DOI: 10.1002/fld.3966 (2014).

[23] T. Buffard, S. Clain, Monoslope and multislope MUSCL methods for unstructured meshes, J. Comput. Phys. 229 (2010) 3745–3776.

[24] R. Abgrall, On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation, J. Comput. Phys. 114 (1994) 45–58.

[25] G. Dal maso, P. LeFloch, F. Murat, Definition and weak stability of nonconservative products, J. Math. Pures Appl. 74 (1995) 483–548.

[26] L. Gosse, A well-balanced flux-vector splitting scheme designed for hypebolic systems of conservation laws with source terms, Comput. Math. Appl. 39 (2000) 135–159.

[27] C. E. Castro, J. A. López-García, C. Parés, High order exactly well-balanced numerical methods for shallow water systems J. Comput. Phys. 246 (2013) 242–264.

[28] M. J. Castro, P. G. LeFloch, M. L. Muñoz-Ruiz, C. Parés, Why many theories of shock waves are necessary: convergence error in formally path-consistent schemes, J. Comput. Phys. 227 (2008) 8107–8129.

[29] A. Bermúdez, M. E. Vázquez , Upwind methods for hyperbolic conservation laws with source terms, Computers & Fluids 24 (1994) 1049–1071.

[30] A. Bermúdez, A. Dervieux, J.-A. Desideri, M. E. Vásquez, Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes, Comput. Methods Appl. Mech. Engrg. 155 (1998) 49–72.

[31] R. J. LeVeque, Balancing source terms and flux gradients on high-resolution Godunov methods: the quasi-steady wave-propagation algorithm, J. Comput. Phys. 146 (1998) 346–365.

[32] M. Hubbard, P. Garcia-Navarro, Flux difference splitting and the balancing of source terms and flux gradients, J. Comput. Phys. 165 (2000) 89–125.

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

[34] C. Berthon, F. Marche, R. Turpault, An efficient scheme on wet/dry transitions for Shallow Water Equations with friction, Computer & Fluids, 48 (2011) 192–201.

[35] A. Canestrelli, A. Siviglia, M. Dumbser, E. F- Toro Well-balanced high-order centered schemes for non-conservative hyperbolic systems. Application to shallow water equations with fixed and mobil bed, Advances in water resources 32 (2009) 834–844.

[36] V. Caleffi, A. Valiani, A. Bernini, Fourth-order balanced sourced term treatment in central WENO schemes for shallow water equations, J. Comput. Phys. 218 (2006) 228–245.

[37] A. Kurganov, D. Levy, Central-upwind schemes for the Saint-Venant system, ESAIM: Math. Model. Num. Anal. 36 (2002) 397–425.

[38] Y. Xing, C. W. Shu, High order finite difference WENO schemes with the exact conservation property fot the shallow water equations, J. Comput. Phys. 208 (2005) 206–227.

[39] Y. Xing, C. W. Shu, S. Noelle, On the advantage of well-balanced schemes for moving water equilibira of the shallow water equations, J. Sci. Comput. 48 (2011) 339–349.

[40] F. Alcrudo, F. Benkhaldoum, Exact solution to the Riemann problem of the shallow water equations with bottom step, Computers & Fluids 30 (2001) 643–671.

[41] A. Noussair, Riemann problem with nonlinear resonance effects and well-balanced Godunov scheme for shallow water fluid flow past an obstacle, SIAM J. Numer. Anal. 39 (2001) 52–72.

[42] R. Bernetti, V. A. Titarev, E. F. Toro, Exact solution of the Riemann problem for the shallow water equations with discontinuous bottom geometry, J. Comput. Phys. 22 (2008) 3212–3243.

[43] P. G. LeFloch, M. D. Thanh, A Godunov-type method for the shallow water equations with discontinuous topography in the resonant regime, J. Comput. Phys. 230 (2011) 7631–7660.

[44] M. D. Thanh, Numerical treatment in resonant regime for shallow water equations with discontinuous topography, Commun. nonlinear Sci. Numer. Simulat. 18 (2013) 417–433.

[45] N. Andrianov, Performance of numerical methods on the non-unique solution to the Riemann problem for the shallow water equations, Int. J. Numer. Meth. Fluids 47 (2005) 825–831.

[46] J. G. Zhou, D. M. Causon, C. G. Mingham, D. M. Ingram, The Surface Gradient Method for the Treatment of Source Terms in the Shallow-Water Equations, J. Comput. Phys. 168 (2001) 1–25.

[47] J. G. Zhou, D. M. Causon, C. G. Mingham, D. M. Ingram, Numerical solutions of the shallow water equations with discontinuous bed topography, Int. J. Numer. Meth. Fluids 38 (2002) 769–788.

[48] A. Duran, Q. Liang, F. Marche, On the well-balanced numerical discretization of shallow water equations on unstructured meshes, J. comput. Phys. 235 (2013) 565–586.

[49] S. Clain, G. Machado, J. M. Nóbrega, R. Pereira, A sixth-order finite volume method for multidomain convection-diffusion problem with discontinuous coefficients, Computer Methods in Applied Mechanics and Engineering, 267 (2013) 43–64.

[50] E. F. Toro, Riemann Solvers and Numerical Methods for Fluid Dynamics, 3^{rd} revision, Springer-Verlag Berlin and Heidelberg GmbH & Co. K (2009).

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

[52] E. F. Toro, M. Spruce, W. Spares, Restoration of the contact surface in the HLL Riemann solver, Shock Wave 4 (1994) 25–34.

[53] Y. Huang, N. Zhang and Y. Pei, Well-ballanced finite volume scheme for shallow water flooding and drying over arbitarary topography, Engineering Applications of Computational Fluid Mechanics 7 (2013) 40–54.

[54] C. Geuzaine and J.-F. Remacle, Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities, Int. J. Numer. Meth. Eng. 79 (2009) 1309–1331.

[55] O. Delestre, C. Lucas, P.-A. Ksinant, F. Darboux, C. Laguerre, T.-N.-T. Vo, F. James, S. Cordier, SWASHES: a compilation of shallow water analytic solutions for hydraulic and environmental studies, Int. J. Numer. Meth. Fluids 72 (2013) 269–300.