# 2D Well-balanced Augmented ADER schemes for the Shallow Water Equations with bed elevation and extension to the rotating frame

A. Navas-Montilla [1], J. Murillo

*anavas@unizar.es , Fluid Mechanics. LIFTEC, CSIC-Universidad de Zaragoza. Zaragoza, Spain*

**Abstract**

In this work, an arbitrary order augmented WENO-ADER scheme for the resolution of the 2D Shallow Water Equations (SWE) with geometric source term is presented and its application to other shallow water models involving non-geometric sources is explored. This scheme is based in the 1D Augmented Roe Linearized-ADER (ARL-ADER) scheme, presented by the authors in a previous work and motivated by a suitable compromise between accuracy and computational cost. It can be regarded as an arbitrary order version of the Augmented Roe solver, which accounts for the contribution of continuous and discontinuous geometric source terms at cell interfaces in the resolution of the Derivative Riemann Problem (DRP). The main novelty of this work is the extension of the ARL-ADER scheme to 2 dimensions, which involves the design of a particular procedure for the integration of the source term with arbitrary order that ensures an exact balance between flux fluctuations and sources. This procedure makes the scheme preserve equilibrium solutions with machine precision and capture the transient waves accurately. The scheme is applied to the SWE with bed variation and is extended to handle non-geometric source terms such as the Coriolis source term. When considering the SWE with bed variation and Coriolis, the most relevant equilibrium states are the still water at rest and the geostrophic equilibrium. The traditional well-balanced property is extended to satisfy the geostrophic equilibrium. This is achieved by means of a geometric reinterpretation of the Coriolis source term. By doing this, the formulation of the source terms is unified leading to a single geometric source regarded as an apparent topography. The numerical scheme is tested for a broad variety of situations, including some cases where the first order scheme ruins the solution.

*Keywords:* ADER, Shallow water, Source terms, Coriolis, Well-balanced, High order accuracy.

## 1. Introduction

Finite Volume (FV) numerical schemes are a very common choice for the resolution of complex flows of hyperbolic nature. Such schemes have experienced a great improvement in the past decades thanks to the development of sophisticated schemes that ensure a high order of accuracy both in space and time when computing the solution. In the framework of FV, the introduction of the ENO and, specially, the WENO reconstruction techniques [4, 5] supposed a major step when seeking arbitrary order of accuracy in space. On the other hand, the preservation of high order in time was generally done by means of a Runge-Kutta time integrators, which sometimes proved to be inefficient due to Butcher's barrier [6]. This issue was circumvented by the ADER approach [7, 8], which provides a fully discrete scheme of arbitrary order. The original ADER idea can be regarded as a high order generalization of Godunov's method by means of a Taylor power series expansion in time, where time derivatives are computed from the reconstructed space derivatives by means of the Cauchy-Kowalewski (CK) procedure. ADER schemes consist of two steps: first, a high-order spatial reconstruction procedure and secondly, the resolution of a high order extension of the Riemann Problem (RP), called Derivative Riemann Problem (DRP) [9]. The initial condition for the DRP consists of piecewise polynomial

data with $K$ nontrivial derivatives, usually constructed by means of non-oscillatory and arbitrary-order reconstruction procedures, such as the aforementioned ENO or WENO methods. The ADER approach successfully allows the construction of arbitrary order schemes for systems of hyperbolic conservation laws [9, 10, 11, 6, 12]. It is of particular interest for this work to consider the application of ADER schemes for the resolution of geophysical problems [13, 14, 15], specially for the resolution of the Shallow Water Equations (SWE) [16, 15, 17, 18, 19], which will be the focus of this work.

An approach for the resolution of the $\text{DRP}_K$ was first presented in [9] and the proposed solver was called Toro–Titarev (TT) solver. This solver allows to reduce the $\text{DRP}_K$ to a series of classical Riemann problems where classic Riemann solvers are of application. The $\text{DRP}_K$ consists of one RP for the leading term, referred to as $\text{DRP}_0$, plus $K$ additional RPs for the derivatives. Up to date, a broad variety of Derivative Riemann solvers have been designed with the aim of providing accurate and fast solutions. Apart from the TT solver, the most common solvers we can find in the literature are the HEOC solver [16, 4] and the Castro–Toro (CT) solver [16]. A semi–implicit version of those schemes was proposed by Montecinos [20], allowing to deal with more stiff source terms. It is worth pointing out that the traditional ADER approach using the CK procedure may become rather cumbersome when dealing with complex systems of equations and may not provide the expected performance when dealing with very stiff source terms. In such a case, a successful solution would be to replace the CK procedure by a local spacetime Galerkin method, as done by the DET solver, proposed by Dumbser in [21].

The TT, CT and HEOC solvers assume that there is a single solution, referred to as the star solution [22], with independence of the presence of source terms. However, when dealing with geometric source terms it was shown that the consideration of the source term in the resolution of the DRP is more convenient in order to ensure certain properties of the numerical solution [23, 24]. Those solvers that account for the source term in the resolution of the (D)RP are called augmented solvers [23] and allow to exactly preserve the equilibrium states [25, 26, 27].

In this work, we will focus on the use of augmented solvers for the resolution of the DRP in order to construct arbitrary order schemes for the 2D SWE with source term. To account for the contribution of the source term in the DRP, the FS and LFS augmented solvers were introduced in [28, 29]. They can be combined with the Augmented Roe (ARoe) solver [24] and HLLS solvers[26], which are the augmented versions of the Roe [30] and HLL solvers [31, 32]. Such solvers, called AR-ADER, ARL-ADER, HLLS-ADER and HLLSL-ADER schemes, provide an arbitrary order approximate solution to the DRP. Here, the LFS solver in combination with the ARoe solver will be used. Such method showed a good performance when solving the SWE with geometric source term and is computationally more efficient than the FS solver [29].

This work has two main aims: (a) to extend the aforementioned methods to 2D to solve the 2D SWE with geometric source term and (b) to explore the resolution of the SWE with source terms of non-geometric nature by doing a geometric reinterpretation of those sources, which allows the preservation of equilibrium states. We first develop the 2D extension of the 1D ARL-ADER scheme in [29] for hyperbolic systems of equations with geometric source terms. This is achieved by means of two techniques: (a) the use of the augmented LFS solver and (b) a particular procedure for the integration of the source term inside cells that ensures an exact balance between flux fluctuations and source terms. This procedure preserves the equilibrium states with machine precision and guarantees an arbitrary order of accuracy thanks to the use of Romberg integration method [51], provided an optimal derivation of time derivatives. Then, the aforestated methods are extended for the numerical treatment of other types of source terms by expressing them in geometric form

When considering the numerical resolution of the SWE with source terms, is must be borne in mind that a suitable discretization of such terms must be considered in order to provide physically feasible solutions. If considering first a quiescent equilibrium, where water is at rest with no rotation, the very first property that the numerical scheme must satisfy is the preservation of such quiescent

equilibrium in the discrete level. Numerical schemes able to preserve still water at rest are called well-balanced methods [34, 35]. There is a broad variety of well-balanced methods based on Riemann solver techniques [36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49]. For a comparison of different well-balanced methods in the framework of DG schemes, see [50]. The well-balanced property can still be enhanced by considering energy conservation criteria in the numerical scheme [51, 52, 53, 54, 55, 56, 57, 58, 28, 29, 55].

The SWE in the rotating frame represents a good model for large scale phenomena in geophysical flows [60, 59]. For such model, the most relevant equilibrium solution is the so-called geostrophic equilibrium state, often referred to as *jet in the rotating frame* [59]. In the last decade, a great effort has been put on the design of FV well-balanced numerical schemes capable to maintain the geostrophic equilibrium by following different approaches [61, 62, 63, 64, 65, 66, 60, 59].

For the particular case of the SWE with bed elevation and Coriolis, the well-balanced property that allowed to correctly simulate the still water at rest is extended to preserve the geostrophic equilibrium. We define two directional primitive variables for the $x$ and $y$ contributions of the Coriolis force, as done in [65, 60, 59], which allows to express again the Coriolis source term as a geometric source term. By doing this, both source terms can be merged into a single geometric source where the scalar variable can be regarded as an apparent topography [65, 59]. The numerical techniques designed here for the discretization of the bed slope source term can be extended to the Coriolis source term while retaining the well-balanced property, even in cases with discontinuous bed topography. However, the high order of accuracy provided by the scheme will help to accurately converge to such solution. It is worth pointing out that the preservation of the geostrophic equilibrium is done at the cost of loosing the optimal accuracy of the integration in time, as the CK procedure has to be carried out in a particular way that combines the use of conserved and primitive source term variables.

The outline of the paper is next presented. In section 2, an introduction to nonlinear systems of conservation laws with geometric source terms is provided. In Section 3, we present the general formulation of 2D ADER schemes in Cartesian grids and provide the details for the construction of the ARL-ADER scheme. Chapter 4 is devoted to the application of the scheme to the SWE and describes a well-balanced integration method for both the bed elevation and Coriolis source term. Chapter 5 includes a variety of test cases for the SWE with bed variation, rotation and with both features at the same time. Finally, in Chapter 6 we present a summary of the work and the concluding remarks.

## 2. Nonlinear systems of equations with source term

The basic ideas underlying this work can be illustrated by examining hyperbolic nonlinear systems of equations with source terms in 2D, that can be expressed in differential formulation as

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{E}(\mathbf{U}) = \mathbf{S}\,, \qquad \forall x, y \in \Omega \subseteq \mathbb{R}^2 \tag{1}$$

where $\mathbf{U} = \mathbf{U}(\mathbf{x}, t) \in \mathcal{C} \subset \mathbb{R}^{N_\lambda}$ is the vector of conserved quantities that takes values on $\mathcal{C}$, with $N_\lambda$ the number of equations. $\mathbf{E}(\mathbf{U}) : \mathcal{C} \longrightarrow \mathbb{R}^{N_\lambda \times 2}$ is the matrix of fluxes, that will hereafter referred to as $\mathbf{E} = (\mathbf{F}, \mathbf{G})$, where $\mathbf{F} = \mathbf{F}(\mathbf{U}) : \mathcal{C} \longrightarrow \mathbb{R}^{N_\lambda}$ and $\mathbf{G} = \mathbf{G}(\mathbf{U}) : \mathcal{C} \longrightarrow \mathbb{R}^{N_\lambda}$ are the physical fluxes on the coordinate directions $x$ and $y$. Note that $\mathbf{x} = (x, y)$.

It is possible to define two Jacobian matrices for the fluxes $\mathbf{F}(\mathbf{U})$ and $\mathbf{G}(\mathbf{U})$ as

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}(\mathbf{U})}{\partial \mathbf{U}}\,, \qquad \mathbf{B}(\mathbf{U}) = \frac{\partial \mathbf{G}(\mathbf{U})}{\partial \mathbf{U}}\,, \tag{2}$$

that allow to rewrite (1) as

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}(\mathbf{U})\frac{\partial \mathbf{U}}{\partial x} + \mathbf{B}(\mathbf{U})\frac{\partial \mathbf{U}}{\partial y} = \mathbf{S}\,, \qquad \forall x, y \in \Omega \subseteq \mathbb{R}^2 \tag{3}$$

and provide sufficient information for the hyperbolicity of the system. The system in (1) is said to be *hyperbolic* if the matrix $\mathcal{J}(\mathbf{U}) \in \mathbb{R}^{N_\lambda \times N_\lambda}$ defined as

$$\mathcal{J}(\mathbf{U}) = k_1 \mathbf{A}(\mathbf{U}) + k_2 \mathbf{B}(\mathbf{U}) \,, \tag{4}$$

is diagonalizable with real eigenvalues for all $\mathbf{k} = (k_1, k_2) \in \mathbb{R}^2$ and for all $\mathbf{U} \in C$ with $C \subseteq \mathbb{R}^{N_\lambda}$ the subset of physically relevant values of $\mathbf{U}$. If the $N_\lambda$ eigenvalues are distinct, then the system is said to be *strictly hyperbolic* [67].

Assuming that the system is hyperbolic with $N_\lambda$ real eigenvalues $\lambda^m(\mathbf{U})$ and $N_\lambda$ linearly independent right eigenvectors $\mathbf{e}^m(\mathbf{U})$, it is possible define two matrices $\mathbf{P}(\mathbf{U}) = (\mathbf{e}^1(\mathbf{U}), \mathbf{e}^2(\mathbf{U}), ..., \mathbf{e}^{N_\lambda}(\mathbf{U}))$ and $\mathbf{P}^{-1}(\mathbf{U})$ with the property that they diagonalize the Jacobian as

$$\mathcal{J}(\mathbf{U}) = \mathbf{P}(\mathbf{U}) \mathbf{\Lambda}(\mathbf{U}) \mathbf{P}^{-1}(\mathbf{U}) \tag{5}$$

with $\mathbf{\Lambda}(\mathbf{U}) = \text{diag}\left(\lambda^1(\mathbf{U}), ..., \lambda^{N_\lambda}(\mathbf{U})\right)$ a diagonal matrix composed by the eigenvalues of the Jacobian.

In this work, we put an special emphasis on the so-called *geometric source terms* for the momentum equation, whose vector components are generally expressed as $S(\mathbf{U}, x, y) = S_s(\mathbf{U}) \partial_\varphi \phi(x, y)$, where $\varphi$ can either be $x$ or $y$ (depending on the vector component of the momentum which the source is applied to). $S_s(\mathbf{U})$ is a function of the conserved quantities and $\phi(x, y)$ is the geometric function, normally a potential that depends upon the position $x, y$ and that can be discontinuous.

## 3. General formulation of 2D ADER schemes in Cartesian grid

Let us consider again the system of conservation laws in (1) to compose the following Initial Boundary Value Problem (IBVP)

$$\begin{cases} \text{PDEs:} & \dfrac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{E}(\mathbf{U}) = \mathbf{S} \\ \\ \text{IC:} & \mathbf{U}(\mathbf{x}, 0) = \mathring{\mathbf{U}}(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega \\ \\ \text{BC:} & \mathbf{U}(\mathbf{x}, t) = \mathbf{U}_{\partial\Omega}(\mathbf{x}, t) \quad \forall \mathbf{x} \in \partial\Omega \end{cases} \tag{6}$$

defined in the domain $\Omega \times [0, T]$, where $\Omega = [a, b] \times [c, d]$ is the spatial domain. Note that the initial condition is given by $\mathring{\mathbf{U}}(\mathbf{x})$ and the boundary condition by $\mathbf{U}_{\partial\Omega}(\mathbf{x}, t)$. The spatial domain is discretized in $N_x \times N_y$ volume cells, defined as $\Omega_{ij} \subseteq \Omega$, such that $\Omega = \bigcup_{i,j=1}^{N} \Omega_{ij}$, with cell edges at

$$a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < ... < x_{N_x - \frac{1}{2}} < x_{N_x + \frac{1}{2}} = b \,, \tag{7}$$

and

$$c = y_{\frac{1}{2}} < y_{\frac{3}{2}} < ... < y_{N_y - \frac{1}{2}} < y_{N_y + \frac{1}{2}} = d \,, \tag{8}$$

Cells and cell are sizes defined as

$$\Omega_{ij} = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right] \times \left[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}\right] \,, \qquad i = 1, ..., N_x, \; j = 1, ..., N_y \tag{9}$$

and

$$\vartheta_{ij} = (x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}) \cdot (y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}) \,, \qquad i = 1, ..., N_x, \; j = 1, ..., N_y \,, \tag{10}$$

respectively, and in the case of regular grid we have $\vartheta_{ij} = \Delta x^2$.

Inside each cell at time $t^n = \sum_{l=1}^{n} \Delta t_l$, with $\Delta t_l$ the time step dynamically calculated, the conserved quantities are defined as cell averages as

$$\overline{\mathbf{U}}_{ij}^{n} = \frac{1}{\vartheta_{ij}} \int_{\Omega_{ij}} \mathbf{U}(\mathbf{x}, t^n) dA \qquad i = 1, ..., N_x,\ j = 1, ..., N_y\,. \tag{11}$$

where $dA = dxdy$ is the differential element of surface. Let us consider again the system in (1) and integrate it over the discrete domain $\Omega_{ij} \times \Delta t$, where $\Delta t = t^{n+1} - t^n$. Application of the Gauss-Ostrogradsky theorem yields

$$\overline{\mathbf{U}}_{ij}^{n+1} = \overline{\mathbf{U}}_{ij}^{n} - \frac{1}{\vartheta_{ij}} \int_0^{\Delta t} \int_{\partial\Omega_{ij}} \mathbf{E} \cdot \hat{\mathbf{n}} dl dt + \frac{1}{\vartheta_{ij}} \int_0^{\Delta t} \int_{\Omega_{ij}} \mathbf{S}\, dA dt\,, \tag{12}$$

where $dl$ is the differential length. If considering a regular Cartesian grid, all cells have a constant cell area $\Delta x^2$ and we obtain the following fully-discrete updating formula

$$\overline{\mathbf{U}}_{ij}^{n+1} = \overline{\mathbf{U}}_{ij}^{n} - \frac{\Delta t}{\Delta x^2} \left( \sum_{r=1}^{4} \mathcal{F}_r^- - \bar{\mathbf{S}}_{ij} \right), \tag{13}$$

where

$$\bar{\mathbf{S}}_{ij} \approx \frac{1}{\Delta t} \int_0^{\Delta t} \int_{x_{i+1/2}}^{x_{i-1/2}} \int_{y_{i+1/2}}^{y_{i-1/2}} \mathbf{S}\; dy dx\; d\tau\,. \tag{14}$$

is the approximation of the space-time integral of the source term inside the cell and $\mathcal{F}_r^-$ is the space-time integral of the numerical fluxes over the $r$-th cell edges. To construct a numerical scheme of order $K+1$-th, it is sufficient to approximate the integral of the flux, $\mathcal{F}_r^-$, using a $K+1$-th order Gaussian quadrature rule as

$$\mathcal{F}_r{}^- = \frac{\Delta x}{2} \sum_{q=1}^{k} w_q \mathcal{F}_{r,q}^-\,, \tag{15}$$

where $w_l$ are the Gaussian weights inside the interval $[-1, 1]$ at the $q = 1, ..., k$ quadrature points along the cell edge and $\mathcal{F}_{r,q}^-$ the numerical fluxes at each of these points, computed by means of the resolution of a 1D approximation to the Cauchy problem with at least $K$ non-trivial derivatives, to ensure high order not only in space but also in time. It is worth recalling that the use of $k$ quadrature points for the Gaussian integration allows to construct a $2k-1$-th order approximation of the integral.

It is worth noting that the notation $\mathcal{F}_{r,q}^-$ for the numerical fluxes has been adopted in order to to express them as cell-leaving fluxes and provide a more compact formulation of the scheme. Generally, the flux $\mathcal{F}$ stands for $\mathbf{F} \cdot n_x + \mathbf{G} \cdot n_y$, but it can only be either $\mathcal{F} = \pm\mathbf{F}$ or $\mathcal{F} = \pm\mathbf{G}$ when considering a Cartesian mesh. In this particular case, the numerical fluxes $\mathcal{F}_{r,q}^-$ can be expressed as $\mathcal{F}_{N,q}^- = \mathbf{G}_{i,j+1/2,q}^-$, $\mathcal{F}_{S,q}^- = -\mathbf{G}_{i,j-1/2,q}^+$, $\mathcal{F}_{E,q}^- = \mathbf{F}_{i+1/2,j,q}^-$ or $\mathcal{F}_{W,q}^- = -\mathbf{F}_{i-1/2\,j,q}^+$. Note that the subscripts for the $r$ loop have been labeled as 3=N, 1=S, 2=E and 4=W, which stand for north, south, east and west.

The numerical fluxes are computed solving an arbitrary order approximation of the Cauchy problem at cell interfaces. This is given by a 1D DRP, which is defined in the $x$ direction for the numerical fluxes on the east and west interfaces and in the $y$ direction for those fluxes on the north and south interfaces. It is worth noting that the source term is included in the definition of the DRP, according to [29]. The DRP$_K$ defined in the $x$ direction, at the interface $i+1/2$ and quadrature point $q$, reads [29]

$$\begin{cases} \dfrac{\partial \mathbf{U}}{\partial t} + \dfrac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = \mathbf{S} \\ \\ \mathbf{U}(x, t=0) = \begin{cases} \mathbf{U}_{ij}(x, y_{ij,q}) & x < 0 \\ \mathbf{U}_{i+1\,j}(x, y_{i+1\,j,q}) & x > 0 \end{cases} \end{cases} \tag{16}$$

where $\mathbf{U}_{ij}(x, y_{i+1/2,j,q})$ and $\mathbf{U}_{i+1\,j}(x, y_{i+1/2,j,q})$ are smooth functions of the position, defined using suitable reconstruction procedures, such as the WENO method. Such functions are evaluated at the particular $y = y_{i+1/2,j,q}$ location where the 1D $x$-oriented DRP is defined.

The solution for the DRP in (16) can be constructed using the flux expansion approach as

$$\begin{aligned}\mathbf{F}^{-}_{i+1/2,j,q} &= \mathbf{F}^{-,(0)}_{i+1/2,j,q} + \sum_{k=1}^{K} \mathbf{F}^{-,(k)}_{i+1/2,j,q} \frac{\Delta t^k}{(k+1)!}\,,\\ \mathbf{F}^{+}_{i+1/2,j,q} &= \mathbf{F}^{+,(0)}_{i+1/2,j,q} + \sum_{k=1}^{K} \mathbf{F}^{+,(k)}_{i+1/2,j,q} \frac{\Delta t^k}{(k+1)!}\,,\end{aligned} \tag{17}$$

where $\mathbf{F}^{-,(0)}_{i+1/2,j,q}$ , $\mathbf{F}^{-,(k)}_{i+1/2,j,q}$, $\mathbf{F}^{+,(0)}_{i+1/2,j,q}$ and $\mathbf{F}^{+,(k)}_{i+1/2,j,q}$ are computed by solving the DRP$_K$. When using the LFS-ARoe solver, the coefficients of (17) read

$$\begin{aligned}\mathbf{F}^{-,(k)}_{i+1/2} &= \mathbf{F}^{(k)}_{i_E} + \textstyle\sum_{m=1}^{N_\lambda} \left(\widetilde{\lambda}^{-}\alpha^{(k)} - \beta^{-,(k)}\right)^m_{i+1/2} \widetilde{\mathbf{e}}^m_{i+1/2}\,, \quad k = 0, K\\ \mathbf{F}^{+,(k)}_{i+1/2} &= \mathbf{F}^{(k)}_{(i+1)_W} + \textstyle\sum_{m=1}^{N_\lambda} \left(\widetilde{\lambda}^{+}\alpha^{(k)} - \beta^{+,(k)}\right)^m_{i+1/2} \widetilde{\mathbf{e}}^m_{i+1/2}\,, \quad k = 0, K\end{aligned} \tag{18}$$

where $\mathbf{F}^{(k)}_{i_E}$ and $\mathbf{F}^{(k)}_{(i+1)_W}$ are the left and right-hand limits to the cell edge of the physical flux ($k = 0$) and their $k$-th time derivatives, $\alpha^{(k)}$ are the wave strengths, $\beta^{-,(k)}$ the source strenghts and $\widetilde{\lambda}^{\pm}$ and $\widetilde{\mathbf{e}}^m_{i+1/2}$ the approximate wave celerities and eigenvectors defined using Roe's averages. The computation of the aforementioned quantities is detailed in Appendix A

Analogously, the DRP in (16) can be defined in the $y$ direction and the numerical fluxes on the north and south interfaces, $\mathbf{G}^{-}_{i,j+1/2,q}$ and $\mathbf{G}^{+}_{i,j-1/2,q}$, can be computed as follows

## 4. Application to the Shallow Water Equations

The SWE with bottom topography in a rotating frame can be expressed in matrix form as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial y} = \mathbf{S}\,, \quad \mathbf{S} = \mathbf{S}_b + \mathbf{S}_c \tag{19}$$

where

$$\mathbf{U} = (h, hu, hv)^T\,, \quad \mathbf{F} = \left(hu, hu^2 + \frac{1}{2}gh^2, huv\right)^T\,, \quad \mathbf{G} = \left(hv, huv, hv^2 + \frac{1}{2}gh^2\right)^T\,, \tag{20}$$

are the vectors of conserved quantities and physical fluxes in the $x$ and $y$ directions and

$$\mathbf{S}_b = \left(0, -gh\frac{dz}{dx}, -gh\frac{dz}{dy}\right)^T\,, \quad \mathbf{S}_c = (0, fhv, -fhu)^T\,, \tag{21}$$

are the vectors of sources due to bed variation and Coriolis force, respectively, where $h$ is the water depth, $u$ and $v$ are the $x$ and $y$ velocities respectively, $z$ is the bed elevation and $f$ the Coriolis coefficient. Equations (19)–(21) represent a good model for both small and large scale phenomena in geophysical flows, the latter being dominated by the Coriolis source term.

In the following subsections, a detailed derivation of the approximation of the source term for the SWE with bottom elevation and Coriolis, to construct a well-balanced scheme in Cartesian grids, is presented. Details on the calculation of the Roe's averages and source strengths for the construction of the approximate solution using the LFS solver are presented in Appendix B.

*4.1. Formulation of a well-balanced scheme for the SWE with bottom topography*

For now, we will consider the SWE in (19)–(21) with $\mathbf{S}_c = 0$, that is, in a fixed frame. The well-balanced formulation can be regarded as a weaker exact conservation property than the energy balanced formulation. The energy balanced formulation ensures the exact conservation property for both quiescent and moving equilibrium cases by considering the mechanical energy, $e = |\mathbf{v}|^2/2g + h + z$. On the other hand, the well-balanced formulation only satisfies the exact conservation property when $\mathbf{v} = 0$, that is, for still water at rest.

Under steady conditions, when the velocity vanishes, $\mathbf{v} = 0$, the equation for the conservation of energy and momentum yield the same result

$$\nabla (h + z) = 0 \tag{22}$$

which is known in the literature as *lake at rest* condition. At the discrete level and considering a Cartesian grid, between two points $\mathbf{x}_1$ and $\mathbf{x}_2$, Equation (22) can be decomposed into the Cartesian directions as

$$\delta (h + z)_{x_2,x_1} = 0\,, \qquad \delta (h + z)_{y_2,y_1} = 0 \tag{23}$$

To construct a well-balanced scheme, the previous discrete conditions in (23) must be satisfied. This can only be achieved if the WENO reconstruction method is applied to $\eta = h + z$ and $z$ first, and $h$ is computed from the difference of these reconstructions as $h^{(0)}_{(\cdot)} = \eta^{(0)}_{(\cdot)} - z^{(0)}_{(\cdot)}$, where $\eta^{(0)}_{(\cdot)}$ and $z^{(0)}_{(\cdot)}$ are the reconstructed water surface elevation and bottom elevation and $h^{(0)}_{(\cdot)}$ the computed water depth. The discharges $hu$ and $hv$ are also reconstructed using the WENO method.

**Theorem 1.** *Let us consider that the initial piecewise constant data for the problem satisfy the quiescent equilibrium condition in (23), that variables $\eta^{(0)}_{(\cdot)}$, $hu^{(0)}_{(\cdot)}$, $hv^{(0)}_{(\cdot)}$ and $z^{(0)}_{(\cdot)}$ are reconstructed using the WENO method and that $h^{(0)}_{(\cdot)}$ is reconstructed as $h^{(0)}_{(\cdot)} = \eta^{(0)}_{(\cdot)} - z^{(0)}_{(\cdot)}$. Then, the reconstructed variables satisfy the quiescent equilibrium condition in (23).*

*Proof.* We will assume that the initial data is in equilibrium state. Therefore, the cell averages $\overline{hu}_{ij} = \overline{hv}_{ij} = 0$ and $\overline{\eta}_{ij} = \overline{h}_{ij} + \overline{z}_{ij} \equiv const$. Then, the WENO reconstruction of such quantities will lead to $hu^{(0)}_{(\cdot)} = hv^{(0)}_{(\cdot)} = 0$ and $\eta^{(0)}_{(\cdot)} \equiv const$ according to the properties of the WENO reconstruction. The water depth, reconstructed as $h^{(0)}_{(\cdot)} = \eta^{(0)}_{(\cdot)} - z^{(0)}_{(\cdot)}$, will satisfy (23), as $\delta \left(h^{(0)} + z^{(0)}\right)_{(\cdot,\cdot)} = \delta \left(\eta^{(0)} - z^{(0)} + z^{(0)}\right)_{(\cdot,\cdot)}$, which yields $\delta \left(\eta^{(0)}\right)_{(\cdot,\cdot)} = 0$ □

The keystone for the design of a well balanced scheme is to ensure that the cell fluctuations, which are discrete differences between flux variations and source contributions between two different points of the grid, are canceled out. In other words, the discrete flux variation between two points must coincide with the approximation of the integral of the source term along the path defined by those points. To this end, a particular discretization of the integral of the source term that fulfills the aforementioned requirement must be found. Such discretization will be derived by decomposing cell fluctuations in the Cartesian directions and locally imposing the equilibrium condition to each of them.

To design a well-balanced scheme, the fluctuation form of the scheme in (13) is more suitable. This formulation reads

$$\overline{\mathbf{U}}^{n+1}_{ij} = \overline{\mathbf{U}}^{n}_{ij} - \frac{\Delta t}{\Delta x^2} \left( \sum_{r=1}^{4} \delta \mathbf{M}^{-}_{r} + \delta \mathbf{M}_{ij} \right), \tag{24}$$

where $\delta \mathbf{M}^{-}_{r}$ are the contribution of the incoming waves at cell interfaces and $\delta \mathbf{M}_{ij}$ is the centered fluctuation, that accounts for the variation of physical fluxes and source terms inside the cell. To construct

a well-balanced scheme, it is required that all fluctuations become nil under quiescent conditions, that is $\delta\mathbf{M}_r^- = \delta\mathbf{M}_{ij} = 0$.

Centered fluctuations read

$$\delta\mathbf{M}_{ij} = \sum_{r=1}^{4} \frac{\Delta x}{2} \sum_{q=1}^{k} w_q \mathcal{F}_{r,q} - \bar{\mathbf{S}}_{ij} \tag{25}$$

where

$$\bar{\mathbf{S}}_{ij} \approx \begin{pmatrix} 0 \\ \dfrac{1}{\Delta t}\displaystyle\int_0^{\Delta t}\int_{x_{i+1/2}}^{x_{i-1/2}}\int_{y_{i+1/2}}^{y_{i-1/2}} S_{ij}^x(x,y,\tau)\, dydx\, d\tau \\ \dfrac{1}{\Delta t}\displaystyle\int_0^{\Delta t}\int_{x_{i+1/2}}^{x_{i-1/2}}\int_{y_{i+1/2}}^{y_{i-1/2}} S_{ij}^y(x,y,\tau)\, dydx\, d\tau \end{pmatrix} \tag{26}$$

is the approximation of the integral of the source term inside the cell. Note that the $q$ loop is done over quadrature points. In this case, a suitable approximation of the integral in (26) is required to ensure $\delta\mathbf{M}_{ij} = 0$.

On the other hand, upwind fluctuations are given by

$$\delta\mathbf{M}_r^- = \sum_{r=1}^{4} \frac{\Delta x}{2} \sum_{q=1}^{k} w_q \left(\mathcal{F}_{r,q}^- - \mathcal{F}_{r,q}\right) \tag{27}$$

where $\mathcal{F}_{r,q} = \mathbf{F}_{i_\eta,j,q} \cdot n_x + \mathbf{G}_{i,j_\eta,q} \cdot n_y$ in Cartesian grid, where $\eta$ can be either $L$ or $R$. For instance $\mathcal{F}_{E,q} = \mathbf{F}_{i_E,j,q}$, $\mathcal{F}_{W,q} = -\mathbf{F}_{i_W,j,q}$, $\mathcal{F}_{N,q} = \mathbf{G}_{i,j_N,q}$ and $\mathcal{F}_{S,q} = -\mathbf{G}_{i,j_S,q}$. For (27), the equilibrium condition is $\mathcal{F}_{r,q}^- = \mathcal{F}_{r,q}$. Remark that $q$ will be hereafter denoted by $\alpha$ in the $x$ direction and by $\beta$ in the $y$ direction.

Let us consider first the equilibrium condition for centered fluctuations, $\delta\mathbf{M}_{ij} = 0$. First of all, we propose to decompose (26) in each coordinate direction by performing a Gaussian integration along the transverse direction with respect to the direction of the variation of the geometric term while seeking a suitable discretization of the source term for the integral along the former direction. This means

$$\bar{\mathbf{S}}_{ij} = \begin{pmatrix} 0 \\ \dfrac{\Delta x}{2}\displaystyle\sum_{\beta=1}^{k} w_\beta \bar{S}_{ij,\beta}^x \\ \dfrac{\Delta x}{2}\displaystyle\sum_{\alpha=1}^{k} w_\alpha \bar{S}_{ij,\alpha}^y \end{pmatrix} \tag{28}$$

where

$$\begin{aligned} \bar{S}_{ij,\beta}^x &\approx \frac{1}{\Delta t}\int_0^{\Delta t}\int_{x_{i-1/2}}^{x_{i+1/2}} S_{ij}^x(x,y_\beta,\tau)\, dx\, d\tau \\ \bar{S}_{ij,\alpha}^y &\approx \frac{1}{\Delta t}\int_0^{\Delta t}\int_{y_{i-1/2}}^{y_{i+1/2}} S_{ij}^y(x_\alpha,y,\tau)\, dy\, d\tau \end{aligned} \tag{29}$$

are the sought approximations. Equation (25) can also be decomposed in each of the Cartesian directions as follows

$$
\begin{aligned}
\delta\mathbf{M}_{ij} = \; & \tfrac{\Delta x}{2}\sum_{\beta=1}^{k} w_\beta \left( \mathbf{F}_{i_E,j,\beta} - \mathbf{F}_{i_W,j,\beta} - \begin{pmatrix} 0 \\ \bar{S}^x_{ij,\beta} \\ 0 \end{pmatrix} \right) + \\
& \tfrac{\Delta x}{2}\sum_{\alpha=1}^{k} w_\alpha \left( \mathbf{G}_{i,j_N,\alpha} - \mathbf{G}_{i,j_S,\alpha} - \begin{pmatrix} 0 \\ 0 \\ \bar{S}^y_{ij,\alpha} \end{pmatrix} \right)
\end{aligned}
\tag{30}
$$

where the source term has been decomposed in each of the coordinate directions. It is worth recalling that the source term is constructed as a power series expansion in time as follows

$$
\mathbf{S}_{ij}(x,y,\tau) = \mathbf{S}_{ij}(x,y,0) + \sum_{k=1}^{K} \left[ \frac{\partial^k \mathbf{S}_{ij}}{\partial t^k} \right]_{x,y,t=0} \frac{\tau^k}{k!}\,, \tag{31}
$$

and so are the physical fluxes $\mathbf{F}$ and $\mathbf{G}$. We can then rewrite the 1D discretizations of the source term as

$$
\begin{aligned}
\bar{S}^x_{ij,\beta} &\approx \int_{x_{i-1/2}}^{x_{i+1/2}} \left( S^x_{ij}(x,y_\beta,0) + \sum_{k=1}^{K} \left[ \frac{\partial^k S^x}{\partial t^k} \right]_{x,y_\beta,t=0} \frac{\Delta t^k}{(k+1)!} \right) dx \\
\bar{S}^y_{ij,\alpha} &\approx \int_{y_{i-1/2}}^{y_{i+1/2}} \left( S^y_{ij}(x_\alpha,y,0) + \sum_{k=1}^{K} \left[ \frac{\partial^k S^y}{\partial t^k} \right]_{x_\alpha,y,t=0} \frac{\Delta t^k}{(k+1)!} \right) dy
\end{aligned}
\tag{32}
$$

or in its compact form

$$
\bar{S}^x_{ij,\beta} = \bar{S}^{x,(0)}_{ij,\beta} + \sum_{k=1}^{K} \bar{S}^{x,(k)}_{ij,\beta} \frac{\Delta t^k}{(k+1)!}\,, \quad \bar{S}^y_{ij,\alpha} = \bar{S}^{y,(0)}_{ij,\alpha} + \sum_{k=1}^{K} \bar{S}^{y,(k)}_{ij,\alpha} \frac{\Delta t^k}{(k+1)!} \tag{33}
$$

To derive the well-balanced formulation, steady conditions are considered. Hence, for the derivation of the suitable approximation of the 1D integrals of the source term in (33), we only consider the leading terms of the equations, as time derivatives vanish in the steady state. Equation (30) is rewritten for the leading term only

$$
\begin{aligned}
\delta\mathbf{M}^{(0)}_{ij} = \; & \tfrac{\Delta x}{2}\sum_{\beta=1}^{k} w_\beta \left( \mathbf{F}^{(0)}_{i_E,j,\beta} - \mathbf{F}^{(0)}_{i_W,j,\beta} - \begin{pmatrix} 0 \\ \bar{S}^{x,(0)}_{ij,\beta} \\ 0 \end{pmatrix} \right) + \\
& \tfrac{\Delta x}{2}\sum_{\alpha=1}^{k} w_\alpha \left( \mathbf{G}^{(0)}_{i,j_N,\alpha} - \mathbf{G}^{(0)}_{i,j_S,\alpha} - \begin{pmatrix} 0 \\ 0 \\ \bar{S}^{y,(0)}_{ij,\beta} \end{pmatrix} \right)
\end{aligned}
\tag{34}
$$

from which we notice the one-dimensional conditions

$$
\mathbf{F}^{(0)}_{i_E,j,\beta} - \mathbf{F}^{(0)}_{i_W,j,\beta} - \begin{pmatrix} 0 \\ \bar{S}^{x,(0)}_{ij,\beta} \\ 0 \end{pmatrix} = 0 \qquad \forall \beta = 1,...,k \tag{35}
$$

$$
\mathbf{G}^{(0)}_{i,j_N,\alpha} - \mathbf{G}^{(0)}_{i,j_S,\alpha} - \begin{pmatrix} 0 \\ 0 \\ \bar{S}^{y,(0)}_{ij,\alpha} \end{pmatrix} = 0 \qquad \forall \alpha = 1,...,k \tag{36}
$$

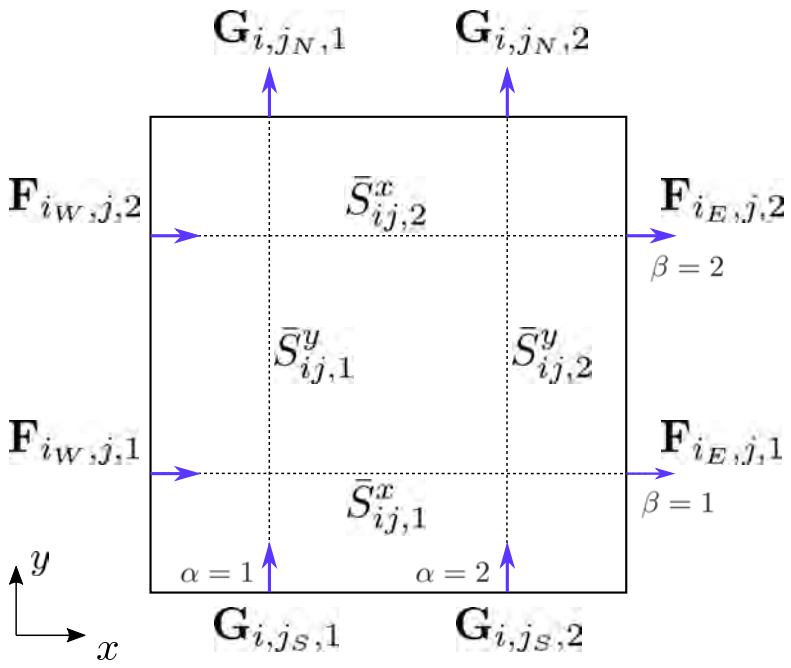


Figure 1: Representation of the Gaussian integration points, the 1D sub-cell integration lines (dashed line) in the Cartesian directions and the the physical fluxes at cell interfaces (blue vectors) for the case of $k = 2$. Integration lines are identified using the notation for the quadrature points $\alpha$ and $\beta$.

The notation for the 1D balancing of fluxes and sources in Equation (35) and (36) is depicted in Figure 1 for the case of $k = 2$. In this figure, the 1D integration lines in the Cartesian directions are depicted using dashed lines and the physical fluxes at both sides of the integration line, as well as the integral of the source term along it, are also represented. It is shown that such 1D integration lines are defined by the quadrature points at cell interfaces. Let us consider now the condition in the $x$-direction in Equation (35) (the condition in $y$ is the same due to the rotational invariance). In what follows, subscripts $j$ and $\beta$ are omitted for the sake of clarity, assuming a pure 1D problem. Under quiescent steady conditions, we have that $\mathbf{F}(\mathbf{U}) = (0, \frac{1}{2}gh^2, 0)^T$ yielding

$$(g\bar{h}\delta h)^{(0)}_{i_E,i_W} - \bar{S}^{x,(0)}_i = 0 \tag{37}$$

If approaching the integral of the source term as

$$\bar{S}^{x,(0)}_i = -g\bar{h}\delta z^{(0)}_{i_W,i_E}\,, \tag{38}$$

it yields $g\bar{h}\delta(h+z)^{(0)}_{i_W,i_E} = 0$, where one notices that $(h+z)^{(0)}_{(\cdot)} = \eta^{(0)}_{(\cdot)}$ is the equilibrium reconstruction variable and therefore $g\bar{h}\delta\eta^{(0)}_{i_W,i_E} = 0$ is satisfied with machine precision.

The discretization of the source term used above, though well-balanced, is only 2-nd order accurate in space. To obtain a $K+1$-th order scheme, it is necessary to extend this integration technique to arbitrary order in space. To this end, we can use Romberg integration, which is a result that can be obtained from Richardson's extrapolation. The main role of Romberg integration is to extend the well-balanced integration in Equation (38) to arbitrary order, without loosing the properties of Equation (38) for reproducing the discrete equilibrium. Note that traditional quadrature rules are not able to maintain the discrete equilibrium.

An arbitrary order integral of the source term is denoted as $\left\{\bar{S}^{x,(0)}_i\right\}^n_m$, where $m$ is the number of subdivisions of the initial interval $\Delta x = [x_{i-1/2}, x_{i+1/2}]$, with step size $h_m = \frac{\Delta x}{2^m}$ and $n$ the number

of Romberg iterations, having a magnitude of the residual of $\mathcal{O}(\Delta x^{2(n+1)})$. To construct a $K+1$-th order ADER scheme, both $n$ and $m$ will take values up to $\lceil \frac{K-1}{2} \rceil$ and the order of accuracy of the method will be either $\mathcal{O}(\Delta x^{K+2})$ if $K$ is even or $\mathcal{O}(\Delta x^{K+1})$ if $K$ is odd. The expression for $\left\{ \bar{S}_i^{x,(0)} \right\}_m^n$ is computed recursively departing from the trapezoid integrals, that is $n=0$ and $m=0,...,\lceil \frac{K-1}{2} \rceil$, and computing the following levels $n=1,...,\lceil \frac{K-1}{2} \rceil$ as

$$\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n+1} = \frac{4^{n+1} \left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n} - \left\{ \bar{S}_i^{x,(0)} \right\}_{m}^{n}}{4^{n+1}-1}. \tag{39}$$

For $n=0$, the integrals are given by $\left\{ \bar{S}_i^{x,(0)} \right\}_0^0 = (-g\bar{h}\delta z)_{i_E,i_W}^{(0)}$ and $\left\{ \bar{S}_i^{x,(0)} \right\}_1^0 = (-g\bar{h}\delta z)_{i_E,i}^{(0)} + (-g\bar{h}\delta z)_{i,i_W}^{(0)}$, for $m=0$ and $m=1$ respectively.

**Theorem 2.** *The extrapolating function $\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n+1}$ in Equation (39) posses the following property: if the two arguments of this function are equal, the function yields the argument of the function, as follows*

$$\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n+1} \left( \left\{ \bar{S}_i^{x,(0)} \right\}_m^n, \left\{ \bar{S}_i^{x,(0)} \right\}_m^n \right) = \left\{ \bar{S}_i^{x,(0)} \right\}_m^n \tag{40}$$

*Proof.* Let us consider that $\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n} = \left\{ \bar{S}_i^{x,(0)} \right\}_m^n$, then (39) becomes

$$\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n+1} = \frac{(4^{n+1}-1) \left\{ \bar{S}_i^{x,(0)} \right\}_m^n}{4^{n+1}-1} \tag{41}$$

which yields $\left\{ \bar{S}_i^{x,(0)} \right\}_{m+1}^{n+1} = \left\{ \bar{S}_i^{x,(0)} \right\}_m^n$.

□

**Theorem 3.** *Let us consider that the initial integrals for the iterative extrapolation in Equation (39) are equal and read $\left\{ \bar{S}_i^{x,(0)} \right\}_m^0 = \left\{ \bar{S}_i^{x,(0)} \right\}_0^0 \forall m = 1, \lceil \frac{K-1}{2} \rceil$. Then, all the higher order iterations $n=1,...,\lceil \frac{K-1}{2} \rceil$ will yield $\left\{ \bar{S}_i^{x,(0)} \right\}_m^n = \left\{ \bar{S}_i^{x,(0)} \right\}_0^0$.*

*Proof.* The integrals are computed iteratively from a set of initial integrals $\left\{ \bar{S}_i^{x,(0)} \right\}_m^0$. If all the initial integrals (at $n=0$) are equal, all the approximations for $m=1,...,\lceil \frac{K-1}{2} \rceil$ at $n=1$ will be equal according to the property in Theorem 2, and so on for $n=1,...,\lceil \frac{K-1}{2} \rceil$.

□

**Theorem 4.** *Let us consider that the initial piecewise constant data for the problem satisfy the quiescent equilibrium condition in (23) and that the integral in Equation (38) is used in combination with the reconstruction procedure in Theorem 1. Then, all the initial approximations of the integrals $\left\{ \bar{S}_{ij,\beta}^{x,(0)} \right\}_m^0$ are equal to $\left\{ \bar{S}_{ij,\beta}^{x,(0)} \right\}_0^0 = -g\bar{h}\delta z_{i_W,i_E}^{(0)}$*

*Proof.* Let us consider that the interval $\Delta x = [x_{i_W}, x_{i_E}]$ where the integral is approached is subdivided $m$ times, leading to an step size of $h_m = \Delta x / 2^m$. Inside that interval, the initial integral is approached as $\left\{ \bar{S}_i^{x,(0)} \right\}_m^0 = \sum_{p=1}^{2^m} -g\bar{h}\delta z_{p-1,p}^{(0)}$. Notice that $p=0$ and $p=2^m$ are equivalent to $i_W$ and $i_E$, as they are the extrema of the interval.

As outlined by Theorem 1, considering that the reconstructed $\eta^{(0)}_{(\cdot)} = h^{(0)}_{(\cdot)} + z^{(0)}_{(\cdot)}$ preserves the discrete equilibrium, we can substitute $\delta z^{(0)}_{p-1,p}$ by $-\delta h^{(0)}_{p-1,p}$, yielding

$$\left\{\bar{S}_i^{x,(0)}\right\}_m^0 = \sum_{p=1}^{2^m} g\bar{h}\delta h^{(0)}_{p-1,p} \tag{42}$$

which can be rewritten as

$$\left\{\bar{S}_i^{x,(0)}\right\}_m^0 = \frac{g}{2}\left(\sum_{p=1}^{2^m}\left(h_p^{(0)}\right)^2 - \sum_{p=1}^{2^m}\left(h_{p-1}^{(0)}\right)^2\right) \tag{43}$$

where $2^m - 1$ terms are canceled and yields

$$\left\{\bar{S}_i^{x,(0)}\right\}_m^0 = \frac{g}{2}\left(\left(h_p^{(0)}\right)^2 - \left(h_{p-1}^{(0)}\right)^2\right) \equiv g\bar{h}\delta h^{(0)}_{i_W,i_E} \tag{44}$$

It is straightforward to notice that $\delta z^{(0)}_{i_W,i_E} = -\delta h^{(0)}_{i_W,i_E}$, therefore $\left\{\bar{S}_i^{x,(0)}\right\}_m^0 = -g\bar{h}\delta z^{(0)}_{i_W,i_E}$. This completes the proof of the theorem.

□

Concerning the derivative terms, there is no need of a particular discretization technique of the source term to ensure the well-balanced property as time derivatives vanish under steady state. Here, we use a 2D Gaussian integration

$$\bar{S}_{ij}^{x,(k)} = \sum_{\beta=1}^{k} w_\beta \sum_{\alpha=1}^{k} w_\alpha \left(-gh^{(k)}\partial_x z\right)_{\alpha,\beta} \tag{45}$$

$$\bar{S}_{ij}^{y,(k)} = \sum_{\beta=1}^{k} w_\beta \sum_{\alpha=1}^{k} w_\alpha \left(-gh^{(k)}\partial_y z\right)_{\alpha,\beta} \tag{46}$$

where $h^{(k)}_{\alpha,\beta}$ is the $k$-th time derivative of $h$ at the quadrature point. The complete integral of the source term will be computed as

$$\bar{\mathbf{S}}_{ij} = \frac{\Delta x}{2}\begin{pmatrix} 0 \\ \sum_{\beta=1}^{k} w_\beta \left\{\bar{S}_{ij,\beta}^{x,(0)}\right\}_m^n + \sum_{k=1}^{K} \bar{S}_{ij}^{x,(k)} \dfrac{\Delta t^k}{(k+1)!} \\ \sum_{\alpha=1}^{k} w_\alpha \left\{\bar{S}_{ij,\alpha}^{y,(0)}\right\}_m^n + \sum_{k=1}^{K} \bar{S}_{ij}^{y,(k)} \dfrac{\Delta t^k}{(k+1)!} \end{pmatrix} \tag{47}$$

where $m = n = \left\lceil \frac{K-1}{2} \right\rceil$.

**Theorem 5.** *Let us consider that the initial piecewise constant data for the problem satisfy the quiescent equilibrium condition in (23). Then, the source term discretization in (38)–(47) in combination with the reconstruction procedure in Theorem 1 satisfies the discrete equilibrium for the centered fluctuations, that is* $\delta\mathbf{M}_{ij} = 0$.

*Proof.* Our goal is to prove that (30) equals zero for any formal order of accuracy of (47). For the sake of brevity, the proof will be done for the second component of (30), since the third component will be analogous and the first component is already zero.

We will assume the result from Theorem 1, that is, the reconstructed $\eta^{(\cdot)}_{(0)}$ is constant and $hu^{(0)}_{(\cdot)} = hv^{(0)}_{(\cdot)} = 0$. We will also assume that the particular derivation of time derivatives using the CK procedure yields nil derivatives when departing from equilibrium data (see [28] for more details). With the latter assumption, we can consider all time derivatives nil, hence the higher order terms of the power series expansion of the fluxes are zero, those of (47) are also zero and the equilibrium condition can be simply stated for (34), instead of (30). Considering the $x$-component of (34) (second component) and omitting subscripts $j$ and $\beta$, the equilibrium condition at any arbitrary formal order of accuracy yields

$$\mathbf{F}^{(0)}_{i_E} - \mathbf{F}^{(0)}_{i_W} - \begin{pmatrix} 0 \\ \left\{\bar{S}^{x,(0)}_i\right\}^n_m \\ 0 \end{pmatrix} = 0 \tag{48}$$

where $\left\{\bar{S}^{x,(0)}_i\right\}^n_m$ is computed using (38) and (39), and $\mathbf{F}(\mathbf{U}) = (0, \frac{1}{2}gh^2, 0)^T$ under steady conditions. If considering $m = n = 0$, we have that

$$\left(\frac{1}{2}gh^2\right)^{(0)}_{i_E} - \left(\frac{1}{2}gh^2\right)^{(0)}_{i_W} - \left\{\bar{S}^{x,(0)}_i\right\}^0_0 = 0 \tag{49}$$

When choosing $\left\{\bar{S}^{x,(0)}_i\right\}^0_0 = -g\bar{h}\delta z^{(0)}_{i_W,i_E}$, as given in (38), Equation (49) yields $g\bar{h}\delta(h+z)^{(0)}_{i_W,i_E} = 0$, where one notices that $(h+z)^{(0)}_{(\cdot)} = \eta^{(0)}_{(\cdot)}$ is the equilibrium reconstruction variable and therefore $g\bar{h}\delta\eta^{(0)}_{i_W,i_E} = 0$. Hence, Equation (49) is satisfied.

When considering a higher order extrapolation of the integral ($m, n > 0$), it is sufficient to show that $\left\{\bar{S}^{x,(0)}_{ij,\beta}\right\}^n_m = \left\{\bar{S}^{x,(0)}_{ij,\beta}\right\}^0_0 = -g\bar{h}\delta z^{(0)}_{i_W,i_E}$, for any $m$ and $n$, to confirm that (48) is satisfied. The proof for the aforemetioned statement is given by Theorems 2, 3 and 4. This completes the proof of this theorem.

□

On the other hand, when considering the upwind fluctuations in (27), it is straightforward to prove that the following approximation for the leading term of the integral of the source term in (A.12)

$$\bar{S}^{x,(0)}_{i+1/2} = \left(-g\bar{h}\delta z\right)^{(0)}_{i+1/2}\,, \tag{50}$$

$$\bar{S}^{y,(0)}_{i+1/2} = 0\,, \tag{51}$$

satisfies the steady state equilibrium condition $\left(\widetilde{\lambda}^{-}\alpha^{(0)} - \beta^{-,(0)}\right)^m_{i+1/2} = 0$ in Equation (A.7), yielding $\mathbf{F}^{-}_{i+1/2,j,\beta} = \mathbf{F}_{i_E,j,\beta}$. Higher order terms are computed as

$$\bar{\mathbf{S}}^{x,(k)}_{i+1/2} = \left(-g\bar{h}^{(k)}\delta z^{(0)}\right)_{i+1/2}\,. \tag{52}$$

Note that the approximation of the integral of the source term for the $y$-oriented DRPs is analogous and the same properties are satisfied.

**Theorem 6.** *Let us consider that the initial piecewise constant data for the problem satisfy the quiescent equilibrium condition in (23). Then, the source term discretization in (50)–(52) in combination with the reconstruction procedure in Theorem 1 satisfies the discrete equilibrium for the upwind fluctuations, that is* $\delta\mathbf{M}^{-}_{r} = 0$.

*Proof.* To prove this theorem, it is sufficient to show that at cell interfaces, the numerical fluxes coincide with the physical fluxes evaluated at such point. According to Equation (27), when the aforementioned hypothesis is satisfied, the equality $\delta\mathbf{M}_r^- = 0$ holds.

For instance, if considering the $x$ flux on the right hand side wall, the equality $\mathbf{F}^-_{i+1/2,j,\beta} = \mathbf{F}_{i_E,j,\beta}$ is sought. In the equilibrium state, the derivative terms for the flux $\mathbf{F}^-_{i+1/2,j,\beta}$ in (17) vanish. Therefore, it is sufficient to prove that $\mathbf{F}^{-,(0)}_{i+1/2,j,\beta} = \mathbf{F}_{i_E,j,\beta}$ holds. According to the properties of the ARoe solver, detailed in [33], it is shown that if using the source term discretization in (50), the numerical fluxes evaluated at the equilibrium state yield $\mathbf{F}^{-,(0)}_{i+1/2}(\mathbf{U}^{(0)}_{i_E}, \mathbf{U}^{(0)}_{(i+1)_W}, \bar{S}^{x,(0)}_{i+1/2}) = \mathbf{F}(\mathbf{U}^{(0)}_{i_E})$. The same result can be derived for the numerical fluxes at the other cell walls.

□

### *4.2. Formulation of a well-balanced scheme for the SWE in the rotating frame*

We now consider the SWE with bed elevation in a rotating frame by means of including both the bed slope and the Coriolis source terms, $\mathbf{S}_b$ and $\mathbf{S}_c$ respectively. Equations (19)–(21) represent a good model for large scale phenomena in geophysical flows, in which oceanic and atmospheric circulations are often perturbations of the so-called geostrophic equilibrium [59].

When designing a numerical scheme for the resolution of a particular system of equations, it is of importance to design the scheme in such a way that allows the preservation of the steady-state equilibrium solutions, since many phenomena of interest are often perturbations of those equilibrium states. It is worth recalling that when the Coriolis effect is neglected, the SWE in (19)–(21) satisfies the quiescent equilibrium steady state in (22). Numerical schemes satisfying (22) in the discrete form are called well-balanced schemes.

When considering the Coriolis source term, equilibrium states become more complex as they now include the circulation of the flow in particular directions. For the system in (19)–(21), the most relevant equilibrium solution to be considered in the design of the numerical scheme is the so-called geostrophic equilibrium state, which arises from the balance of the Coriolis force with the hydrostatic pressure change due to the surface elevation gradient. This steady state is often referred to as *jet in the rotating frame* [59]. According to [59], the geostrophic equilibrium satisfies

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0\,, \quad g\frac{\partial}{\partial x}(h+z) = fv\,, \quad g\frac{\partial}{\partial y}(h+z) = -fu\,, \tag{53}$$

which can be rewritten as

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0\,, \quad g\frac{\partial}{\partial x}(h+z-V) = 0\,, \quad g\frac{\partial}{\partial y}(h+z+U) = 0\,, \tag{54}$$

with $\partial_x V = fv/g$ and $\partial_y U = fu/g$ the primitive functions of the Coriolis force, also called *apparent topography* [62, 59]. For the sake of simplicity, we will define the potentials

$$L = h + z - V\,, \quad K = h + z + U\,, \tag{55}$$

which are functionals that are conserved in the geostrophic equilibrium. We can identify two particular jets in the rotating frame [59], which satisfy

$$u = 0 \quad \frac{\partial v}{\partial y} = 0\,, \quad \frac{\partial h}{\partial y} = 0\,, \quad \frac{\partial z}{\partial y} = 0\,, \quad L \equiv constant\,, \tag{56}$$

and

$$v = 0 \quad \frac{\partial u}{\partial x} = 0\,, \quad \frac{\partial h}{\partial x} = 0\,, \quad \frac{\partial z}{\partial x} = 0\,, \quad K \equiv constant\,, \tag{57}$$

In this section, a well-balanced WENO-ADER scheme using the Augmented Roe solver, which preserves the jets in (56) and (57), is proposed. The keystone of this scheme is the treatment of the Coriolis source terms as geometric sources in order to discretize them in the same way than the bed elevation source term. As done for the SWE with variable bed elevation and no Corolis terms, it is necessary to identify first which quantities the reconstruction procedure will be applied to and which other quantities will be computed from the reconstructed data in order to satisfy the discrete equilibrium. When Coriolis was not considered, the reconstruction technique was applied to $hu$, $hv$, $z$ and $h+z$ (the equilibrium variable) and then, the water depth $h$ was computed from the reconstructions. When Coriolis forces are present, the equilibrium variable is not anymore $h+z$ and instead, the exact conservation must be ensured for both $K$ and $L$. In order to satisfy this, the WENO reconstruction will be carried out for $hu$, $hv$, $h$, $z$, $L$ and $K$ first, and then, $V$ and $U$ will be computed from the reconstructed data as

$$V_{(\cdot)}^{(0)} = h_{(\cdot)}^{(0)} + z_{(\cdot)}^{(0)} - L_{(\cdot)}^{(0)}\,, \qquad U_{(\cdot)}^{(0)} = K_{(\cdot)}^{(0)} - h_{(\cdot)}^{(0)} - z_{(\cdot)}^{(0)} \tag{58}$$

where $h_{(\cdot)}^{(0)}$, $z_{(\cdot)}^{(0)}$, $L_{(\cdot)}^{(0)}$ and $K_{(\cdot)}^{(0)}$ are the reconstructed water depth, bottom elevation, $L$ potential and $K$ potential, and $V_{(\cdot)}^{(0)}$ and $U_{(\cdot)}^{(0)}$ the computed Coriolis primitive variables. Note that the proposed approach requires to perform $2\times$ more reconstructions than a traditional SWE model. With WENO limiting, this may be a large additional expense which is worth accepting as the well-balanced property is totally required to reproduce most practical cases.

When using WENO-ADER schemes, the problem data is discretized in the form of cell averages, which are required in the first step of the scheme to carry out the WENO reconstructions. The discretization of $h$, $z$, $hu$ and $hv$ as cell averages is straightforward, however, for $V$ and $U$ is not that simple and has to be done in a more elaborate way. According to the definitions of $V$ and $U$, we can express them as the following integrals

$$V(x,y) = \int_0^x \frac{fv(\chi,y)}{g} d\chi + V(0)\,, \quad U(x,y) = \int_0^y \frac{fu(x,\chi)}{g} d\chi + U(0) \tag{59}$$

where $V(0) = U(0) = 0$. When considering piecewise constant data in a Cartesian grid with cell size $\Delta x$, we can compute $V$ and $U$ at cell interfaces as

$$V_{i+1/2,j} = \sum_{t=0}^{i} \left(\frac{f\overline{v}}{g}\right)_{tj} \Delta x\,, \quad U_{i,j+1/2} = \sum_{t=0}^{j} \left(\frac{f\overline{u}}{g}\right)_{it} \Delta x \tag{60}$$

and then calculate the cell averages as

$$\overline{V}_{ij} = \frac{1}{2}\left(V_{i+1/2,j} + V_{i-1/2,j}\right)\,, \quad \overline{U}_{ij} = \frac{1}{2}\left(U_{i,j+1/2} + U_{i,j-1/2}\right) \tag{61}$$

After computing the cell averages for $V$ and $U$, those for $K$ and $L$ can be computed and the reconstruction procedure for all variables can be carried out. It is worth noting that the use of the global integrals $U$ and $V$ enforces global communication in a parallelized code, which increases the computational expense.

**Theorem 7.** *Let us consider that the initial piecewise constant data for the problem satisfy the geostrophic equilibrium in (56) and (57), that variables $L_{(\cdot)}^{(0)}$, $K_{(\cdot)}^{(0)}$, $h_{(\cdot)}^{(0)}$, $hu_{(\cdot)}^{(0)}$, $hv_{(\cdot)}^{(0)}$ and $z_{(\cdot)}^{(0)}$ are reconstructed using the WENO method and that $V_{(\cdot)}^{(0)}$ and $U_{(\cdot)}^{(0)}$ are reconstructed using Equation (58). Then, the reconstructed variables satisfy the quiescent equilibrium condition in (23).*

*Proof.* We will assume that the initial data is in equilibrium state. Therefore, the cell averages $\overline{hu}_{ij} = \overline{hu}_{ij} = 0$, $\overline{L}_{ij} = \overline{h}_{ij} + \overline{z}_{ij} - \overline{V}_{ij} \equiv const$ and $\overline{K}_{ij} = \overline{h}_{ij} + \overline{z}_{ij} + \overline{U}_{ij} \equiv const$, where $\overline{U}_{ij}$ and $\overline{V}_{ij}$ are computed using (61). Then, the WENO reconstruction of such quantities will lead to $hu^{(0)}_{(\cdot)} = hv^{(0)}_{(\cdot)} = 0$, $L^{(0)}_{(\cdot)} \equiv const$ and $K^{(0)}_{(\cdot)} \equiv const$ according to the properties of the WENO reconstruction. The Coriolis primitives $V$ and $U$, reconstructed using Equation (58), will satisfy (56) and (57) as $\delta\left(h^{(0)} + z^{(0)} - V^{(0)}\right)_{(\cdot,\cdot)} = \delta\left(L^{(0)}\right)_{(\cdot,\cdot)}$ and $\delta\left(h^{(0)} + z^{(0)} + U^{(0)}\right)_{(\cdot,\cdot)} = \delta\left(K^{(0)}\right)_{(\cdot,\cdot)}$, and from the WENO reconstruction we have that $\delta\left(L^{(0)}\right)_{(\cdot,\cdot)} = \delta\left(K^{(0)}\right)_{(\cdot,\cdot)} = 0$.

□

Concerning the integration of the source terms, $\mathbf{S}_b$ and $\mathbf{S}_c$, to satisfy the well-balance property, we must follow the same approach of the previous section for the SWE with variable bed. The numerical scheme must be written in fluctuation form (24) and a particular discretization of the source term must be sought in such a way that all fluctuations, those at cell interfaces and those inside the cell, must become nil under geostrophic conditions, satisfying (56) and (57). As done in the previous section, the scheme is designed to satisfy the discrete equilibrium in the Cartesian directions and therefore the source term integration is reduced to a one-dimensional formulation. For the $x$-geostrophic balance in (56), the 0-th discretization of the source term reads

$$\bar{S}^{x,(0)}_i = -(g\bar{h}\delta z)^{(0)}_{i_W,i_E} + (g\bar{h}\delta V)^{(0)}_{i_W,i_E} \tag{62}$$

yielding $g\bar{h}\delta(h+z-V)^{(0)}_{i_W,i_E} = 0$ where $(h+z-V)^{(0)}_{(\cdot)} = L^{(0)}_{(\cdot)}$ is the equilibrium reconstruction variable and therefore $g\bar{h}\delta L^{(0)}_{i_W,i_E} = 0$ is satisfied with machine precision. To extend the 2-nd order integral in (62) to higher order of accuracy, we use the Romberg integration method detailed in the previous section. Concerning the derivative terms, there is no need of a particular discretization technique of the source term to ensure the well-balanced property, as outlined in the previous section. Here, we use a 2D Gaussian integration

$$\bar{S}^{x,(k)}_{ij} = \sum_{\alpha=1}^{k} w_\alpha \sum_{\beta=1}^{k} w_\beta \left(-gh^{(k)}\partial_x z + f(hv)^{(k)}\right)_{\alpha,\beta} \tag{63}$$

$$\bar{S}^{y,(k)}_{ij} = \sum_{\alpha=1}^{k} w_\alpha \sum_{\beta=1}^{k} w_\beta \left(-gh^{(k)}\partial_y z - f(hu)^{(k)}\right)_{\alpha,\beta} \tag{64}$$

where $h^{(k)}_{\alpha,\beta}$ and $hu^{(k)}_{\alpha,\beta}$ are the $k$-th time derivative of $h$ and $hu$ at the quadrature point.

All time derivatives should vanish under steady regime. However, this is only achieved when using a particular expression of the CK procedure for the computation of time derivatives in terms of space derivatives. The keystone to ensure that all time derivatives vanish in the geostrophic equilibrium is to use spatial derivatives of $V$ and $U$ when constructing the first derivative using the CK procedure, instead of directly using $fhv$ and $fhu$. For instance, the first derivative for the discharge in $x$ yields

$$\partial_t(hu) = -\partial_x\left(hu^2 + \frac{1}{2}gh^2\right) - \partial_y\left(huv\right) - gh\partial_x z + gh\partial_x V \tag{65}$$

noticing that under the $x$-geostrophic equilibrium in (56), it becomes

$$\partial_t(hu) = -\partial_x\left(\frac{1}{2}gh^2\right) - gh\partial_x z + gh\partial_x V \tag{66}$$

If substituting $L = h+z-V$ in Equation (66), it yields $\partial_t(hu) = -gh\partial_x L = 0$, hence the equilibrium is satisfied. Second and higher order derivatives cannot be computed using $V$ and $U$ because there is not evolution equation for such variables.

Let us consider now the upwind fluctuations in (27). As done in the previous section, the approximate integrals of the source term at cell interfaces are constructed considering variations in the direction of definition of the DRP. For an $x$-oriented DRP we have

$$\bar{S}^{x,(0)}_{i+1/2} = \left(-g\bar{h}\delta z\right)^{(0)}_{i+1/2} + \left(g\bar{h}\delta V\right)^{(0)}_{i+1/2} \tag{67}$$

$$\bar{S}^{y,(0)}_{i+1/2} = 0 \tag{68}$$

satisfying the steady state equilibrium condition. Higher order terms are computed as in (52), neglecting the contribution of the derivatives of the Coriolis term due to its non-geometric nature. When considering an $y$-oriented DRP, we have

$$\bar{S}^{x,(0)}_{j+1/2} = 0 \tag{69}$$

$$\bar{S}^{y,(0)}_{j+1/2} = \left(-g\bar{h}\delta z\right)^{(0)}_{j+1/2} - \left(g\bar{h}\delta U\right)^{(0)}_{i+1/2} \tag{70}$$

which satisfies the steady state equilibrium condition.

**Theorem 8.** *Let us consider that the initial piecewise constant data for the problem satisfy the geostrophic equilibrium condition in equilibrium in (56) and (57). Then, the source term discretizations proposed in Section 4.2 in combination with the reconstruction procedure in Theorem 7 satisfy the discrete equilibrium for both cell centered and upwind fluctuations, that is* $\delta\mathbf{M}_{ij} = 0$ *and* $\delta\mathbf{M}^{-}_{r} = 0$.

*Proof.* The proof for this theorem is similar to that for Theorems 5 and 6, if considering that the variable $z$ in the aforementioned theorems is substituted by an apparent topography, $z^x_{ap} = z - V$ and $z^y_{ap} = z + U$, in the $x$ and $y$ directions respectively.

□

## 5. Numerical results

### *5.1. Convergence test for the SWE with bed elevation*

In this section, a convergence rate test for the ARL-ADER well-balanced scheme is presented. The following initial condition is imposed

$$z(x,y) = 0.1\exp\left(-\frac{(x-50)^2+(y-50)^2}{80}\right), \qquad \forall (x,y)\in\Omega \tag{71}$$

$h(x,y,0) = 1$ m and $hu(x,y,0) = hv(x,y,0) = 0$ m$^2$/s, $\forall (x,y)\in\Omega$, with $x$ and $y$ given in m. The computational domain is $\Omega = [0,100]\times[0,100]$ m and the solution is computed at $t = 5$ s setting $CFL = 0.2$ using the 1-st and 3-rd order scheme.

Numerical errors and convergence rates for $h$ and $hu$ computed in four different grids composed of $50\times 50$, $100\times 100$, $200\times 200$ and $400\times 400$ cells are presented in Tables 1 and 2, respectively. Numerical errors have been computed using a reference solution computed by the 3-rd order scheme in a $2000\times 2000$ grid and are measured using the $L_1$, $L_2$ and $L_\infty$ error norms. In Figure 2, a logarithmic plot of the numerical errors for the water depth and discharge provided by the 1-st and 3-rd order schemes are represented against the computational time. It can be observed that the theoretical convergence rates are achieved.

| Scheme | N | $L_1$ error | Order | $L_2$ error | Order | $L_\infty$ error | Order |
|---|---|---|---|---|---|---|---|
| 1st | 50 | 1.13E-03 | | 2.16E-05 | | 8.65E-03 | |
| | 100 | 6.43E-04 | 0.81 | 1.25E-05 | 0.79 | 4.65E-03 | 0.89 |
| | 200 | 3.46E-04 | 0.90 | 6.79E-06 | 0.88 | 2.44E-03 | 0.93 |
| | 400 | 1.80E-04 | 0.94 | 3.55E-06 | 0.94 | 1.26E-03 | 0.96 |
| 3rd | 50 | 6.45E-05 | | 1.24E-06 | | 4.09E-04 | |
| | 100 | 8.69E-06 | 2.89 | 1.71E-07 | 2.86 | 5.83E-05 | 2.81 |
| | 200 | 1.10E-06 | 2.99 | 2.17E-08 | 2.98 | 7.39E-06 | 2.98 |
| | 400 | 1.36E-07 | 3.01 | 2.71E-09 | 3.00 | 9.21E-07 | 3.00 |

Table 1: Section 5.1. Convergence rate test for $h$ using $L_1$ and $L_2$ and $L_\infty$ error norms for the 1-st and 3-rd order ARL-ADER schemes. CFL=0.2.

| Scheme | N | $L_1$ error | Order | $L_2$ error | Order | $L_\infty$ error | Order |
|---|---|---|---|---|---|---|---|
| 1st | 50 | 2.17E-03 | | 4.52E-05 | | 1.86E-02 | |
| | 100 | 1.25E-03 | 0.80 | 2.67E-05 | 0.76 | 1.12E-02 | 0.73 |
| | 200 | 6.74E-04 | 0.89 | 1.47E-05 | 0.87 | 6.22E-03 | 0.85 |
| | 400 | 3.51E-04 | 0.94 | 7.70E-06 | 0.93 | 3.29E-03 | 0.92 |
| 3rd | 50 | 1.37E-04 | | 3.02E-06 | | 1.30E-03 | |
| | 100 | 1.82E-05 | 2.91 | 4.12E-07 | 2.87 | 1.80E-04 | 2.85 |
| | 200 | 2.30E-06 | 2.98 | 5.25E-08 | 2.97 | 2.35E-05 | 2.94 |
| | 400 | 2.86E-07 | 3.01 | 6.56E-09 | 3.00 | 3.00E-06 | 2.97 |

Table 2: Section 5.1. Convergence rate test for $hu$ using $L_1$ and $L_2$ and $L_\infty$ error norms for the 1-st and 3-rd order ARL-ADER schemes. CFL=0.2.

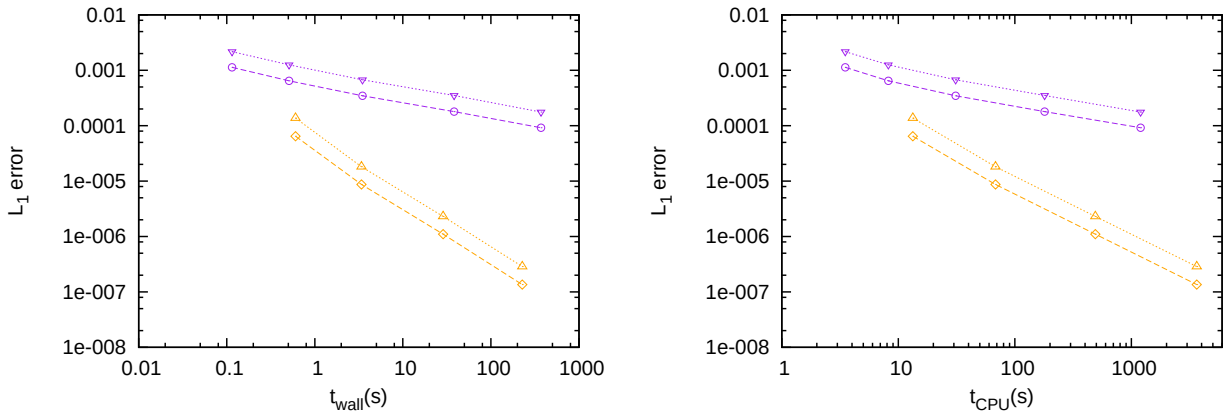


Figure 2: Section 5.1. Convergence rate test: logarithmic plot of the $L_1$ error against the wall-clock time (bottom-left) and CPU time (bottom-right). Solution computed using a 1-st (purple) and 3-rd (orange) order schemes.

### *5.2. Steady supercritical flow through a cross-section discontinuity*

The *Mach 3 wind tunnel with a step* test problem, introduced [70], has proven to be a useful test for a large number of methods and a large number of years in the framework of Euler equations. Here, we propose the analogous problem for the SWE with variable bed elevation and use it to test the numerical performance of the ADER schemes when solving complex 2D shock structures. The configuration of the case is detailed below.

We consider a computational domain given by $\Omega = [0, 70] \times [0, 25] \setminus \Pi_s$ where $\Pi_s = [15, 70] \times [0, 5]$ (units in m) is a solid obstacle which mimics the cross-section contraction and is not included in the computational domain. The boundary condition at the inlet ($x = 0$) is defined as supercritical flow setting $h = 1$ m and $hu = 10.5$ m$^2$/s, while at the outlet ($x = 75$ m), free flow condition is imposed. On the remaining boundaries, we impose solid wall conditions. The initial condition is given

by $h(x,y,0) = 1$ m, $hu(x,y,0) = 10.5$ m$^2$/s and $hv(x,y,0) = 0$ m$^2$/s $\forall (x,y) \in \Omega$.

Bed elevation is not constant in this case and is given by the following piecewise function

$$z(x,y) = \begin{cases} 0 & \text{if} \quad x \leq 5 \\ 0.3\,x^{0.3} - 0.486 & \text{if} \quad x > 5 \end{cases} \tag{72}$$

Two different grids are used in this test: a coarse grid, composed of $100 \times 280$ cells and a fine grid, composed of $400 \times 1120$ cells. The numerical solution for the water surface elevation at $t = 100$ s computed by the 1-st and 3-rd order well-balanced ARL-ADER scheme in each grid is presented in Figure 3. We have used $CFL = 0.3$.

In order to give a closer insight to the structure and features of the solution, a representation of the numerical solution of $|\nabla(h+z)|$ (top), Froude number (middle) and velocity magnitude (bottom) is presented in Figure 4. The solution is represented at the time when the Kelvin-Helmholtz instability is better observed. From such figures we can see that the initial effect of the presence of the obstacle in the supercritical flow is the formation of a bow shock (hydraulic jump) that is reflected on the top solid wall and eventually forms a Mach stem that is joined to the incident wave (bow shock) and reflected wave at the so-called triple point. It is also worth mentioning that at the triple point, a shear layer appears and a Kelvin-Helmholtz instability develops, triggered by the numerical shockwave instabilities at the Mach stem and amplified by the physical instability at the slip line.

In Figure 3, it is observed that the position of the triple point does depend on the grid and the accuracy of the numerical scheme. The coarser the grid and the lower the accuracy is, the higher the triple point is located. Moreover, we observe that the solution provided by the 3-rd order scheme in the coarse grid is of about the same accuracy than the solution provided by the 1-st order scheme in the fine grid, which has 16 times more cells. The position of the triple point/upper Mach stem proved to be a good estimation for the accuracy of the scheme and can be used to assess the convergence [71]. In Figure 5, the $x$ position of the Mach stem at $y = 24.5$ m computed by the 1-st and 3-rd order schemes is plotted against the number of cells in the $x$-direction, denoted by $N_x$, chosen in the discretization. It is observed that for the 3-rd order scheme, when discretizing the domain with more than 1000 cells in the $x$-direction, the variation in the position of the Mach stem is of the order of $10^{-2}$ m, while it is more than an order of magnitude higher for the 1-st order scheme.

### *5.3. 2D Riemann problem*

Compared with the relatively simple 1D RPs, the solution of 2D RPs include complex geometric wave patterns that pose a computational challenge for high-resolution numerical schemes. In this section, we solve a 2D RP whose initial condition is given by piecewise constant data in each of the four quadrants of the Cartesian plane. Hereafter, such quadrants will be denoted by $Q_1$, $Q_2$, $Q_3$, $Q_4$. The problem is solved using the 1-st and 3-rd order ADER scheme in a grid composed of $800 \times 800$ cells using $CFL = 0.4$ in the domain $\Omega = [0,100] \times [0,100]$ m, at a time $t = 3$ s.

The initial condition is given by

$$h(x,y) = \begin{cases} 10 & \text{if} \quad (x,y) \in Q_1 \\ 3 & \text{if} \quad (x,y) \in Q_2 \\ 0.2 & \text{if} \quad (x,y) \in Q_3 \\ 3 & \text{if} \quad (x,y) \in Q_4 \end{cases}, \; z(x,y) = \begin{cases} 1 & \text{if} \quad (x,y) \in Q_1 \\ 0 & \text{if} \quad (x,y) \in Q_2 \\ -0.5 & \text{if} \quad (x,y) \in Q_3 \\ 0 & \text{if} \quad (x,y) \in Q_4 \end{cases} \tag{73}$$

$$u(x,y) = \begin{cases} 0 & \text{if} \quad (x,y) \in Q_1 \\ 3 & \text{if} \quad (x,y) \in Q_2 \\ 3 & \text{if} \quad (x,y) \in Q_3 \\ 0 & \text{if} \quad (x,y) \in Q_4 \end{cases}, \; v(x,y) = \begin{cases} 0 & \text{if} \quad (x,y) \in Q_1 \\ 0 & \text{if} \quad (x,y) \in Q_2 \\ 3 & \text{if} \quad (x,y) \in Q_3 \\ 3 & \text{if} \quad (x,y) \in Q_4 \end{cases} \tag{74}$$

with $h$ and $z$ in m and $u$ and $v$ in m/s.

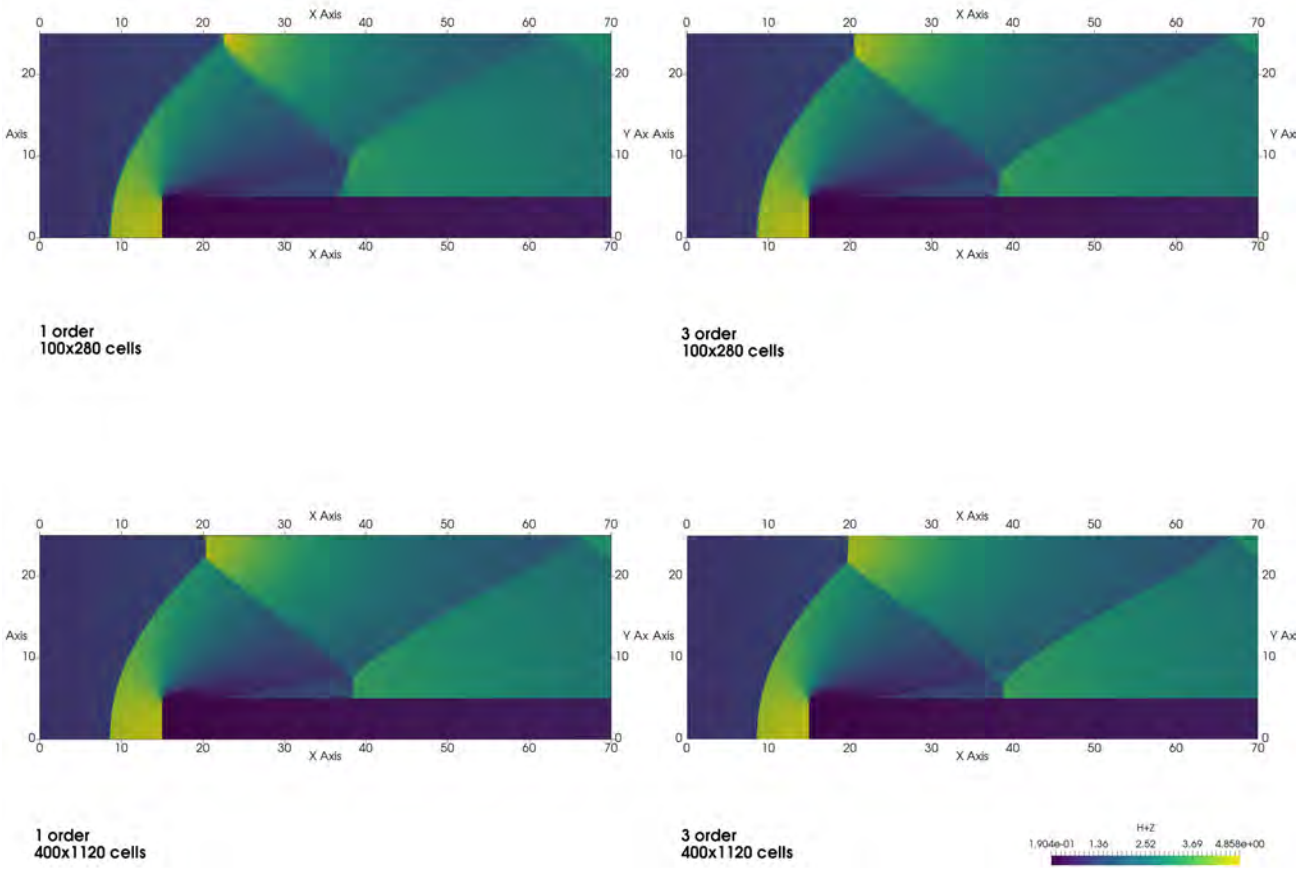


Figure 3: Section 5.2. Numerical solution for $h + z$ (m) at $t = 100$ s provided by the 1-st order (left) and 3-rd order scheme (right) using the $100 \times 280$ grid (top) and the $400 \times 1120$ grid (bottom).

The structure of the solution is composed by a complex variety of waves: at the four edges of the quadrants limiting with the axis, there are four contact discontinuities due to the bed steps. In the first quadrant, there are two interacting rarefaction waves moving in the $x$ and $y$ directions. In the second and fourth quadrants, there are two rarefaction waves moving in the $y$ and $x$ direction respectively, which interact with a transverse shock wave. Finally, the most complex wave pattern can be found in the third quadrant. Here, there are two pairs of shocks in each coordinate direction, moving perpendicularly with respect to each other, and from their interaction a jet stream pointing to the bottom-left corner of the domain is generated. Such jet is bounded by multiple shocks and a strong recirculation is observed at both sides of the jet. It is worth pointing out that this RP is a resonant problem as the number of waves is greater than the number of eigenvalues.

A shadowgraph for the numerical $h + z$ and the velocity vector field has been represented in Figure 6. A cross-sectional representation of the numerical $h + z$ and $q$ computed using a 1-st and 3-rd order scheme in a $200 \times 200$ and $800 \times 800$ grid is also presented in Figure 7 and compared with a reference solution that has been computed using the 3-rd order scheme in a very fine mesh.

It is observed that only when using the 3-rd order scheme, the finest details such as the normal shock produced by the deceleration of the reverse flow in the jet in the third quadrant, can be captured. Moreover, the velocity peak in the jet is underestimated when using the 1-st order scheme. Important

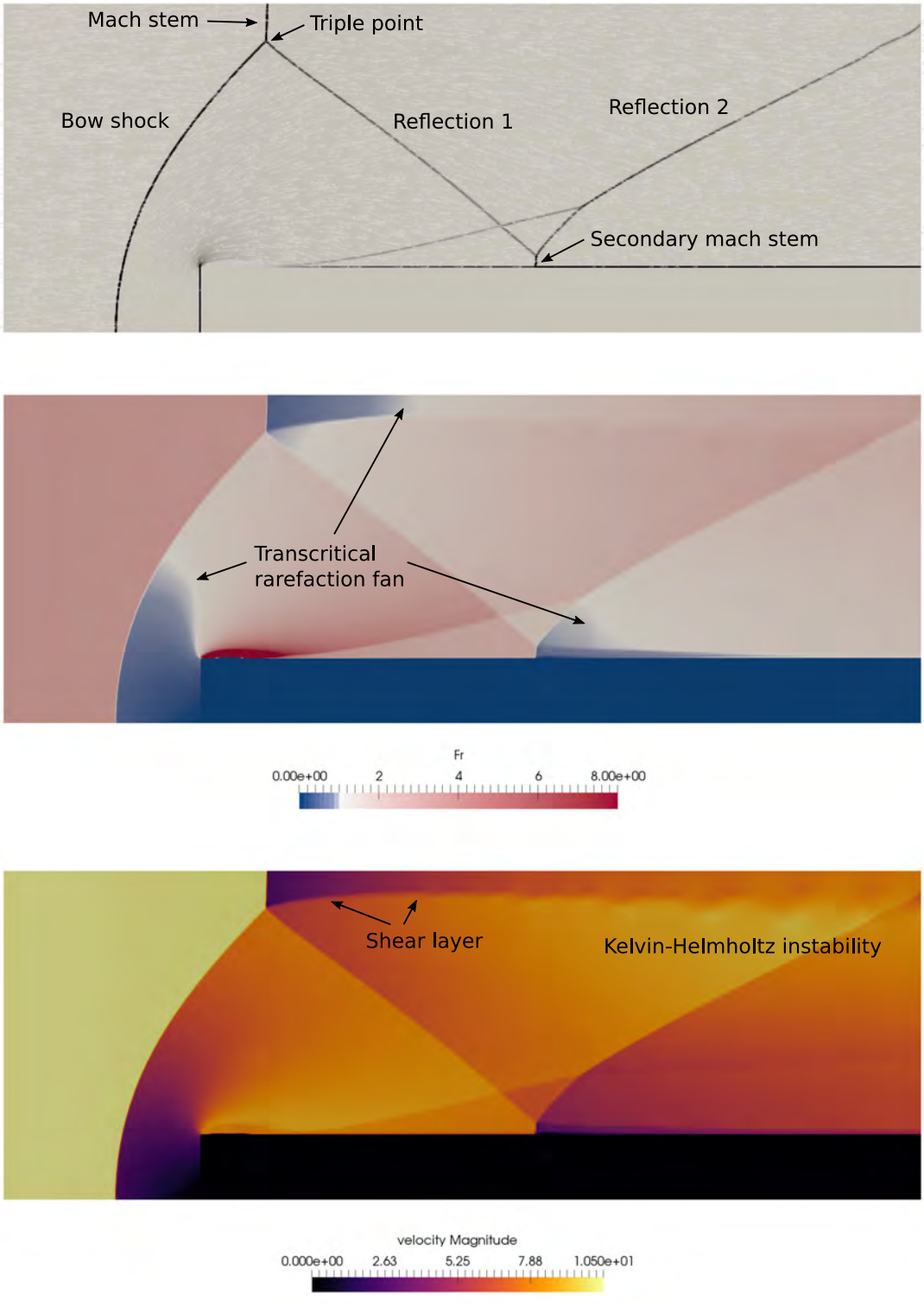


Figure 4: Section 5.2. Numerical solution for the water surface elevation gradient (top), Froude number (middle) and velocity magnitude ($m^2/s$) (bottom), including the relevant feartures of this particular flow.

differences can also be found in the first quadrant, where the solution provided by the 1-st order scheme shows larger diffusion and does not capture the correct shape of the interaction of the rarefaction fans.

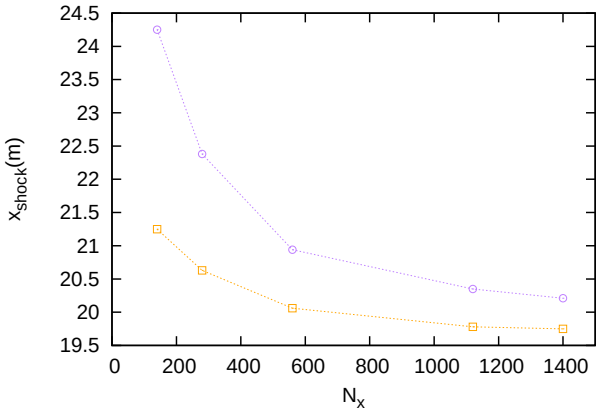


Figure 5: Section 5.2. Representation of the $x$ position of the Mach stem at $y = 24.5$ m computed by the 1-st (purple) and 3-rd order (orange) ARL-ADER schemes against the number of cells in the $x$-direction

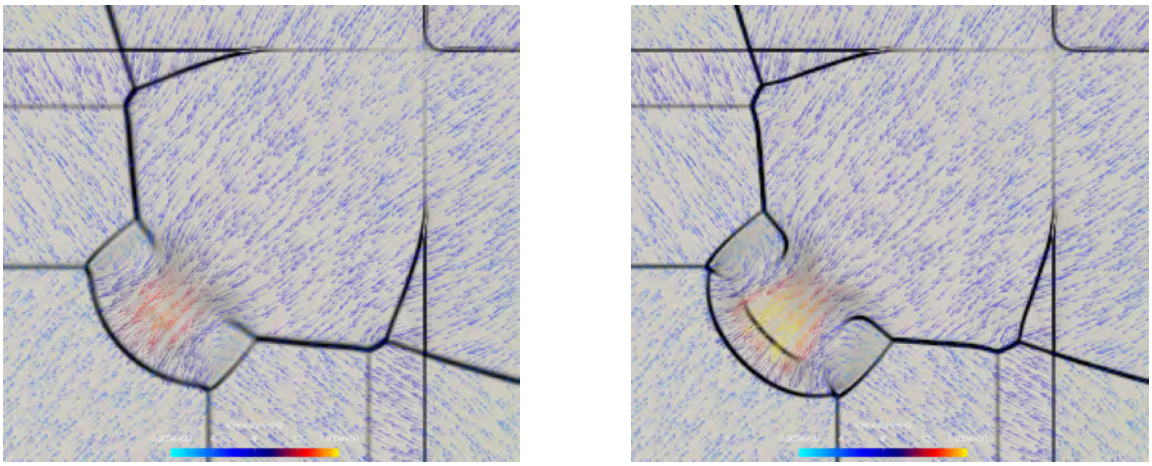

Figure 6: Section 5.3. Water surface elevation gradient and velocity field (m/s) provided by the 1-st order scheme (left) and 3-rd order scheme (right) in a $200 \times 200$ grid.

### *5.4. 1D geostrophic equilibrium with bed variation*

This test case consist of a one-dimensional flow in the $y$-direction that is initially at geostrophic equilibrium [59]. The computational domain is $y \in [-5, 5]$ m and the initial condition is given by

$$K(x, y) = 2\,, \quad u(x, y) = \frac{2g}{f} y \exp\left(-y^2\right), \quad v(x, y) = 0\,, \tag{75}$$

where $K$ is given in m, $u$ and $v$ in m$^2$/s, $f = 1$ s$^{-1}$ and $g = 9.8$ ms$^{-2}$. The bottom elevation is given by

$$z(x, y) = 0.1 \sin(0.2\pi y)\,. \tag{76}$$

The numerical solution for $h + z$, $hu$, $hv$ and $K$ at time $t = 10$ s provided by the 3-rd order ARL-ADER scheme is presented in Figure 8 and is compared with the exact solution. The numerical solution provided by a non well-balanced version of the ARL-ADER scheme is also presented in Figure 8. This non well-balanced scheme is constructed using the WENO reconstruction over $h+z$, $hu$, $hv$ and $z$, the traditional Roe solver for the fluxes, the ADER approach for the time stepping and a traditional Gaussian integration for the source term. Primitive variables are not used and the source term is only accounted for inside cells and not at cell interfaces.

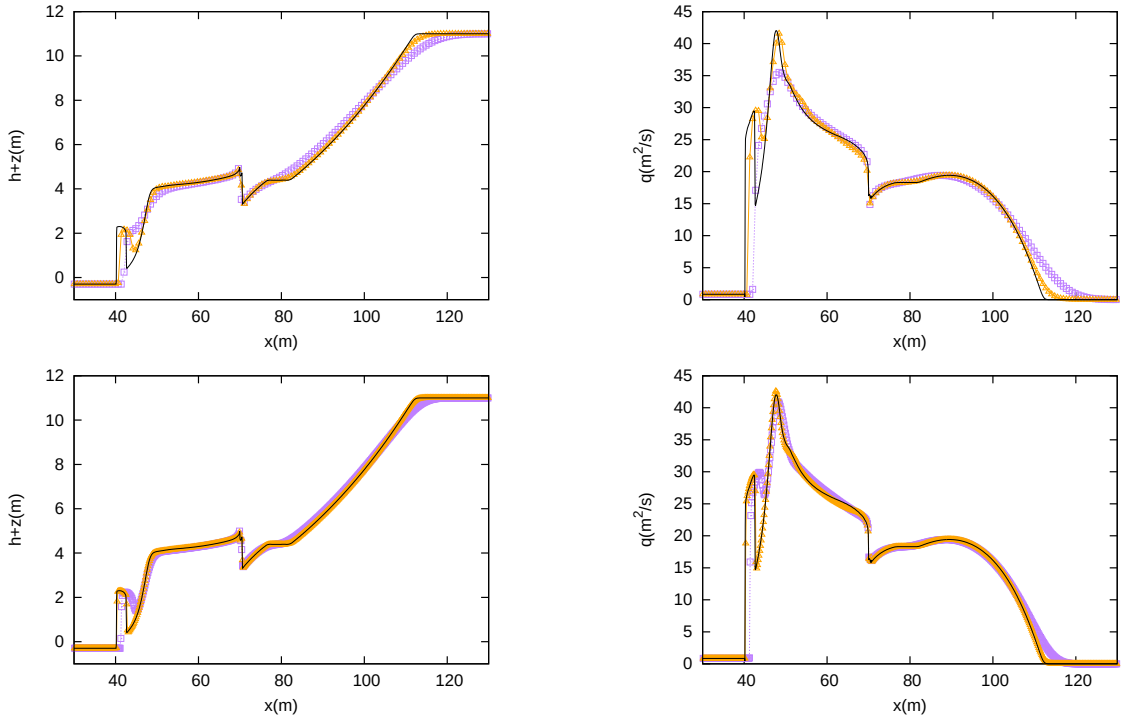


Figure 7: Section 5.3. Cross-sectional representation of the numerical $h + z$ (m) (left) and $q$ (m$^2$/s) (right) computed using a 1-st (−□−) and 3-rd order (−△−) scheme in a $200 \times 200$ (top) and $800 \times 800$ (bottom) grid.

| | 1-st order | 3-rd order |
|---|---|---|
| $t$(s) | $L_\infty$ error | $L_\infty$ error |
| 5 | 4.44088E-016 | 4.44089E-016 |
| 10 | 4.44088E-016 | 4.44089E-016 |

Table 3: Section 5.4. Numerical errors for $K$ provided by the 1-st and 3-rd order ARL-ADER scheme, measured with $L_\infty$ error norm at $t = 5$ and $t = 10$ s. Double precision is used.

According to the numerical results, the proposed ARL-ADER scheme preserves the geostrophic equilibrium up to machine precision, with round-off errors for $K$ of magnitude $10^{-16}$ m, as shown in Table 3. Hence, the scheme satisfies the well-balanced property. On the other hand, the non well-balanced ADER scheme leads to the appearance of artificial waves and is not able to preserve the geostrophic equilibrium.

In order to study the convergence properties of the non well-balanced ADER scheme, the evolution in time of the $L_1$ error norm for $K$ computed by a 1-st and 3-rd order scheme in two different grids with $\Delta x = 0.2$ m and $\Delta x = 0.1$, is depicted in Figure 9. It is observed that the numerical solution converges to the exact geostrophic state as the grid is refined and the order is increased, however, it can never achieve the level of accuracy of the well-balanced scheme.

### *5.5. Convergence test for the SWE with bed elevation and Coriolis*

A convergence rate test for the ARL-ADER well-balanced scheme with bed elevation and Coriolis is presented. The following initial condition is imposed

$$z(x,y) = 0.1 \exp\left(-\frac{(x-50)^2 + (y-50)^2}{80}\right), \qquad \forall (x,y) \in \Omega \tag{77}$$

$h(x,y,0) = 0.5$ m and $hu(x,y,0) = hv(x,y,0) = 0$ m$^2$/s $\forall (x,y) \in \Omega$, $g = 1$ m/s$^2$ and $f = 0.05$ s$^{-1}$.

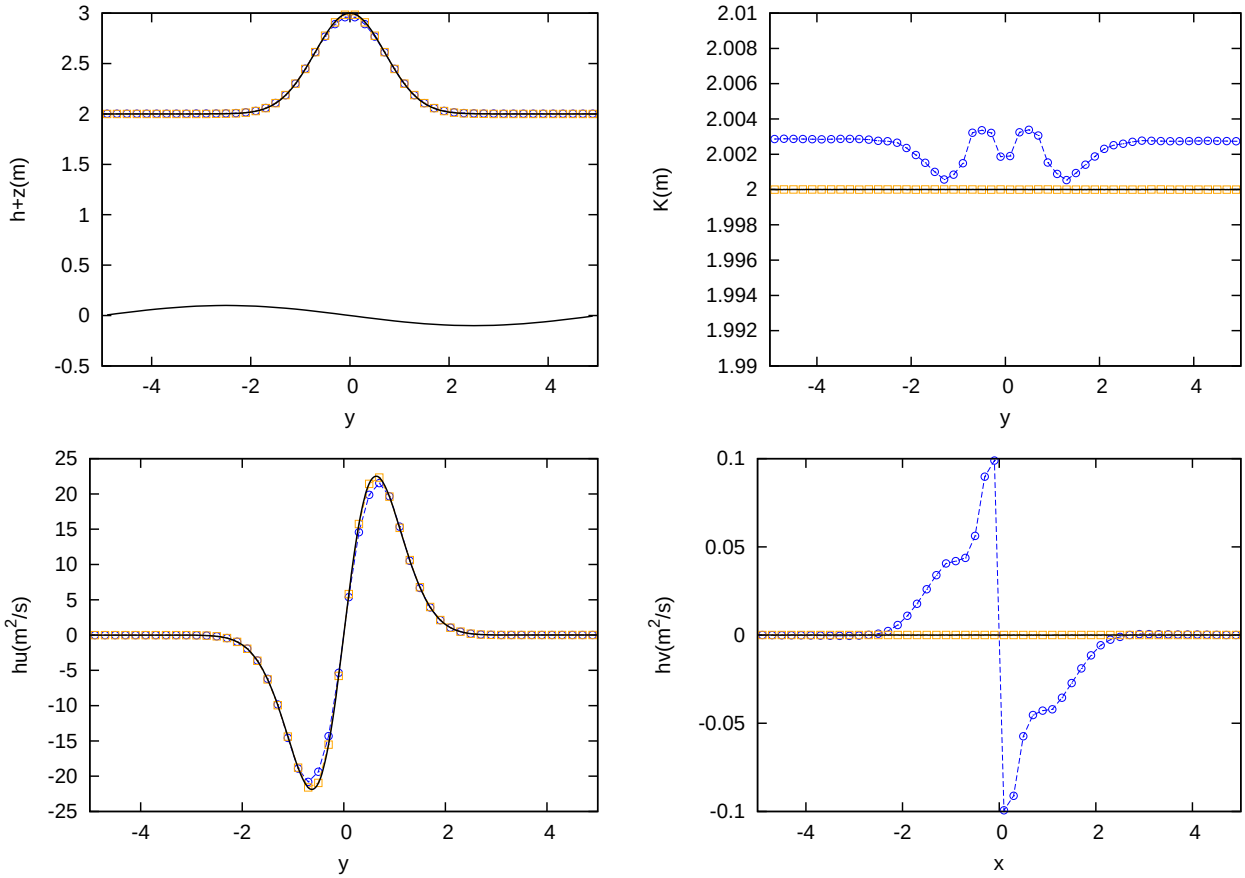


Figure 8: Section 5.4. Numerical solution for $h+z$, $K$, $hu$ and $hv$ at $t = 10$ s provided by the 3-rd order well-balanced ARL-ADER scheme (−□−) and by its non well-balanced version (− ∘ −), using $\Delta x = 0.2$ m.

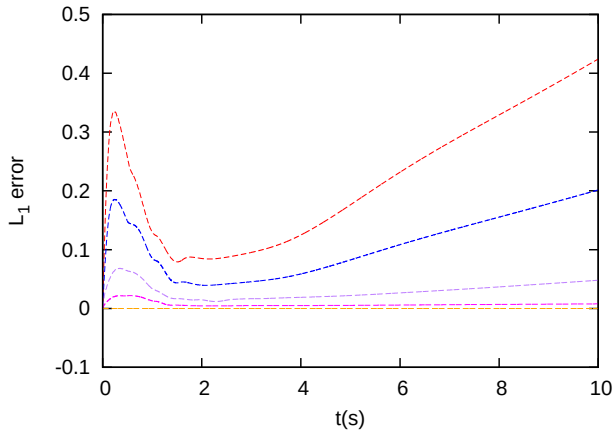


Figure 9: Section 5.4. Evolution in time of $L_1$ error norm for $K$ computed by the well-balanced ARL-ADER scheme (−) and by its non well-balanced version using a 1-st order scheme and $\Delta x = 0.2$ m (−), a 1-st order scheme and $\Delta x = 0.1$ (−), 3-rd order scheme and $\Delta x = 0.2$ (−) and 3-rd order scheme and $\Delta x = 0.1$ (−).

The computational domain is $\Omega = [0, 100] \times [0, 100]$ m and the solution is computed at $t = 3$ s setting $CFL = 0.1$ using the 3-rd order scheme.

Numerical errors and convergence rates for $h$ and $hu$ computed in fours different grids composed of $50 \times 50$, $100 \times 100$, $200 \times 200$ and $400 \times 400$ cells are presented in Table 4. Numerical errors have been

| Variable | N | $L_1$ error | Order | $L_2$ error | Order | $L_\infty$ error | Order |
|---|---|---|---|---|---|---|---|
| $h$ | 50 | 5.39E-06 | | 2.06E-07 | | 2.50E-04 | |
| | 100 | 7.33E-07 | 2.88 | 2.89E-08 | 2.83 | 3.82E-05 | 2.71 |
| | 200 | 9.51E-08 | 2.95 | 3.82E-09 | 2.92 | 5.27E-06 | 2.86 |
| | 400 | 1.21E-08 | 2.97 | 4.97E-10 | 2.94 | 7.18E-07 | 2.88 |
| $hu$ | 50 | 2.69E-06 | | 1.07E-07 | | 9.74E-05 | |
| | 100 | 3.25E-07 | 3.05 | 1.30E-08 | 3.05 | 1.23E-05 | 2.98 |
| | 200 | 3.84E-08 | 3.08 | 1.52E-09 | 3.09 | 1.42E-06 | 3.11 |
| | 400 | 4.57E-09 | 3.07 | 1.76E-10 | 3.11 | 1.68E-07 | 3.08 |

Table 4: Section 5.5. Convergence rate test for $h$ and $hu$ using $L_1$ and $L_2$ and $L_\infty$ error norms for the 3-rd order ARL-ADER scheme. CFL=0.1.

| | | $t = 0.3$ s | | | | $t = 3$ s | | | |
|---|---|---|---|---|---|---|---|---|---|
| Scheme | N | $L_1$ error | Order | $L_\infty$ error | Order | $L_1$ error | Order | $L_\infty$ error | Order |
| 3-rd | 25 | 4.26E-06 | | 2.23E-04 | | 3.73E-06 | | 5.90E-05 | |
| | 50 | 6.50E-07 | 2.71 | 3.50E-05 | 2.67 | 5.74E-07 | 2.70 | 1.04E-05 | 2.50 |
| | 100 | 8.53E-08 | 2.93 | 4.90E-06 | 2.84 | 1.41E-07 | 2.03 | 2.47E-06 | 2.07 |
| 5-th | 25 | 1.21E-06 | | 6.23E-05 | | 3.11E-06 | | 1.25E-04 | |
| | 50 | 5.78E-08 | 4.39 | 3.13E-06 | 4.31 | 1.68E-07 | 4.21 | 7.10E-06 | 4.14 |
| | 100 | 2.08E-09 | 4.80 | 1.20E-07 | 4.71 | 2.20E-08 | 2.93 | 1.10E-06 | 2.69 |

Table 5: Section 5.5. Convergence rate test for $h$ using $L_1$ and $L_\infty$ error norms for the 3-rd and 5-th order ARL-ADER scheme. The solution is computed at $t = 0.3$ and $t = 3$ s using CFL=0.02.

computed for $h$ and $hu$ using a reference solution computed by the 3-rd order scheme in a $2000 \times 2000$ grid and are measured using the $L_1$, $L_2$ and $L_\infty$ error norms. Numerical errors for $hv$ are not presented due to the symmetry of the case. The theoretical convergence rate is achieved.

### *5.5.1. Computation of stiff non-geometric source terms*

The numerical method proposed here considers a geometric reinterpretation of the Coriolis source term that allows to apply the well-balancing techniques derived for the SWE with bed elevation to the SWE in the rotating frame. To ensure the discrete equilibrium, first order time derivatives are derived expressing the source term in terms of the primitive variables. By contrast, second and higher order time derivatives are derived considering the original source terms $fhu$ and $fhv$. This yields to a suboptimal integration in time that can be evidenced when computing very stiff source terms.

If the test case in the previous section is repeated, setting $f = 2$ s$^{-1}$, a suboptimal convergence rate is observed in the solution. Table 5 shows the numerical errors provided by a 3-rd and 5-th order ARL-ADER scheme at $t = 0.3$ and $t = 3$ s, using CFL=0.02. It is observed that the convergence rates at the shorter time, $t = 0.3$ s, are optimal. However, as the solution advances in time, a degradation of accuracy is observed. This is because at shorter times, the spatial error is much higher than the temporal error, hence the overall convergence is maintained. When moving to longer times, the temporal error, which has a suboptimal behavior, grows over the spatial error and hence the overall convergence is reduced.

It is worth recalling that this test case considers a ratio between the Coriolis parameter and the gravity constant of $f/g = 2$ while in the atmospheric and oceanic circulation we usually find that $f/g = 0.00001$. Therefore, the present test case is beyond the range of application of the model, however the authors consider of importance to show the actual limitations of the scheme.

### *5.6. 2D geostrophic adjustment*

Here we consider the test case proposed in [61] (see also [59, 64]), which consists of a initial asymmetrical column of water that falls under a strong rotation that leads to a 2D geostrophic adjustment the

numerical scheme must be able to reproduce. The computational domain is $\Omega = [-10, 10] \times [-10, 10]$ m, the bottom topography is flat and the initial condition is given by

$$h(x, y) = 1 + 0.5A_0 \left( 1 - \tanh \left( \frac{\sqrt{(\sqrt{\lambda}x)^2 + (y/\sqrt{\lambda})^2} - R_i}{R_E} \right) \right) , \tag{78}$$

$$u(x, y) = v(x, y) = 0 \, , \tag{79}$$

where $A_0 = 0.5$, $\lambda = 2.5$, $R_E = 0.1$, $R_i = 1$ and the gravity and Coriolis parameters are set to $g = 1$ m/s$^2$ and $f = 1$ s$^-1$. The numerical solution is computed using the 3-rd order ARL-ADER scheme using a grid of $400 \times 400$ cells and setting CFL=0.4. Initially, the elliptical column of water is not at equilibrium and evolves in an nonaxisymetric way due to the rotation effect. Two successive shock (gravity) waves are generated and leave behind a small smooth hump that is slowly spinning clockwise. The numerical result for the water surface elevation at times $t = 0$, $t = 4$, $t = 8$, $t = 12$, $t = 16$ and $t = 20$ s are presented in Figure 10. It is observed that the expected behavior of the evolution of the solution is reproduced by the numerical scheme and that the numerical results are qualitatively very similar to those presented in [59, 64].

A cross sectional representation of the solution for $h + z$ and $L$ at $y = 10$ m and $t = 4$ s, provided by a 1-st and 3-rd order ARL-ADER scheme in two grids composed of $101 \times 101$ and $401 \times 401$ cells is depicted in Figure 11.

### *5.7. 2D propagation of Rossby waves on the equatorial $\beta$-plane*

This test case considers the propagation of a Rossby soliton on the equatorial beta-plane, for which an asymptotic solution exists to the inviscid SWE. Theoretically, the soliton should propagate to the west at fixed phase speed, without change of shape. Since the uniform propagation and shape preservation of the soliton are achieved through a fine balance between linear wave dynamics and nonlinearity, this is a good context in which to look for erroneous wave dispersion and/or numerical damping and has proven to be a good benchmark for atmosphere and ocean models (http://marine.rutgers.edu/po /index.php?model=test-problems) [72]. The interest in this test problem is to assess the spurious dispersion and dissipation effects of the numerical scheme, and how they relate to the choice of grid resolution and the accuracy of the scheme.

Long, weakly nonlinear, equatorial Rossby waves are governed by either Korteweg–de Vries (KDV) or the modified Korteweg–de Vries (MKDV) equation [72, 73]. Here, a *zero*-th order asymptotic solution to the SWE is used [73]. The initial condition for a dipole at $(x, y) = (0, 0)$ m can be found in [72, 73]. Here, the dipole is translated to $(x, y) = (72, 6)$ m by means of evaluating the initial condition at $x' = x - 72$ m and $y' = y - 6$ m. The gravity constant is set to $g = 1$ m/s$^2$ and the Coriolis parameter is calculated using the $\beta$-plane approximation $f(y) = f_0 + \beta y$ with $f_0 = 0$ s$^{-1}$ and $\beta = 1$ m$^{-1}$s$^{-1}$.

The case considered here consists of the *zero*-th order soliton described above over a flat bed, that is $z(x, y) = 0$ m, computed inside the domain $\Omega = [0, 96] \times [0, 12]$, units in m, at time $t = 120$ s. The numerical solution provided by the 1-st order (12) and 3-rd order ARL-ADER scheme (13) at times $t = 0$, $t = 30$, $t = 60$, $t = 90$ and $t = 120$ s using two grid sizes of $\Delta x = 0.2$ and $\Delta x = 0.1$ m, are presented and compared with the exact solution at $t = 120$ s. CFL is set to 0.4. It is observed that the 1-st order scheme in the coarsest grid does not perform well as it generates spurious waves. When moving to the finest grid size, the performance of the scheme is improved though it is still very diffusive and dispersive. The 3-rd order scheme does provide an accurate solution with both grids and ensures a much lower dispersion and diffusion of the soliton.

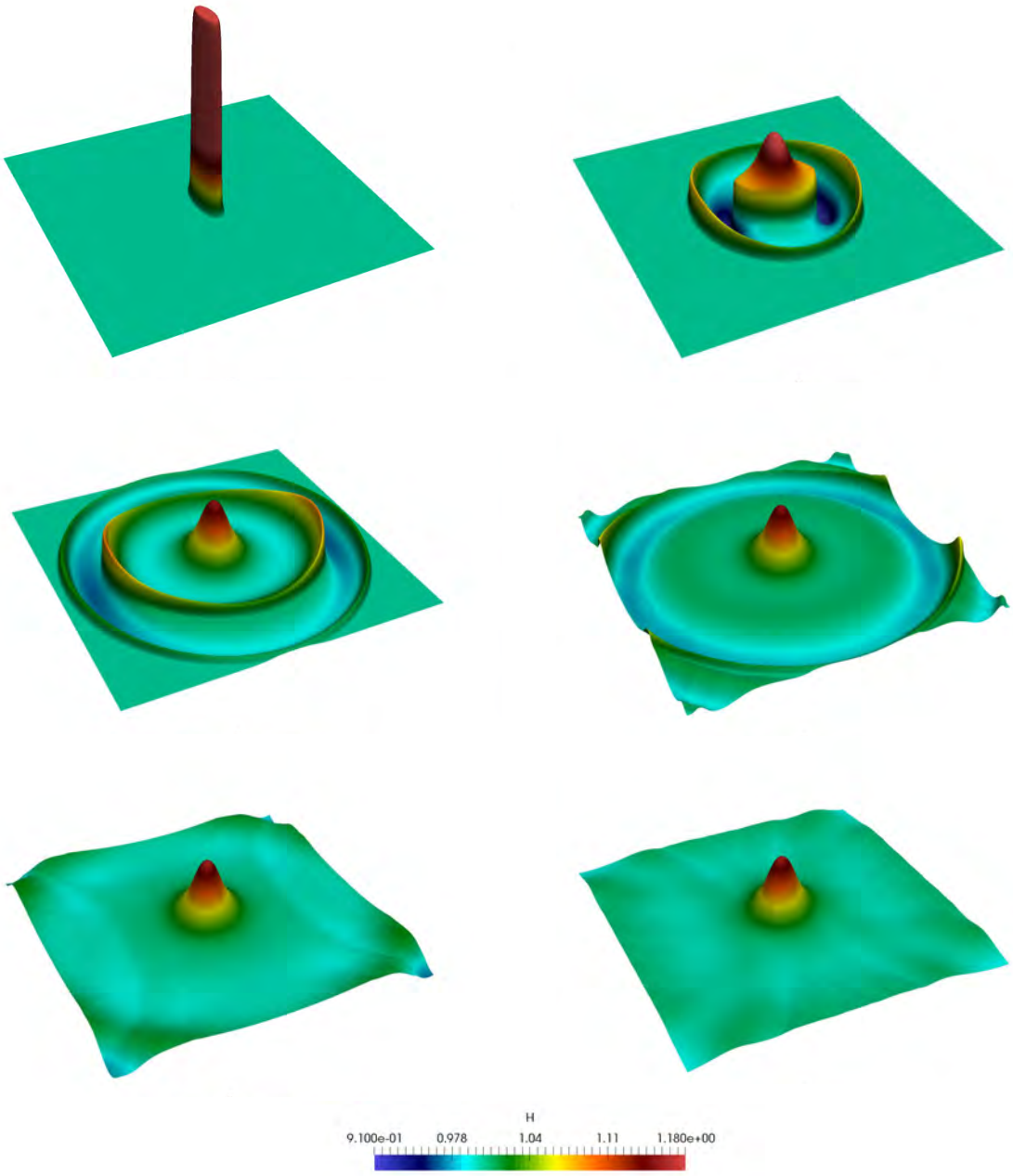


Figure 10: Section 5.6. Numerical $h+z$ (m) at $t = 0$ (top-left), $t = 4$ (top-right), $t = 8$ (middle-left), $t = 12$ (middle-right), $t = 16$ (bottom-left) and $t = 20$ s (bottom-right) provided by the 3-rd order ARL-ADER scheme in a $400 \times 400$ grid.

To assess the performance of the numerical schemes, we have used the following metrics: the damping factor, $\nu$, which accounts for the numerical damping of the solution and the relative speed, $c_r$, which accounts for the numerical dispersion of the solution, defined in [72].

It is worth noting that all maximum and minimum water depth values are cell-averaged values and no interpolation is used. Numerical values for the metrics described above and other related data is presented in Table 6 using the results provided by the 1-st and 3-rd AR-ADER scheme in two grids of $\Delta x = 0.2$ and $\Delta x = 0.1$ m. It is evidenced that the 3-rd order scheme outperforms the 1-st order scheme in terms of numerical dispersion and damping, as it was expected. If comparing with the

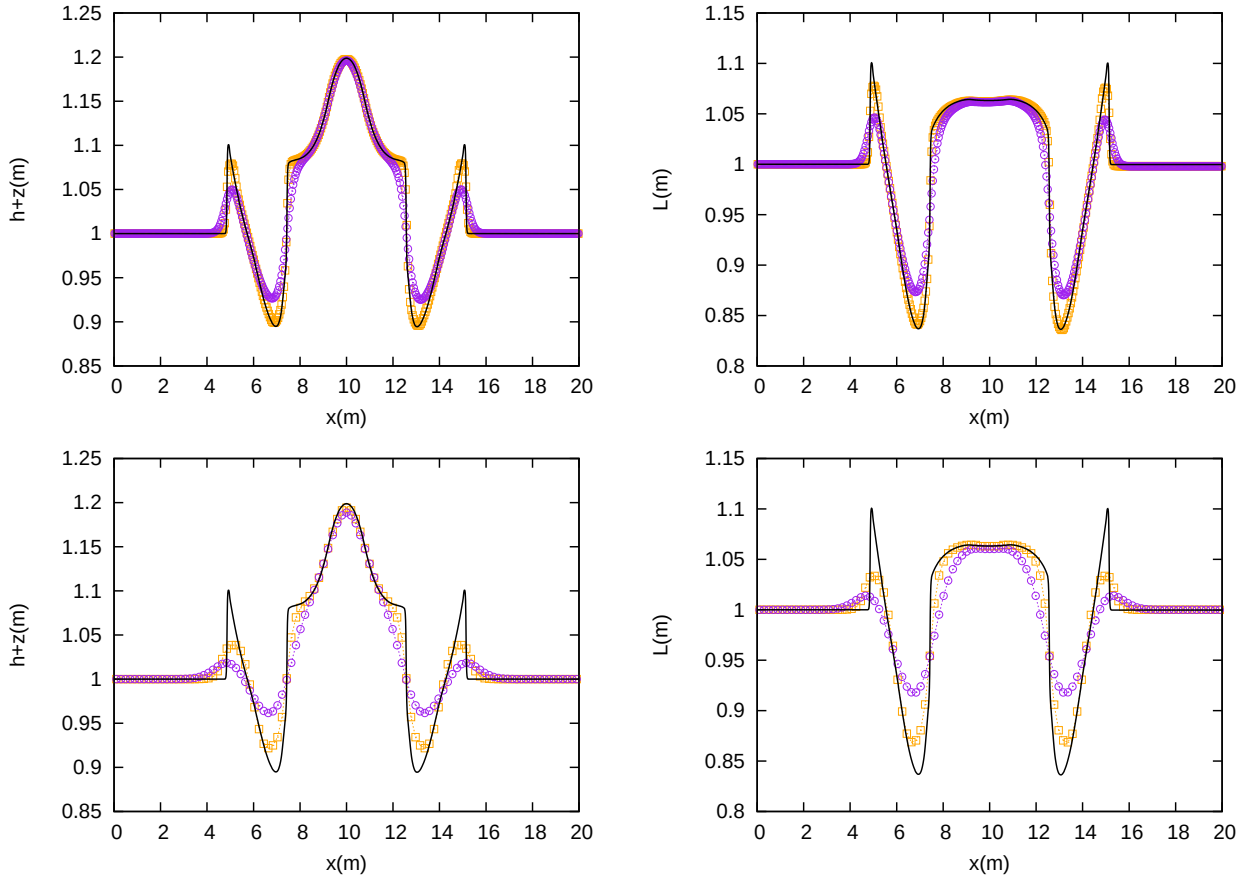


Figure 11: Section 5.6. Cross sectional representation of the solution for $h + z$ (m) and $L$ (m) at $y = 10$ (m) and $t = 4$ s provided by a 1-st and 3-rd order ARL-ADER scheme in a $101 \times 101$ and $401 \times 401$ grid.

numerical results in [72] it is observed that the measures for dispersion and damping are of the same magnitude.

Figure 14 shows a comparison between the numerical solution at $t = 30$ s provided by the 3-rd order ARL-ADER scheme and the non well-balanced version of this scheme, in a $\Delta x = 0.1$ m grid. It is observed that the non well-balanced scheme produces spurious waves arising from the imbalance of the Coriolis terms in the scheme. This result motivates the necessity of well-balanced schemes, specially for the simulation of such kind of geophysical events that consist of the evolution in time of perturbations of a certain equilibrium state. Only when the equilibrium state is accurately preserved, those perturbations will be properly captured.

### *5.8. Kelvin front generation on the equatorial $\beta$-plane*

In this section, the generation of nonlinear planetary (Rossby) and Kelvin waves at Earth's equatorial line is reproduced by the ARL-ADER numerical scheme. When the equatorial area is perturbed (by changing winds, for instance), its adjustment to the new equilibrium state is done by means of wave propagation. Such perturbations are usually of a very low frequency and therefore gravity waves are not excited, instead, only certain type of waves such as Kelvin waves, mixed waves and planetary waves (Rossby waves) appear. The short wavelength Kelvin waves carry energy eastward direction, whereas the long wavelength planetary waves carry energy to westward direction.

An additional phenomena is considered in this test case. It has been recognized for some time that nonlinear equatorial Kelvin waves can steepen and break, forming a broken wave, or front, propagating

| | 3-rd order | | 1-st order | |
|---|---|---|---|---|
| | $\Delta x = 0.2$ | $\Delta x = 0.1$ | $\Delta x = 0.2$ | $\Delta x = 0.1$ |
| $h_{max,t=0}(m)$ | 1.16948 | 1.17032 | 1.16948 | 1.17032 |
| $h_{min,t=0}(m)$ | 1.00000 | 1.00000 | 1.00000 | 1.00000 |
| $h_{max,t=120}(m)$ | 1.14019 | 1.15332 | 1.05989 | 1.08135 |
| $h_{min,t=120}(m)$ | 0.99509 | 0.99263 | $\approx 0$ | 0.99835 |
| $\nu$ | 0.975 | 0.985 | 0.906 | 0.924 |
| $x_{end}(m)$ | 24.9 | 24.84 | 26.9 | 26.2 |
| $c_r$ | 0.981 | 0.983 | 0.940 | 0.954 |
| $\text{Time}_{CPU}$ (s) | 212.8 | 1777.5 | 7.2 | 39.5 |
| $\text{Time}_{wall-clock}$ (s) | 12.55 | 100.75 | 0.95 | 6.7 |

Table 6: Section 5.7. Numerical values of the relevant metrics for the assessment of numerical dispersion and damping of the scheme.

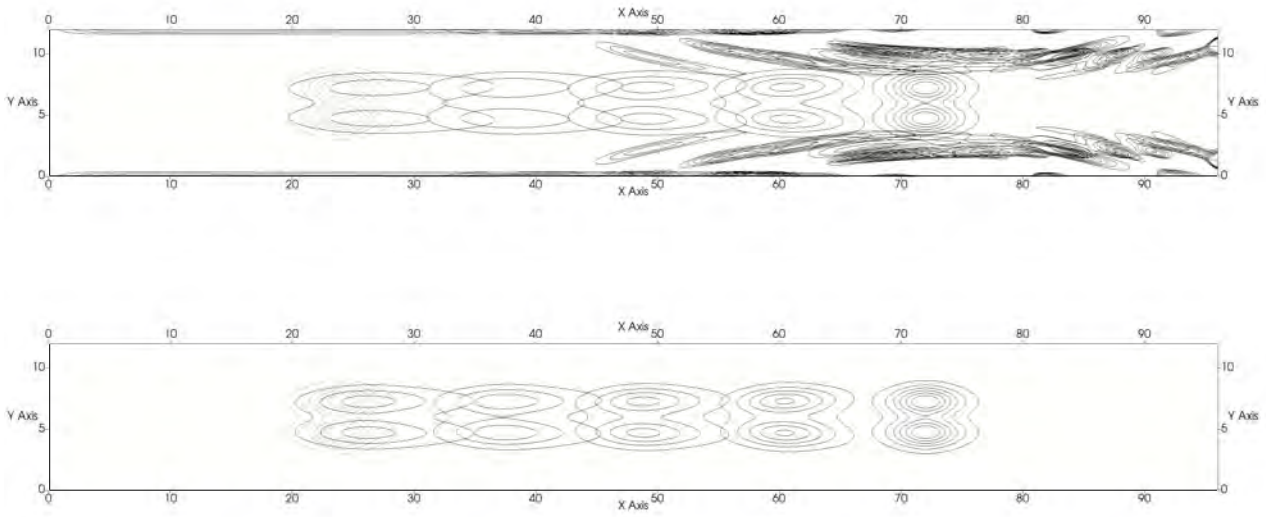


Figure 12: Section 5.7. Numerical solution for $h$ (m) provided by the 1-st order scheme at times $t = 0$, $t = 30$, $t = 60$, $t = 90$ and $t = 120$ s, using two different grids with $\Delta x = 0.2$ (top) and $\Delta x = 0.1$ m (bottom). The contour plot has been generated using 6 intervals from 1.02 to 1.14 m.

eastward [74]. This leads to the generation of equatorially trapped inertial-gravity (or Poincare) waves, which is an analogous mechanism than for nonlinear coastal Kelvin waves.

In this test case, we aim to show that the proposed numerical scheme is able to reproduce the formation and propagation of both Rossby and Kelvin waves over a non-flat bed elevation and eventually the generation of the Kelvin front and resonant formation of Poincare waves. The computational domain for this test case is $\Omega = [0, 70] \times [0, 12]$, units in m. The initial perturbation is given by a Gaussian water surface elevation anomaly, which reads

$$h(x, y) = h_0 + 0.8 \exp\left(-\frac{(x-30)^2 + (y-6)^2}{3}\right) \tag{80}$$

where $h_0 = 2$ m. The bed elevation is given by

$$z(x, y) = \begin{cases} 0 & \text{if } x \geq 40 \\ 0.025x - 1 & \text{if } x > 40 \end{cases} \tag{81}$$

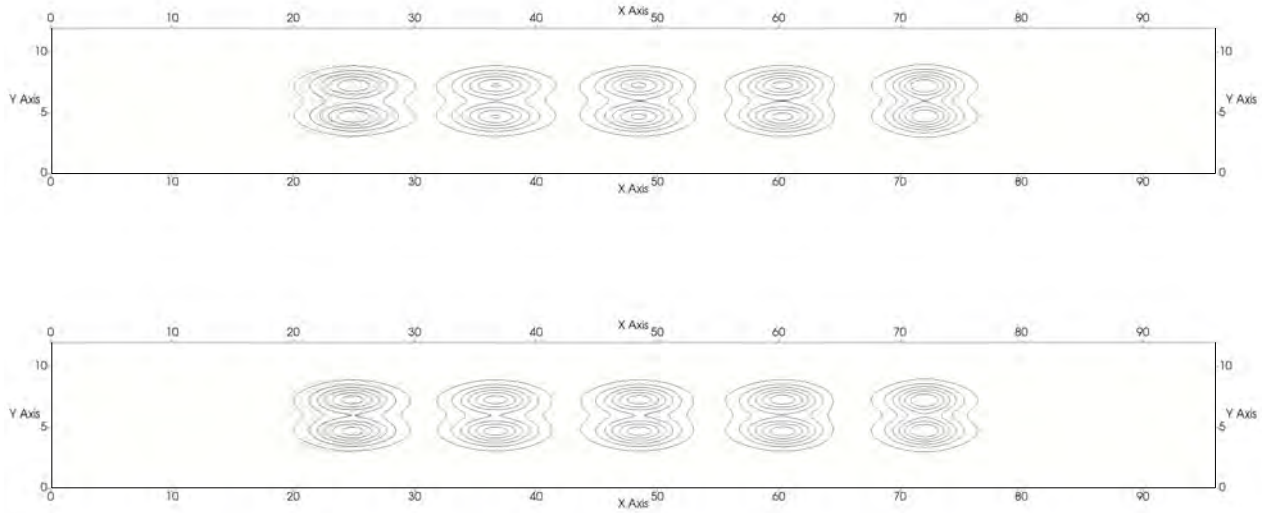


Figure 13: Section 5.7. Numerical solution for $h$ (m) provided by the 3-rd order scheme at times $t = 0$, $t = 30$, $t = 60$, $t = 90$ and $t = 120$ s, using two different grids with $\Delta x = 0.2$ (top) and $\Delta x = 0.1$ m (bottom). The contour plot has been generated using 6 intervals from 1.02 to 1.14 m.

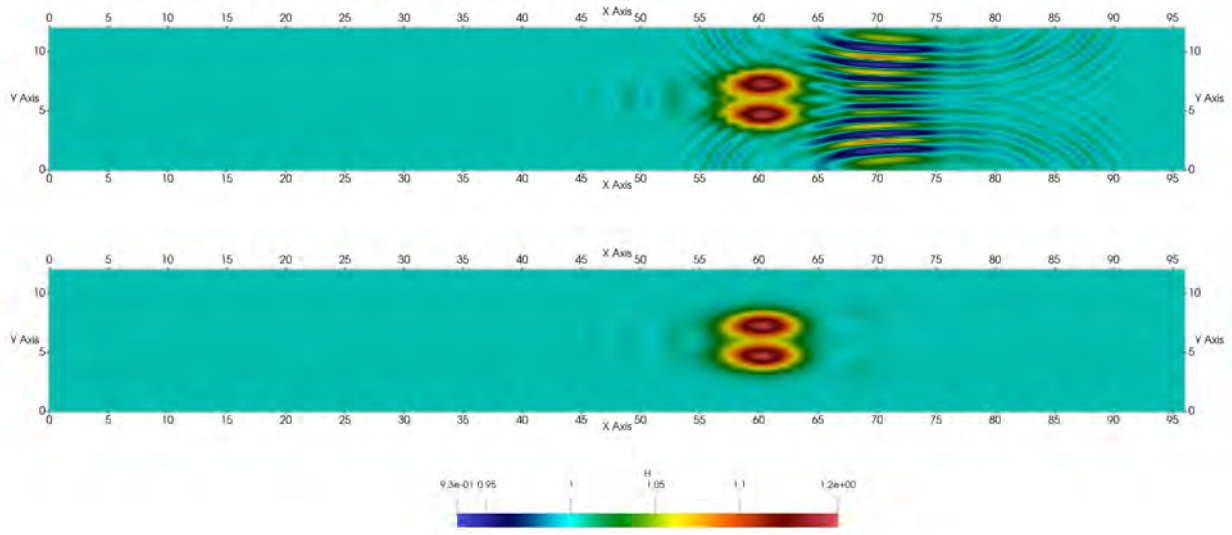


Figure 14: Section 5.7. Numerical solution for $h$ (m) provided by a well-balanced (bottom) and non well-balanced (top) 3-rd order ADER scheme at time $t = 30$ s using $\Delta x = 0.1$ m.

The numerical solution is computed using the 1-st and 3-rd order ARL-ADER scheme in a $1400 \times 240$ grid, using $CFL = 0.4$. The solution for $h + z$ at $t = 40$ s is depicted in Figure 15 using a contour plot with 20 intervals from 1.94 to 2.36 m. It is observed that only when using the 3-rd order scheme, an accurate resolution of the Kelvin front formation is possible and Poincare waves are captured. Regarding the planetary waves moving westward, it is worth mentioning that both schemes are able to reproduce the expected physical behavior, though the first order scheme is more diffusive and

dispersive.

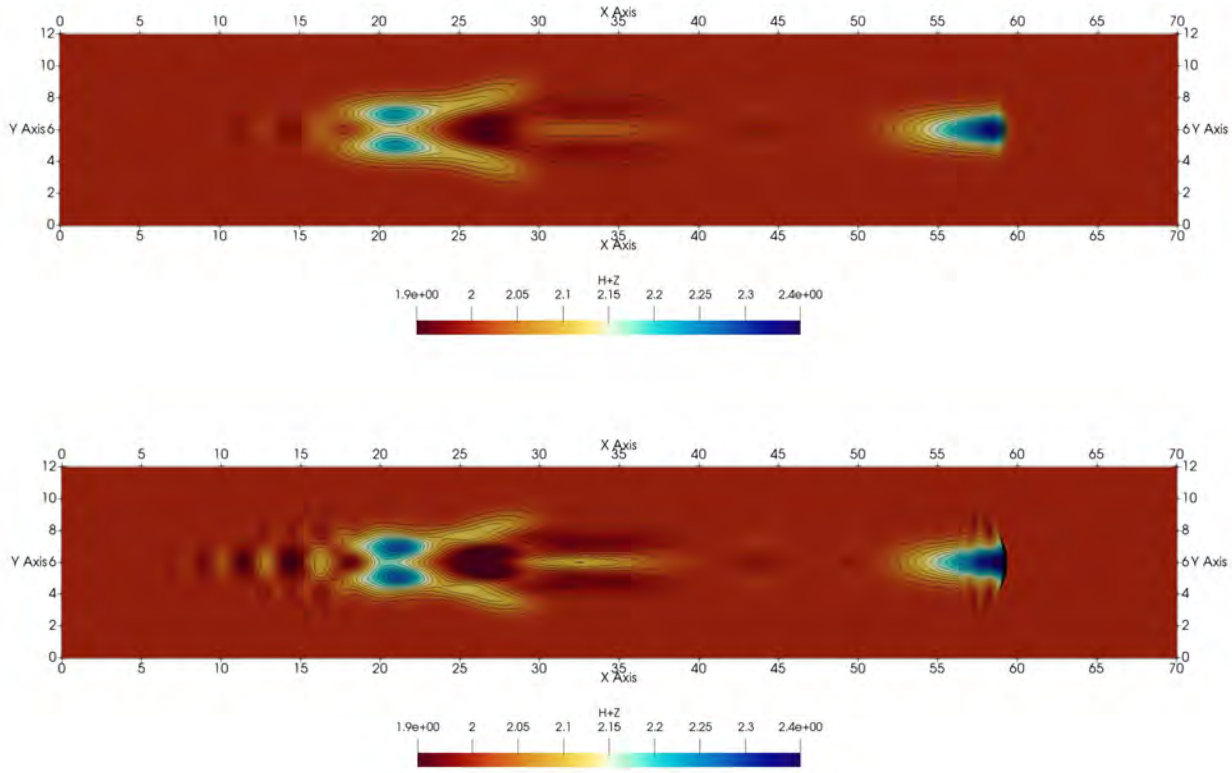


Figure 15: Section 5.8. Numerical solution for $h+z$ (m) provided by the 1-st (top) and 3-rd order (bottom) ARL-ADER scheme in the fine mesh at $t = 40$ s using $CFL = 0.4$. The contour plot has been generated using 20 intervals from 1.94 to 2.36 m.

### *5.9. Anticyclonic propagation in the $\beta$-plane*

The proposed scheme is applied here to a more realistic case from [75] that consists of a initially symmetric vortex propagating westward due to the effect of the variation of the Coriolis coefficient in the $y$-direction. The domain extent is an idealized 2000 km $\times$ 1200 km rectangular basin and the initial condition is given by a Gaussian distribution of the free surface centered at the origin of the domain, prescribed together with a velocity field which is in geostrophic balance. The water depth at the initial time is given by

$$h(x,y) = h_0 + \zeta(x,y)\,, \tag{82}$$

where $h_0 = 1.631$ m is the water depth reference level and $\zeta(x,y)$ is the surface height anomaly given by

$$\zeta(x,y) = Ae^{-(x^2+y^2)/B^2} \tag{83}$$

with $A = 0.95$ m, $B = 130$ km and $x$ and $y$ given in km. The initial velocity field is given by

$$u_1(x,y) = 2A\frac{g}{f(y)}\frac{y}{B^2}e^{-(x^2+y^2)/B^2}\,, \quad u_2(x,y) = -2A\frac{g}{f(y)}\frac{x}{B^2}e^{-(x^2+y^2)/B^2}\,, \tag{84}$$

with $f(y)$ the Coriolis parameter, evaluated at $25^o$ N using the $\beta$-plane approximation with $f_0 = 6.1635 \times 10^{-5}$ s$^{-1}$ and $\beta = 2.0746 \times 10^{-11}$ m$^{-1}$s$^{-1}$. The gravity constant is set to $g = 9.81$ m/s$^2$.

The solution is computed using the 1-st and 3-rd order ARL-ADER schemes at $t = 8$ weeks. Figure 16 shows the numerical solution for the 1-st (left) and 3-rd order scheme (right) using a $100 \times 60$ grid (top) and a $200 \times 120$ grid (bottom). As reported in [64], the first order scheme is not able to reproduce the physical behavior of the solution and is much more mesh dependent than the 3-rd order scheme, which provides a rather accurate solution even for the coarsest grid.

To compare the performance of the designed well-balanced 3-rd order scheme to that of a non well-balanced scheme, the solution has been computed using the 3-rd order non well-balanced scheme used in Section 5.4. The numerical solution at $t = 8$ weeks is presented in Figure 17 for a $200 \times 120$ grid (left) and a $400 \times 240$ grid (right). It is evidenced that the non well-balanced scheme provides a poorer resolution of the moving eddy and would require a very fine mesh to capture the solution with a similar accuracy than the well-balanced scheme. Note that a level of refinement higher than those of Figure 16 has been used, as the solution for the $100 \times 60$ grid was very inaccurate.

In Figure 18, the trajectory of the center of the moving vortex computed by the 1-st and 3-rd order well-balanced schemes is plotted in the $x-y$ plane. The trajectories are computed using three different grids composed of $100 \times 60$ cells, $200 \times 120$ cells and $400 \times 240$ cells. It is observed that the 3-rd order scheme provides an accurate prediction of the trajectory of the westward moving eddy, even for the coarsest grid, while the 1-st order scheme has a much lower convergence rate and requires more than $400 \times 240$ cells to predict the trajectory within an acceptable level of accuracy.

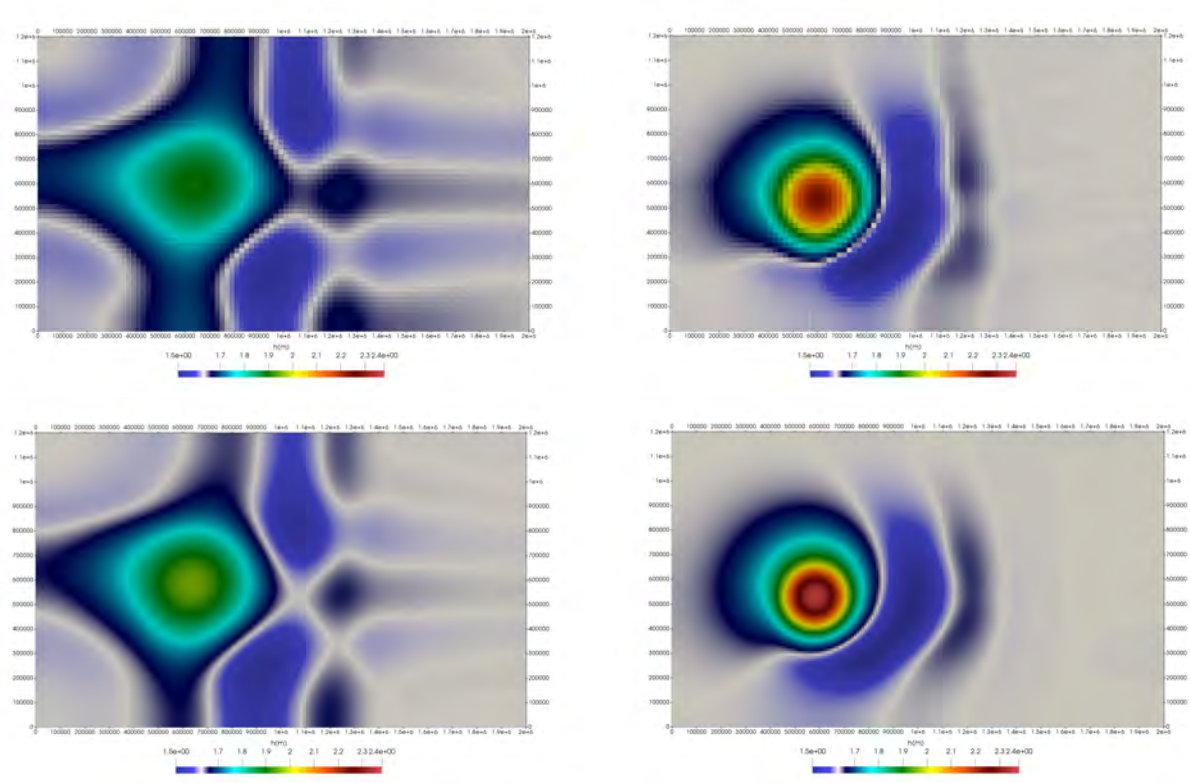

Figure 16: Section 5.9. Numerical solution for $h + z$ (m) computed by the 1-st (left) and 3-rd order scheme (right) using a $100 \times 60$ grid (top) and a $200 \times 120$ grid.

## 6. Summary and concluding remarks

In this work, an arbitrary order augmented WENO-ADER scheme for the resolution of the 2D

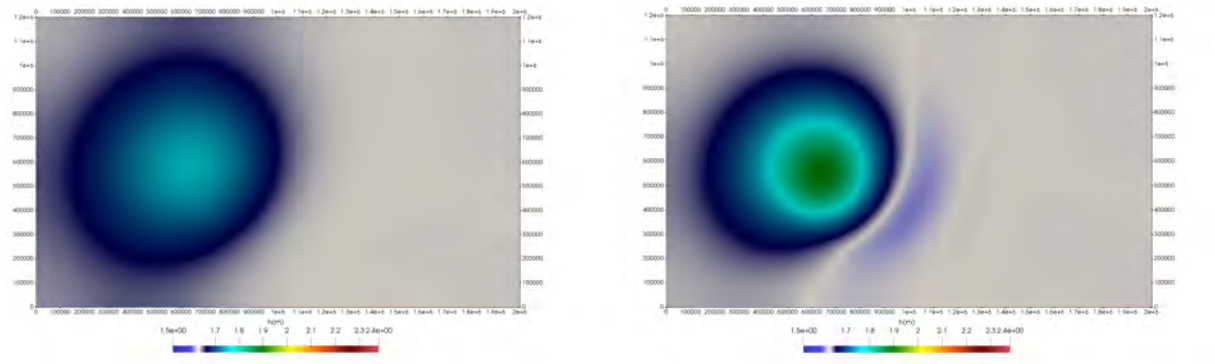

Figure 17: Section 5.9. Numerical solution for $h + z$ (m) computed by the non well-balanced 3-rd order scheme using a $200 \times 120$ grid (left) and a $400 \times 240$ grid (right).

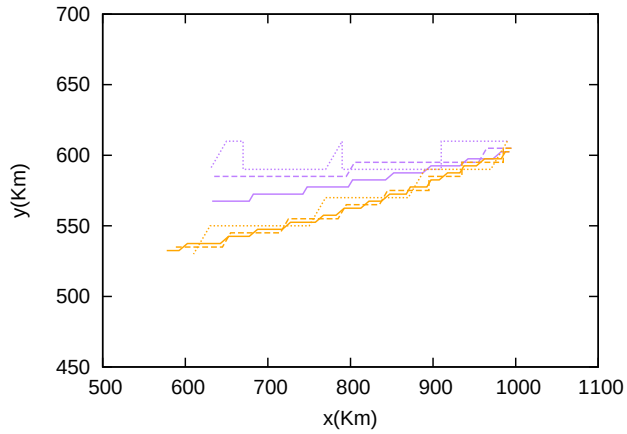


Figure 18: Section 5.9. Trajectory of the center of the moving vortex in the $x - y$ plane computed by the 1-st (purple) and 3-rd order scheme (orange) in a $100 \times 60$ cell grid (dotted line), $200 \times 120$ cell grid (dashed line) and $400 \times 240$ cell grid (solid line).

SWE with geometric source term has been presented. This scheme is the 2D extension of the ARL-ADER scheme in [29] which is based on the LFS Derivative Riemann solver. The solution of the DRP is computed using an Augmented solver, namely the ARoe solver. Compared to other traditional homogeneous solvers, the ARoe solver enhances the performance of the numerical scheme as it accounts for the contribution of the source term at cell interfaces. This allows to provide more accurate numerical approximations when dealing with geometric source terms, as well as to ensure the exact balance between fluxes and sources without requiring any additional correction. The extension of the aforementioned methods to the resolution of problems non-geometric source terms has also been explored in this work. This is the case of the SWE in the rotating frame.

The main novelty presented in this work is the 2D extension of the ARL-ADER scheme by using a particular arbitrary order discretization of the source term inside cells that ensures the exact preservation of the equilibrium states of relevance. This is achieved by reducing the 2D problem to two 1D problems in the Cartesian directions. This allows to derive a 2D arbitrary order approximation of the integral of the source term by means of the combination of 1D arbitrary order integrals using Gaussian quadrature and Romberg integration. The keystone for the preservation of equilibrium with very high order is the use of the Romberg integration method, which allows to extend the well–balanced discretization of the source term to arbitrary order. As a result, the proposed scheme is able to preserve the steady states of relevance while retaining a high order of accuracy for the resolution of transient

wave propagation.

The ARL-ADER scheme has been applied to the resolution of the SWE with bed variation and rotation, for which the relevant equilibrium states are the quiescent equilibrium and the geostrophic equilibrium. When considering rotation, the Coriolis source term has been reinterpreted as a geometric source term, allowing to apply the well-balanced algorithms applied for general geometric source terms. To ensure an exact equilibrium in the discrete level, special attention must be paid to the formulation of the CK procedure for the reconstruction of time derivatives.

The 2D ARL-ADER scheme has been assessed in a broad variety of cases involving bed variation and rotation. Such cases comprise smooth and discontinuous transient events, steady flows in complex geometries and steady and transient problems in the rotating frame including realistic scenarios with Rossby and Kelvin wave propagation. Most of the test cases propose the exercise of the scheme in situations where source terms are dominant. The numerical results show that the proposed scheme performs well in all cases, ensuring convergence to the reference solutions. The effectiveness of the well-balanced property is made visible in most of the presented tests. The capabilities of the augmented solver are evidenced in the test number 3, where the scheme is able to reproduce the full wave structure of the solution of a 2D resonant RP. It is also remarkable to mention the good performance of the proposed scheme for the resolution of the SWE in the rotating frame, compared to other non well-balanced schemes. In test case number 9 a realistic scenario involving the propagation of an anticyclonic eddy in the northern hemisphere is considered. Here, only when using the high order version of the well-balanced scheme, the physical solution is reproduced.

The convergence rate of the solution has been experimentally assessed. It is observed that the scheme achieves the expected accuracy when considering smooth cases with bed variation and Coriolis forces. It has also been shown that when considering an stiff Coriolis source term, the convergence rate may be reduced. The numerical results evidence that such loss of accuracy has to do with the integration in time, which is not optimally done due to the CK procedure. This underscores the importance of an optimal derivation of time derivatives and shows that when considering complex source terms, such as the geometric reinterpretation of the Coriolis source term, the CK procedure may become rather cumbersome and other techniques may be worth being used [21].

Concerning the overall computational expenses of the scheme, it is worth noting that the achievement of the well-balanced property while ensuring high order in space and time is done at the cost of using complex procedures. Such procedures increase the computational cost when compared to a typical non well-balanced method. The increased expense is mainly due to three aspects: the Romberg integration, doubling the number of spatial WENO reconstructions and the calculation of the Coriolis primitive variables. Though the latter does not involve a great computational cost, it enforces global communication in a parallelized code, which increases the computational expense.

The numerical results evidence that high order schemes are not only recommended, due to the improved efficiency of the methods, but sometimes necessary when the first order schemes are not able to reproduce the physical solution. The simulation tool proposed in this work provides an appropriate balance between computational efficiency and complexity of implementation. It is able to reproduce a broad variety of flows dominated by source terms by using a unified strategy based on the consideration of geometric source terms that allow to satisfy certain equilibrium conditions. The scheme uses a Cartesian mesh, but no mesh–dependency of the solution is observed due to the high order of accuracy. The scheme can be easily extended to other systems of equations with source terms of a broader variety.

**Acknowledgment**

This work has been funded by the Spanish Ministry of Science and Innovation under the research project CGL2015-66114-R.

| Decoration | Meaning |
|---|---|
| $(\cdot)_{i\pm1/2,j}$ | Relative to cell's right and left interfaces |
| $(\cdot)_{i,j\pm1/2}$ | Relative to cell's top and bottom interfaces |
| $(\cdot)_{i\pm1/2,j,q}$ | Relative to cell's right and left interfaces, at Gaussian point $q$ |
| $(\cdot)_{i,j\pm1/2,q}$ | Relative to cell's top and bottom interfaces, at Gaussian point $q$ |
| $(\cdot)_{i_r,j,q}$ | Left ($r = W$) and right ($r = E$) limit to the cell's interface, at Gaussian point $q$ |
| $(\cdot)_{i,j_r,q}$ | Upper ($r = N$) and lower ($r = S$) limit to the cell's interface, at Gaussian point $q$ |
| $(\cdot)_{r,q}$ | Relative to the location at the interface $r = N, S, E, W$ and Gaussian point $q$ |
| $\overline{(\cdot)}^n_{ij}$ | Cell average at time $t^n$ |
| $(\cdot)_{j\pm1/2}$ | Relative to cell's top and bottom interfaces |
| $\underline{(\cdot)}^{\pm}$ | Relative to the numerical solution at left ($-$) and right ($+$) limits to the RP interface |
| $\overline{(\cdot)}_{i+1/2}$ | Arithmetic mean at $i + 1/2$, that is, using cell averages at $i$ and $i + 1$ |
| $\widetilde{(\cdot)}_{i+1/2}$ | Roe's average at $i + 1/2$, that is, using cell averages at $i$ and $i + 1$ |
| $\delta(\cdot)_{i+1/2}$ | Discrete difference at $i + 1/2$, that is, $\delta(\cdot)_{i+1/2} = (\cdot)_{i+1} - (\cdot)_i$ |
| $(\cdot)^{(k)}$ | $k$-th time derivative |
| $\partial^{(k)}_{\phi}$ | $k$-th derivative with respect to variable $\phi$ |
| $\{\overset{\circ}{(\cdot)}\}^n_m$ | High order extension of quadrature rule $(\cdot)$ using Romberg integration |
| $\overset{\circ}{(\cdot)}$ | Evaluation of a variable at the initial time $t = 0$ |
| $(\cdot)_{\partial\Omega}$ | Evaluation of a variable at the boundary of the domain $\Omega$ |

Table 7: Dictionary of decorations

| Variable | Meaning |
|---|---|
| $x, y$ | Spatial coordinates |
| $\Delta x$ | Cell size |
| $t$ | Time |
| $\Delta t$ | Time step |
| $\mathbf{U}$ | Vector of conserved quantities |
| $\mathbf{E}$ | Matrix of physical fluxes |
| $\mathbf{S}$ | Vector of source terms |
| $\mathbf{F}$, $\mathbf{G}$ | Vectors of physical fluxes in $x$ and $y$ directions. |
| $\delta\mathbf{M}$ | Flux fluctuation |
| $\mathbf{A}$, $\mathbf{B}$ | Jacobian matrices in $x$ and $y$ directions. |
| $\mathbf{J}$ | General notation for the Jacobian matrix |
| $\mathcal{J}$ | Jacobian matrix composed of $\mathbf{A}$ and $\mathbf{B}$ |
| $\mathbf{e}^m$ | $m$-th eigenvector of the Jacobian |
| $\mathbf{P}$ | Matrix containing all eigenvectors $\mathbf{e}^m$ in columns |
| $\mathbf{\Lambda}$ | Diagonal matrix of the Jacobian |
| $\lambda^m$ | $m$-th eigenvalue |
| $\alpha^m$ | $m$-th wave strength |
| $\beta^m$ | $m$-th source strength |
| $\Omega$ | Spatial domain |
| $N_x$ | Number of cells in the $x$ direction |
| $N_y$ | Number of cells in the $y$ direction |
| $N^\lambda$ | Number of equations/eigenvalues |
| $\mathcal{F}$ | Normal flux to a cell interface |
| $h$ | Water depth |
| $u$, $v$ | Velocities in $x$ and $y$ directions |
| $\eta$ | Water surface elevation $h + z$ |
| $z$ | Bed elevation |
| $g$ | Gravity |
| $f$ | Coriolis coefficient |
| $V, U$ | Primitive variables of the $x$ and $x$ Coriolis forces |
| $L, K$ | Equilibrium variables in $x$ and $y$ directions in presence of Coriolis forces |
| $\omega$ | Gaussian quadrature weight |

Table 8: Dictionary of variables

## Appendix A. Resolution of the 1D DRP$_K$ using the LFS solver

The coefficients of the polynomial time series expansion of the numerical fluxes in (17) are computed by solving the DRP$_K$. The keystone for the resolution of the DRP$_K$ is that it can be decomposed in $K+1$ conventional RPs, one for the evolution of the conserved quantities and $K$ more RPs for the evolution of their derivatives. The former, hereafter referred to as DRP$_0$, allows to compute the leading terms of the series, $\mathbf{F}^{-,(0)}_{i+1/2}$ and $\mathbf{F}^{+,(0)}_{i+1/2}$, whereas the latter allow to compute the higher order terms, $\mathbf{F}^{-,(k)}_{i+1/2}$ and $\mathbf{F}^{+,(k)}_{i+1/2}$. Note that the subscripts $j$ and $q$ have been omitted for the sake of clarity and will not be displayed throughout this section. It is worth recalling that the DRP$_K$ includes the contribution of the source term at the origin and a suitable solver, namely the FS or LFS solver in [29], is required to compute the solution. The details for the decomposition of the DRP$_K$ in conventional RPs using the LFS are presented below.

The DRP$_0$ corresponds to the following RP

$$\begin{cases} \dfrac{\partial \mathbf{U}}{\partial t} + \dfrac{\partial \mathbf{F}}{\partial x} = \mathbf{S} \\ \mathbf{U}(x,0) = \begin{cases} \mathbf{U}^{(0)}_{i_E} & \text{if} \quad x<0 \\ \mathbf{U}^{(0)}_{(i+1)_W} & \text{if} \quad x>0 \end{cases} \end{cases} \tag{A.1}$$

with $\mathbf{U}^{(0)}_{i_E} = \lim_{x\to 0^-} \mathbf{U}_i(x)$ and $\mathbf{U}^{(0)}_{(i+1)_W} = \lim_{x\to 0^+} \mathbf{U}_{i+1}(x)$ the initial data for the conserved quantities at cell interfaces provided by an adequate reconstruction procedure, such as the WENO method.

The $K$ RPs associated to the high order terms of the DRP$_K$ are composed of the linearized evolution equations for time derivatives of the conserved variables, leading to the following RP

$$\begin{cases} \dfrac{\partial}{\partial t}\left(\partial_t^{(k)}\mathbf{U}\right) + \widetilde{\mathbf{J}}_{i+1/2}\dfrac{\partial}{\partial x}\left(\partial_t^{(k)}\mathbf{U}\right) = \partial_t^{(k)}\mathbf{S} \\ \partial_t^{(k)}\mathbf{U}(x,0) = \begin{cases} \mathbf{D}^{(k)}_{i_E} & \text{if} \quad x<0 \\ \mathbf{D}^{(k)}_{(i+1)_W} & \text{if} \quad x>0 \end{cases} \end{cases} \tag{A.2}$$

where $\widetilde{\mathbf{J}}_{i+1/2} = \tilde{\mathbf{J}}_{i+1/2}(\mathbf{U}^{(0)}_{i_E}, \mathbf{U}^{(0)}_{(i+1)_W})$ is an approximation of the Jacobian matrix of the flux in the $x$ direction, $\mathbf{A}(\mathbf{U})$, at each interface $x_{i+1/2}$, and satisfies

$$\delta\mathbf{F}^{(0)}_{i+1/2} = \widetilde{\mathbf{J}}_{i+1/2}\delta\mathbf{U}^{(0)}_{i+1/2}\,, \tag{A.3}$$

by using the Roe averages [30].

The Jacobian in (A.3) is used here to approximate time derivatives of the flux as

$$\partial_t^{(k)}\mathbf{F}(\mathbf{U}) = \widetilde{\mathbf{J}}_{i+1/2}\partial_t^{(k)}\mathbf{U}\,. \tag{A.4}$$

with $\partial_t^{(k)}\mathbf{U}$ the time derivatives of the conserved quantities, which will be denoted as $\mathbf{D}^{(k)}$.

The CK procedure is used to express time derivatives of the variables in terms of their spatial derivatives, which are computed using a suitable derivative reconstruction procedure [68]. This procedure is based on the recursive computation of time derivatives of $\mathbf{U}$ using (1), as follows

$$\partial_t^{(k)}\mathbf{U} = \partial_t^{(k-1)}\left(-\partial_x\mathbf{F}(\mathbf{U}) - \partial_y\mathbf{G}(\mathbf{U}) + \mathbf{S}(\mathbf{U})\right)\,. \tag{A.5}$$

For the sake of simplicity, the CK procedure is carried out using absolute coordinates and is expressed only in terms of space derivatives of the conserved quantities, as $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{S}$ depend upon $\mathbf{U}$. Time derivatives of the source term will be denoted by $\mathbf{Q}^{(k)} = \partial_t^{(k)}\mathbf{S}$.

Due to the presence of the source term in the DRP, an augmented solver must be used to solve (A.1) and (A.2). Following [29], the ARoe solver is chosen to compute $\mathbf{F}^{-,(0)}_{i+1/2}$, $\mathbf{F}^{-,(k)}_{i+1/2}$, $\mathbf{F}^{+,(0)}_{i+1/2}$ and $\mathbf{F}^{+,(k)}_{i+1/2}$. The details for this procedure are summarized below.

In what follows, $\delta(\cdot)_{i+1/2}$ operator will represent the difference between the right and left state of the DRP centered in $i+1/2$ for a given variable, as $\delta(\cdot)_{i+1/2} = (\cdot)_{(i+1)_W} - (\cdot)_{i_E}$ and $\delta(\cdot)_{i-1/2} = (\cdot)_{i_W} - (\cdot)_{(i-1)_E}$.

Considering that (A.1) is hyperbolic, a set of approximated eigenvalues $\tilde{\lambda}^m_{i+1/2}$ and eigenvectors $\widetilde{\mathbf{e}}^m_{i+1/2} = (e^m_1, ..., e^m_{N_\lambda})^T$ can be calculated. The approximate Jacobian $\widetilde{\mathbf{J}}_{i+1/2}$ can be expressed as

$$\widetilde{\mathbf{J}}_{i+1/2} = \widetilde{\mathbf{P}}_{i+1/2}\widetilde{\mathbf{\Lambda}}_{i+1/2}\widetilde{\mathbf{P}}^{-1}_{i+1/2}\,, \tag{A.6}$$

with $\widetilde{\mathbf{P}}_{i+1/2} = \left(\widetilde{\mathbf{e}}^1, ..., \widetilde{\mathbf{e}}^{N_\lambda}\right)_{i+1/2}$ an invertible matrix composed by the eigenvectors of $\widetilde{\mathbf{J}}_{i+1/2}$ and $\widetilde{\mathbf{\Lambda}}_{i+1/2}$ the diagonal matrix composed by the eigenvalues of $\widetilde{\mathbf{J}}_{i+1/2}$.

When using the ARoe solver, the coefficients of (17) read

$$\begin{aligned}
\mathbf{F}^{-,(k)}_{i+1/2} &= \mathbf{F}^{(k)}_{i_E} + \textstyle\sum_{m=1}^{N_\lambda} \left(\widetilde{\lambda}^- \alpha^{(k)} - \beta^{-,(k)}\right)^m_{i+1/2} \widetilde{\mathbf{e}}^m_{i+1/2}\,, \quad k = 0, K \\
\mathbf{F}^{+,(k)}_{i+1/2} &= \mathbf{F}^{(k)}_{(i+1)_W} + \textstyle\sum_{m=1}^{N_\lambda} \left(\widetilde{\lambda}^+ \alpha^{(k)} - \beta^{+,(k)}\right)^m_{i+1/2} \widetilde{\mathbf{e}}^m_{i+1/2}\,, \quad k = 0, K
\end{aligned} \tag{A.7}$$

with

$$\left(\tilde{\lambda}^\pm\right)^m_{i+1/2} = \left(\frac{\widetilde{\lambda} \pm |\widetilde{\lambda}|}{2}\right)^m_{i+1/2}, \qquad \left(\beta^{\pm,(0)}\right)^m_{i+1/2} = \left(\frac{\widetilde{\lambda}^\pm}{\widetilde{\lambda}}\beta^{(0)}\right)^m_{i+1/2}, \tag{A.8}$$

$\mathbf{A}^{(k)}_{i+1/2} = \left(\alpha^{(k),1}, ..., \alpha^{(k),N_\lambda}\right)^T_{i+1/2}$ the wave strengths and $\mathbf{B}^{(k)}_{i+1/2} = \left(\beta^{(k),1}, ..., \beta^{(k),N_\lambda}\right)^T_{i+1/2}$ the source strengths.

It is worth showing that the term $\sum_{m=1}^{N_\lambda} \left(\widetilde{\lambda}\alpha^{(k)} - \beta^{(k)}\right)^m_{i+1/2} \widetilde{\mathbf{e}}^m_{i+1/2}$ stands for the flux fluctuation

$$\delta\mathbf{M}_{i+1/2} = \delta\mathbf{F}_{i+1/2} - \bar{\mathbf{S}}_{i+1/2} \tag{A.9}$$

The fluctuation in (A.9) can be rewritten as

$$\delta\mathbf{F}_{i+1/2} - \bar{\mathbf{S}}_{i+1/2} = \mathbf{P}(\mathbf{\Lambda}\mathbf{P}^{-1}\delta\mathbf{U} - \mathbf{P}^{-1}\bar{\mathbf{S}})_{i+1/2} \tag{A.10}$$

which helps understanding how the source term is introduced in the formulation of the numerical fluxes. This is the basis of the augmented approach.

Physical fluxes as well as wave and source strengths are computed differently for (A.1) and (A.2). Following [29] they read

$$\mathbf{F}^{(k)}_{i_E} = \begin{cases} \mathbf{F}\left(\mathbf{U}^{(0)}_{i_E}\right) & \text{if } k = 0 \\ \widetilde{\mathbf{J}}_{i+1/2}\mathbf{D}^{(k)}_{i_E} & \text{if } k > 0 \end{cases}, \quad \mathbf{A}^{(k)}_{i+1/2} = \begin{cases} \widetilde{\mathbf{P}}^{-1}_{i+1/2}\delta\mathbf{U}^{(0)}_{i+1/2} & \text{if } k = 0 \\ \widetilde{\mathbf{P}}^{-1}_{i+1/2}\delta\mathbf{D}^{(k)}_{i+1/2} & \text{if } k > 0 \end{cases} \tag{A.11}$$

and $\mathbf{B}^{(k)}_{i+1/2} = \widetilde{\mathbf{P}}^{-1}_{i+1/2}\bar{\mathbf{S}}^{(k)}_{i+1/2}$. A suitable approximation of the integral of the source term across the interface

$$\bar{\mathbf{S}}^{(0)}_{i+1/2} \approx \int_{x^-_{i+1/2}}^{x^+_{i+1/2}} \mathbf{S}(x, y, 0)\, dx\,, \quad \bar{\mathbf{S}}^{(k)}_{i+1/2} \approx \int_{x^-_{i+1/2}}^{x^+_{i+1/2}} \mathbf{Q}^{(k)}\, dx\,, \tag{A.12}$$

must be found in order to satisfy an exact equilibrium of fluxes and sources at the interface.

## Appendix B. Approximate solution using the linearized ARoe (ARL) solver

The approximate Jacobian $\tilde{\mathbf{J}}_{\xi+1/2}(\mathbf{U}^{(0)}_{\xi_R}, \mathbf{U}^{(0)}_{(\xi+1)_L})$ for the DRP$_K$ in (A.3) is defined at the interface $x_{\xi+1/2}$, with $\xi = i, j$, using the left and right states of the reconstructed variables and reads

$$\widetilde{\mathbf{J}}_{\xi+1/2} = \begin{pmatrix} 0 & n_x & n_y \\ \widetilde{c}^2 n_x - \widetilde{u}(\widetilde{\mathbf{v}}\cdot\hat{\mathbf{n}}) & \widetilde{\mathbf{v}}\cdot\hat{\mathbf{n}} + \widetilde{u}n_x & \widetilde{u}n_y \\ \widetilde{c}^2 n_y - \widetilde{v}(\widetilde{\mathbf{v}}\cdot\hat{\mathbf{n}}) & \widetilde{v}n_x & \widetilde{\mathbf{v}}\cdot\hat{\mathbf{n}} + \widetilde{v}n_y \end{pmatrix}, \tag{B.1}$$

It is constructed using the Roe averages [30] and is diagonalized by the eigenvectors $\widetilde{\mathbf{e}}^1_{\xi+1/2}, \widetilde{\mathbf{e}}^2_{\xi+1/2}, \widetilde{\mathbf{e}}^3_{\xi+1/2}$, leading to the diagonal matrix $\widetilde{\mathbf{\Lambda}}_{\xi+1/2} = diag(\widetilde{\lambda}^1_{\xi+1/2}, \widetilde{\lambda}^2_{\xi+1/2}, \widetilde{\lambda}^3_{\xi+1/2})$.

Regarding the source strengths for the leading term, it is worth recalling that only when considering geometric source terms the source strengths are not nil. Otherwise, the source term is not included in the definition of the DRP. When dealing with geometric source terms, we only consider non-zero source components for the momentum equations as follows

$$\bar{\mathbf{S}}^{(0)}_{\xi+1/2} = \left(0, \bar{S}^{x,(0)}_{\xi+1/2}, \bar{S}^{y,(0)}_{\xi+1/2}\right)^T, \tag{B.2}$$

with $\bar{\mathbf{S}}^{(0)}_{\xi+1/2}$ defined in (A.12). Then, we can define the following vector $\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2} \in \mathbb{R}^2$, where $^M$ stands for *momentum*, as

$$\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2} = \left(\bar{S}^{x,(0)}_{\xi+1/2}, \bar{S}^{y,(0)}_{\xi+1/2}\right)^T, \tag{B.3}$$

and use it for the definition of the source strengths as

$$\beta^{(0),1}_{\xi+1/2} = -\frac{1}{2\widetilde{c}}\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}, \quad \beta^{(0),2}_{\xi+1/2} = -\frac{1}{\widetilde{c}}\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}_\perp, \quad \beta^{(0),3}_{\xi+1/2} = -\beta^{(0),1}_{\xi+1/2} \tag{B.4}$$

where $\hat{\mathbf{n}}_\perp = \mathbf{R}_{\pi/2}\hat{\mathbf{n}}$ is the unitary vector parallel to the cell interface and $\mathbf{R}_{\pi/2}$ a $\pi/2$ rad rotation matrix. The expression of the source strengths for the higher order terms can be derived analogously.

If considering a geometric source term, the projection $\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}$ must be an approximation of

$$\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}} \approx \frac{1}{\Delta t}\int_0^{\Delta t}\int_{\breve{x}^-_{i+1/2}}^{\breve{x}^+_{i+1/2}} S_s(\mathbf{U})\nabla\phi\cdot\hat{\mathbf{n}}d\breve{x}, \tag{B.5}$$

and using the relations $dx = d\breve{x}n_x$ and $dy = d\breve{x}n_y$, it can be approached by

$$\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}} = S_s(\bar{\mathbf{U}}^{(0)}_{\xi+1/2})\int_{\mathbf{x}^-_{i+1/2}}^{\mathbf{x}^+_{\xi+1/2}}\nabla\phi d\mathbf{x} = S_s(\bar{\mathbf{U}}^{(0)}_{v+1/2})\delta(\phi)^{(0)}_{\xi+1/2}, \tag{B.6}$$

On the other hand, the projection $\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}_\perp$ must be an approximation of

$$\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}_\perp \approx \frac{1}{\Delta t}\int_0^{\Delta t}\int_{\breve{x}^-_{\xi+1/2}}^{\breve{x}^+_{\xi+1/2}} S_s(\mathbf{U})\nabla\phi\cdot\hat{\mathbf{n}}_\perp d\breve{x}, \tag{B.7}$$

where $\nabla\phi\cdot\hat{\mathbf{n}}_\perp$ is the directional derivative of $\phi$ in the direction parallel to the cell interface. According to the definition of the DRP$_K$, we only consider variations of the variables in the normal direction to the cell interface, hence $\bar{\mathbf{S}}^{M,(0)}_{\xi+1/2}\cdot\hat{\mathbf{n}}_\perp = 0$ and $\beta^{(0),2}_{\xi+1/2} = 0$.

## References


[1] B. van Leer, Towards the ultimate conservative difference scheme V. A second order sequel to Godunovs method, J. Comput. Phys. 32 (1979) 101–136.

[2] A. Harten, High resolution schemes for hyperbolic conservation laws, J. Comput. Phys. 49 (1983) 357–393.

[3] P. Colella and P.R. Woodward, The piecewise parabolic method (PPM) for gas dynamical simulations, J. Comput. Phys. 54 (1984) 174–201.

[4] A. Harten, B. Engquist, S. Osher, S.R. Chakravarthy, Uniformly high order accuracy essentially non-oscillatory schemes III, J. Comput. Phys. 71 (1987) 231–303.

[5] X.-D. Liu, S. Osher, T. Chan, Weighted essentially non-oscillatory schemes, J. Comput Phys. 115 (1994) 200–212.

[6] T. Schwartzkopff, M. Dumbser, C.-D. Munz, Fast high order ADER schemes for linear hyperbolic equations, J. Comput. Phys. 197 (2004) 532–539.

[7] E.F. Toro, R.C. Millington, and L.A.M. Nejad. Primitive upwind methods for hyperbolic partial differential equations. In C. H. Bruneau, editor, Sixteenth International Conference on Numerical Methods for Fluid Dynamics. Lecture Notes in Physics, pages 421–426. Springer-Verlag, 1998.

[8] E.F. Toro, R.C. Millington, and L.A.M. Nejad. Towards very high order Godunov schemes. In E. F. Toro, editor, Godunov Methods. Theory and Applications, pages 907–940. Kluwer/Plenum Academic Publishers, 2001.

[9] E.F. Toro, V.A. Titarev, Solution of the generalised Riemann problem for advection-reaction equations, Proc. Roy. Soc. London A 458 (2002) 271–281.

[10] E.F. Toro, V.A. Titarev, ADER schemes for scalar hyperbolic conservation laws with source terms in three space dimensions, J. Comput. Phys. 202 (1) (2005) 196–215.

[11] E.F. Toro, V.A. Titarev, Derivative Riemann solvers for systems of conservation laws and ADER methods, J. Comput Phys. 212 (1) (2006) 150–165.

[12] O. Zanotti, F. Fambri, M. Dumbser, A. Hidalgo, Spacetime adaptive ADER discontinuous Galerkin finite element schemes with sub-cell finite volume limiting, Comput. Fluids. 118 (2015) 204–224.

[13] M. Dumbser, M. Kser and E.F. Toro, An arbitrary high-order Discontinuous Galerkin method for elastic waves on unstructured meshes-V. Local time stepping and p-adaptivity, Geophys. J. Int. 171 (2007) 695–717.

[14] M. Dumbser, M. Castro, C. Pars, E. F. Toro, ADER schemes on unstructured meshes for non-conservative hyperbolic systems: Applications to geophysical flows, Comput. Fluids. 38 (2009) 1731–1748.

[15] C.E. Castro, E.F. Toro and M. Käser, ADER scheme on unstructured meshes for shallow water: simulation of tsunami waves, Geophys. J. Int. 189 (2021) 1505–1520.

[16] C.E. Castro, E.F. Toro, Solvers for the high-order Riemann problem for hyperbolic balance laws, J. Comput. Phys. 227 (2008) 2481–2513.

[17] G. Vignoli, V.A. Titarev, E.F. Toro, ADER schemes for the shallow water equations in channel with irregular bottom elevation, J. Comput. Phys. 227 (2008) 2463–2480.

[18] G. Montecinos, C.E. Castro, M. Dumbser and E.F. Toro, Comparison of solvers for the generalized Riemann problem for hyperbolic systems with source terms, J. Comput. Phys. 231 (2012) 6472 – 6494.

[19] A. Canestrelli, A. Siviglia, M. Dumbser, E. F. Toro, Well-balanced high-order centred schemes for non-conservative hyperbolic systems. Applications to shallow water equations with fixed and mobile bed, Adv. Water Resour. 32 (2009) 834–844.

[20] E.F. Toro, G. Montecinos, Implicit, semi-analytical solution of the generalized Riemann problem for stiff hyperbolic balance laws, J. Comput. Phys. 303 (2015) 146–172.

[21] M. Dumbser, C. Enaux, E.F. Toro, Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws, J. Comput. Phys. 227 , Issue (2008) 3971–4001.

[22] E.F. Toro, Riemann solvers and numerical methods for fluid dynamics: a practical introduction, third ed., Springer-Verlag, Berlin, Heidelberg, 2009.

[23] D.L. George. Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation, J. Comput. Phys. 227 (2008) 3089–3113.

[24] J. Murillo, P. García-Navarro, Weak solutions for partial differential equations with source terms: application to the shallow water equations, J. Comput. Phys. 229 (2010) 4327–4368.

[25] J. Murillo, J. Burguete, P. Brufau, and P. García-Navarro. The influence of source terms on stability, accuracy and conservation in two-dimensional shallow flow simulation using triangular finite volumes, Int. J. Numer. Meth. Fluids (2007) 54 543–590.

[26] J. Murillo, P. García-Navarro, Augmented versions of the HLL and HLLC Riemann Solvers including source terms in one and two dimensions for shallow flow applications, J. Comput. Phys. 231 (2012) 6861–6906.

[27] J. Murillo, P. García-Navarro, Augmented Roe's approaches for Riemann problems including source terms: definition of stability region with application to the shallow water equations with rigid and deformable bed. In M. E. Vázquez-Cendón and A. Hidalgo and P. García-Navarro and L. Cea, eds., Numerical Methods for Hyperbolic Equations. Theory and Applications, pages 149–154. Taylor-Francis Group, 2013.

[28] A. Navas-Montilla, J. Murillo, Energy balanced numerical schemes with very high order. The Augmented Roe Flux ADER scheme. Application to the shallow water equations, J. Comput. Phys. 290 (2015) 188–218.

[29] A. Navas-Montilla, J. Murillo, Asymptotically and exactly energy balanced augmented flux-ADER schemes with application to hyperbolic conservation laws with geometric source terms, J. Comput. Phys. 317 (2016) 108–147.

[30] P.L. Roe, Approximate Riemann solvers, parameter vectors, and difference schemes, J. Comput. Phys. 43 (1981) 357–372.

[31] A. Harten, P. Lax and B. van Leer, On upstream differencing and Godunov type methods for hyperbolic conservation laws, SIAM review. 25 (1983) 35–61.

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

[33] J. Murillo and A. Navas-Montilla, A comprehensive explanation and exercise of the source terms in hyperbolic systems using Roe type solutions. Application to the 1D-2D shallow water equations, Advances in Water Resources 98 (2016) 70–96.

[34] A. Bermudez and M.E. Vázquez-Cendón, Upwind methods for hyperbolic conservation laws with source terms, Comput. Fluids. 23 (1994) 1049–1071.

[35] J.M. Greenberg, A.Y. Leroux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations, SIAM J. Numer. Anal. 33 (1996) 1–16.

[36] P. García-Navarro, M.E. Vázquez-Cendón. On numerical treatment of the source terms in the shallow water equations, Comput. and Fluids. 29 (2000) 951–979.

[37] J. Burguete and P. García-Navarro. Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows, Int. J. Numer. Meth. Fluids 37 (2001) 209–248.

[38] J. Burguete and P. García-Navarro. Implicit schemes with large time step for non-linear equations: application to river flow hydraulics, Int. J. Numer. Meth. Fluids 46 (2004) 607–636.

[39] A. Chinnayya, A.-Y. LeRoux, N. Seguin, A well-balanced numerical scheme for the approximation of the shallow water equations with topography: the resonance phenomenon, Int. J. Finite Vol. 1 (2004) 1–33.

[40] 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. 227 (2008) 3212–3243.

[41] G. Rosatti, L. Begnudelli, The Riemann Problem for the one-dimensional, free-surface Shallow Water Equations with a bed step: theoretical analysis and numerical simulations, J. Comput. Phys. 229 (2010) 760-787.

[42] P.G. LeFloch, M.D. Thanh, The Riemann problem for shallow water equations with discontinuous topography, Commun. Math. Sci. 5 (2007) 865–885.

[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] X. Yulong and S. Chi-Wang , High order finite difference WENO schemes with the exact conservation property for the shallow water equations, J. Comput. Phys. 208 (2005) 206–227.

[45] F. Alcrudo, F. Benkhaldoun, Exact solutions to the Riemann problem of the shallow water equations with a bottom step, Comput. Fluids 30 (2001) 643–671.

[46] B.F. Sanders, D.A. Jaffe and A.K. Chu. Discretization of integral equations describing flow in nonprismatic channels with uneven beds, J. Hydraul. Eng–ASCE 129(3) (2003) 235–244.

[47] G. Kesserwani, R. Ghostine, J. Vazquez, A. Ghenaim and R. Mosé. Application of a second-order Runge-Kutta discontinuous Galerkin scheme for the shallow water equations with source terms, Int. J. Numer. Meth. Fluids 56 (2008) 805–821.

[48] M. Catella, E. Paris and L. Solari. Conservative scheme for numerical modeling of flow in natural geometry, J. Hydraul. Eng–ASCE 134(6) (2008) 736–748.

[49] S.H. Lee and N.G. Wright. Simple and efficient solution of the shallow water equations with source terms, Int. J. Numer. Meth. Fluids 63 (2010) 313–340.

[50] V. Caleffi, A. Valiani, G. Li, A comparison between bottom-discontinuity numerical treatments in the DG framework, Appl. Math. Model. (2015).

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

[52] U.S. Fjordholm, S. Mishra, E. Tadmor,Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography, J. Comput. Phys. 230 (2011) 5587–5609.

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

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

[55] V. Caleffi, A. Valiani, Well balancing of the SWE schemes for moving-water steady flows, J. Comput. Phys., 342 (2017) 85–116.

[56] Y. Cheng, A. Chertock, M. Herty, A. Kurganov, A New Approach for Designing Moving-Water Equilibria Preserving Schemes for the Shallow Water Equations. Submitted preprint.

[57] J. Murillo, P. García-Navarro, Energy balance numerical schemes for shallow water equations with discontinuous topography, J. Comput. Phys. 236 (2012) 119–142.

[58] J. Murillo, P. García-Navarro, Accurate numerical modeling of 1D flow in channels with arbitrary shape. Application of the energy balanced property, J. Comput. Phys. 260 (2014) 222–248.

[59] A. Chertock, M. Dudzinski, A. Kurganov, M. Lukácová-Medvid'ová, Well-balanced schemes for the shallow water equations with Coriolis forces, Numer. Math, (2017) 1–35.

[60] E. Audusse, R. Klein, D. D. Nguyen, S. Vater, Preservation of the discrete geostrophic equilibrium in shallow water flows. Finite Volumes for Complex Applications VI Problems and Perspectives, 4 (2011) 59–67.

[61] A. C. Kuo, L. M. Polvani, Nonlinear geostrophic adjustment, cyclone/anticyclone asymmetry, and potential vorticiy rearrangement. Phys. of Fluids. 12 (2000) 1087–1100.

[62] F. Bouchut, J. Le Sommer, V. Zeitlin, Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. Part 2. High-resolution numerical simulations. J. Fluid Mech., 514 (2004) 35–63.

[63] N. Pankratz, J. R. Natvig, B. Gjevik, S. Noelle, High-order well-balanced finite-volume schemes for barotropic flows: Development and numerical comparisons, Ocean Model., 18 (2007) 53–79.

[64] Castro, M. J., Lpez, J. A., and Pars, C., Finite volume simulation of the geostrophic adjustment in a rotating shallow-water system. SIAM J. Sci. Comput., 31 (2008), 444–477.

[65] E. Audusse, R. Klein, A. Owinoh, Conservative discretization of Coriolis force in a finite volume framework, J. Comp. Phys., 228 (2009) 2934–2950.

[66] F. Bouchut, V. Zeitlin, A robust well-balanced scheme for multi-layer shallow water equations. Discrete Continuous Dyn. Syst. Ser. B, 13 (2010) 739–758.

[67] E. Godlewski, P.-A. Raviart *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer Science and Business Media, Berlin, 2013.

[68] J.B. Cheng, E. F. Toro, S. Jiang, W. Tang, A sub-cell WENO reconstruction method for spatial derivatives in the ADER scheme, J. Comput. Phys. 251 (2013) 53–80.

[69] High Order ADER FV/DG Numerical Methods for Hyperbolic Equations, PhD Thesis, Monographs of the School of Doctoral Studies in Environmental Engineering, University of Trento (2007).

[70] A. Emery, An evaluation of several differencing methods for inviscid fluid flow problems, J. Comput. Phys., 2 (1968) 306–331.

[71] P. Woodward, P. Colella, The numerical simulation of two-dimensional fluid flow with strong shocks, J. Comput. Phys., 54 (1984), 115–173.

[72] W.-Y. Sun, O. M. Sun, Numerical simulation of Rossby wave in shallow water, Comput. Fluids, 76 (2013), 116–127.

[73] Boyd JP, Equatorial solitary waves. Part-1: Rossby solitons, J. Phys. Oceanogr. 10 (1980), 699-717.

[74] Fedorov, A. V. and Melville, W. K., Kelvin fronts on the equatorial thermocline, J. Phys. Oceanogr., 30 (2000), 1692–1705.

[75] D. Y. Le Roux, V. Rostand, and B. Pouliot, Analysis of Numerically Induced Oscillations in 2D FiniteElement ShallowWater Models Part I: InertiaGravity Waves, SIAM J. Sci. Comp., 29 (2007), 331–360.