



**HAL**  
open science

## Numerical approximation of the shallow water equations with Coriolis source term

Emmanuel Audusse, Virgile Dubos, Arnaud Duran, Noémie Gaveau, Youssef Nasser, Yohan Penel

► **To cite this version:**

Emmanuel Audusse, Virgile Dubos, Arnaud Duran, Noémie Gaveau, Youssef Nasser, et al.. Numerical approximation of the shallow water equations with Coriolis source term. *ESAIM: Proceedings*, 2021, 70, pp.31-44. 10.1051/proc/202107003 . hal-03182659v2

**HAL Id: hal-03182659**

**<https://hal.science/hal-03182659v2>**

Submitted on 7 Jul 2021

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

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

## NUMERICAL APPROXIMATION OF THE SHALLOW WATER EQUATIONS WITH CORIOLIS SOURCE TERM

E. AUDUSSE<sup>1</sup>, V. DUBOS<sup>2</sup>, A. DURAN<sup>3</sup>, N. GAVEAU<sup>4</sup>, Y. NASSERI<sup>5</sup> AND Y. PENEL<sup>2</sup>

**Abstract.** We investigate in this work a class of numerical schemes dedicated to the non-linear Shallow Water equations with topography and Coriolis force. The proposed algorithms rely on Finite Volume approximations formulated on collocated and staggered meshes, involving appropriate diffusion terms in the numerical fluxes, expressed as discrete versions of the linear geostrophic balance. It follows that, contrary to standard Finite-Volume approaches, the linear versions of the proposed schemes provide a relevant approximation of the geostrophic equilibrium. We also show that the resulting methods ensure semi-discrete energy estimates. Numerical experiments exhibit the efficiency of the approach in the presence of Coriolis force close to the geostrophic balance, especially at low Froude number regimes.

### INTRODUCTION

The question of the accuracy of numerical schemes around stationary solutions or/and in asymptotic regimes has been a subject of great interest over the last two decades, see the seminal works [4, 10, 12] in late nineties and the books [5, 9] ten years later. In the context of geophysical flows and more particularly for finite-volume methods applied to shallow water equations, a lot of works have been devoted to the accuracy around the *lake-at-rest* equilibrium and more recently extended to nonzero velocity one dimensional stationary states. The accuracy of numerical schemes around geostrophic equilibria which are of fundamental importance for large scales atmospheric or oceanographic flows was less investigated, see [16] for a general introduction to geophysical rotating fluid dynamics. To our knowledge, the first work in this field is due to Bouchut, Le Sommer and Zeitlin [6] but was fully accurate only for one-dimensional flows, as exhibited in [2] where accurate and stable Godunov type schemes were designed for linear two-dimensional shallow water equations. Recently two independent works [13, 15] proposed an IMEX type scheme for fully nonlinear equations.

Here we aim at designing explicit schemes that are proved to be accurate and stable in the nonlinear framework. Our work is mostly based on the ideas developed in [2] for the linear rotating case and in [7, 8] for the nonlinear non-rotating case. We investigate schemes applied to collocated and staggered schemes.

---

<sup>1</sup> LAGA, Institut Galilée, Université Sorbonne Paris Nord – 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse.

<sup>2</sup> INRIA Paris – Sorbonne Université – CNRS (LJLL), team ANGE, 2 rue Simone Iff, CS 42112, 75589 Paris cedex 12.

<sup>3</sup> Institut Camille Jordan, Université Claude Bernard Lyon 1 – 43, boulevard du 11 novembre 1918, 69622 Villeurbanne.

<sup>4</sup> Institut Denis Poisson, Université D'Orléans, rue de Chartres, 45067 Orléans.

<sup>5</sup> Institut de Mathématiques de Marseille, Université d'Aix-Marseille – Technopôle Château Gombert, 13453, Marseille.

Let  $\Omega$  be an open bounded domain of  $\mathbb{R}^2$  and let  $T > 0$ . The nonlinear Shallow Water equations with Coriolis force formulated on  $\Omega \times (0, T)$  read:

$$\begin{cases} \partial_t h + \operatorname{div}(h\mathbf{u}) = 0, \\ \partial_t(h\mathbf{u}) + \operatorname{div}(h\mathbf{u} \otimes \mathbf{u}) + h(\nabla\phi + \omega\mathbf{u}^\perp) = 0, \end{cases} \quad (1)$$

where  $h$  is the water height and  $\mathbf{u} = (u_x, u_y)$  the horizontal velocity<sup>1</sup>. The Coriolis force is accounted for in the momentum equations through the angular speed  $\omega$ . Following [7, 14], the pressure forces appear under a non conservative form through the scalar potential  $\phi = gh$ , where  $g$  is the standard gravity constant. For the sake of simplicity, a flat topography is considered in the present work, but the proposed approaches naturally extended to varying bottoms.<sup>2</sup>

It is well-known that the total energy of the system decomposes as  $E = \mathcal{E} + \mathcal{K}$  where

$$\mathcal{E} = \frac{1}{2}gh^2 \quad \text{and} \quad \mathcal{K} = \frac{1}{2}h\|\mathbf{u}\|^2$$

stand respectively for potential and kinetic energies. We recall that the energy  $E$  plays the role of a mathematical entropy associated to the hyperbolic system (1) and regular solutions satisfy the following conservation law:

$$\partial_t E + \operatorname{div} \left[ \left( \phi + \frac{1}{2}\|\mathbf{u}\|^2 \right) h\mathbf{u} \right] = 0. \quad (2)$$

When developing numerical methods, main objectives are accuracy and stability. To get stability, a crucial objective is to build numerical approximations satisfying a discrete counterpart of (2) that ensures that the discrete energy is nonincreasing. To achieve this, a general strategy is to consider a sufficient amount of numerical diffusion in the scheme. But in some physical contexts such as low Froude number regimes or near specific stationary states, these diffusive terms may considerably degrade the accuracy of the approximations and specific schemes are needed. Here we are interested in flows around the geostrophic balance

$$\nabla\phi + \omega\mathbf{u}^\perp = 0, \quad \operatorname{div} \mathbf{u} = 0 \quad (3)$$

To address such an issue, based on the linear case [2], and on the non-linear case without Coriolis force [7, 8], we propose a numerical approach involving discrete versions of these equilibria in the numerical fluxes. As a preliminary step, the strategy can be understood at the continuous level by investigating how the model (1) behaves with respect to some generic perturbations  $(\mathbf{q}, \pi)$ :

$$\begin{cases} \partial_t h + \operatorname{div}(h\mathbf{u} - \mathbf{q}) = 0, \\ \partial_t(h\mathbf{u}) + \operatorname{div}(\mathbf{u} \otimes (h\mathbf{u} - \mathbf{q})) + (h\nabla\phi - \nabla\pi) + \omega(h\mathbf{u} - \mathbf{q})^\perp = 0. \end{cases} \quad (4)$$

Hence  $\mathbf{q}$  and  $\pi$  must be respectively seen as (small) perturbations with respect to the flow rate and to the hydrostatic pressure.

The solutions to the modified equations (4) satisfy the following energy balance:

$$\partial_t E + \operatorname{div} \left[ \left( \phi + \frac{1}{2}\|\mathbf{u}\|^2 \right) (h\mathbf{u} - \mathbf{q}) - \pi\mathbf{u} \right] = -\mathbf{q} \cdot (\nabla\phi + \omega\mathbf{u}^\perp) - \pi \operatorname{div} \mathbf{u}, \quad (5)$$

which motivates a choice for  $\mathbf{q}$  and  $\pi$  involving resp. the quantities  $\nabla\phi + \omega\mathbf{u}^\perp$  and  $\operatorname{div} \mathbf{u}$ . Let us remark that these quantities govern the geostrophic equilibrium (3) associated to System (1) linearized around the steady

<sup>1</sup> $\mathbf{u}^\perp = (-u_y, u_x)$ .

<sup>2</sup>In case of a non-flat topography  $z_b$ ,  $\phi = g(h + z_b)$ .

state  $(\tilde{h}, \tilde{\mathbf{u}}) = (h_0, 0)$  for a constant  $h_0$ :

$$\begin{cases} \partial_t h = -h_0 \operatorname{div} \mathbf{u}, \\ \partial_t \mathbf{u} = -(\nabla \phi + \omega \mathbf{u}^\perp). \end{cases}$$

From a numerical point of view, diffusion terms are thus expected to have regularizing effects in the sense that they allow to recover a discrete counterpart of (5). Moreover, such terms are intended to vanish close to the geostrophic equilibrium, which must improve the quality of the approximations in this regime. Hence, the present approach can be seen as a non-linear extension of the linear case [2].

The paper is split into two parts: in the first section, we present the two geometric frameworks (collocated *vs.* staggered), we build some discrete operators with mimetic properties and we propose two well-balanced numerical schemes whose semi-discrete (in time) versions are proved to be stable. The second section is dedicated to the assessment of those schemes by means of the stationary vortex test case.

## 1. DESIGN OF THE NUMERICAL SCHEMES

### 1.1. Collocated scheme

#### 1.1.1. Mesh and notations

To discretize System (1), we consider a tessellation  $\mathbb{T}$  of the computational domain  $\Omega \subset \mathbb{R}^2$  made of non-overlapping rectangular cells of sizes  $\Delta x, \Delta y$ . The set of all edges of the mesh is denoted by  $\mathcal{E}$  and the set of vertices by  $\mathcal{V}$ .

- A generic cell of  $\mathbb{T}$  is denoted by  $K$ , its area by  $m_K$  and its boundary by  $\partial K$ .
- The length of any edge  $e$  is denoted by  $m_e$ .
- Given a cell  $K$  and an edge  $e \in \partial K$ ,  $K_e$  is the neighbouring cell to  $e$  (other than  $K$ ) and  $\mathbf{n}_{e,K}$  the outward normal pointing to  $K_e$ .
- Given a vertex  $\mathbf{x}$ , a dual cell is built by joining the centers of the neighbouring cells. Its area is denoted by  $m_{\mathbf{x}}$ .

Notations are pictured on Figs. 1.

In particular, the discrete Green formula yields

$$\forall K \in \mathbb{T}, \quad \sum_{e \subset \partial K} m_e \mathbf{n}_{e,K} = \mathbf{0}. \quad (6)$$

In what follows, for any piecewise constant function  $\varphi$  we will denote

$$\bar{\varphi}_e = \frac{\varphi_{K_e} + \varphi_K}{2} \quad \text{and} \quad \bar{\varphi}_{\mathbf{x}} = \frac{1}{4} (\varphi_{K_{NW}} + \varphi_{K_{NE}} + \varphi_{K_{SE}} + \varphi_{K_{SW}})$$

and introduce the vector quantity

$$\delta \varphi_e = \frac{\varphi_{K_e} - \varphi_K}{2} \mathbf{n}_{e,K}.$$

In the same way, for a piecewise constant vector function  $\Phi$ , we will note

$$\delta \Phi_e = \frac{\Phi_{K_e} - \Phi_K}{2} \cdot \mathbf{n}_{e,K}.$$

Finally, the quantity

$$\varphi^\pm = \frac{1}{2} (\varphi \pm |\varphi|)$$

will refer to the the positive and negative parts of any scalar function  $\varphi$ . Before describing the schemes, we also need to define the following discrete divergence and gradient operators:

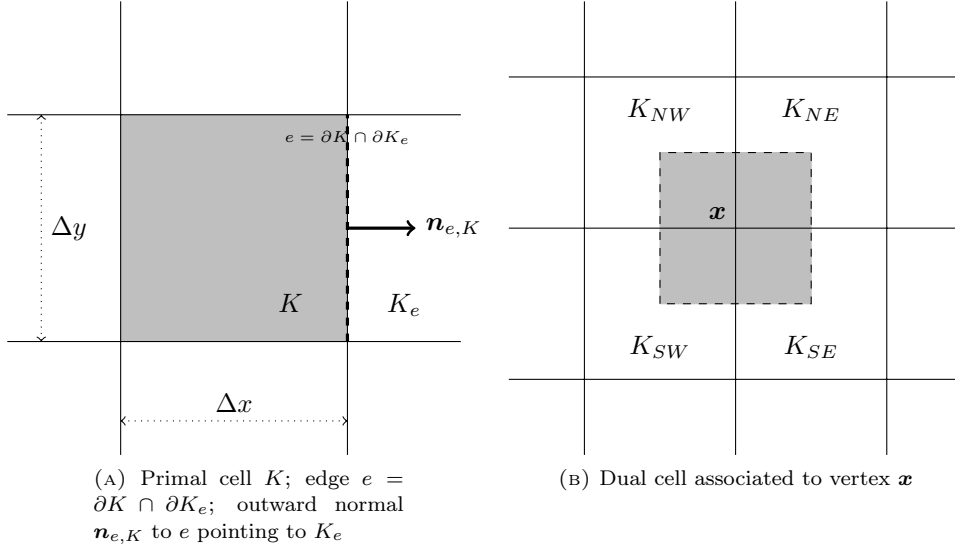


FIGURE 1. Geometric settings

- Given a vector function  $(\Phi_e)_{e \in \mathcal{E}}$ , the discrete divergence operator  $\text{div}_K$  reads

$$\forall K \in \mathbb{T}, \text{div}_K \Phi = \frac{1}{m_K} \sum_{e \subset \partial K} m_e \Phi_e \cdot \mathbf{n}_{e,K}. \quad (7)$$

if  $(\Phi_K)_{K \in \mathbb{T}}$  is rather considered, then  $\Phi_e$  is replaced by  $\bar{\Phi}_e$  in (7) and the corresponding operator is denoted by  $\text{div}_K^c \Phi$ ;

We can also define a divergence operator at the vertices when  $(\Phi_K)_{K \in \mathbb{T}}$  is considered by applying Formula (7) to the dual cell (see Fig.1(b)):

$$\forall \mathbf{x} \in \mathcal{V}, \text{div}_{\mathbf{x}}^c \Phi = \left( \frac{\mathbf{u}_{KSE} + \mathbf{u}_{KNE}}{2} - \frac{\mathbf{u}_{KSW} + \mathbf{u}_{KNW}}{2} \right) \cdot \frac{\mathbf{e}_1}{\Delta x} + \left( \frac{\mathbf{u}_{KNE} + \mathbf{u}_{KNW}}{2} - \frac{\mathbf{u}_{KSE} + \mathbf{u}_{KSW}}{2} \right) \cdot \frac{\mathbf{e}_2}{\Delta y}. \quad (8)$$

- In the same way, considering  $(\Phi_e)_{e \in \mathcal{E}}$  and  $(\Psi_K)_{K \in \mathbb{T}}$ , the upwind divergence operator applied to the formal product  $\Psi \otimes \Phi$  is defined by

$$\forall K \in \mathbb{T}, \text{div}_K^{up}(\Psi \otimes \Phi) = \frac{1}{m_K} \sum_{e \subset \partial K} m_e (\Psi_K (\Phi_e \cdot \mathbf{n}_{e,K})^+ + \Psi_{K_e} (\Phi_e \cdot \mathbf{n}_{e,K})^-). \quad (9)$$

Likewise, if  $(\Phi_K)_{K \in \mathbb{T}}$  is rather considered, then  $\Phi_e$  is replaced by  $\bar{\Phi}_e$  in (9) and the corresponding operator is denoted by  $\text{div}_K^{c,up}(\Psi \Phi)$ ;

- Given a scalar function  $(\phi_e)_{e \in \mathcal{E}}$ , the discrete gradient reads

$$\forall K \in \mathbb{T}, \nabla_K \phi = \frac{1}{m_K} \sum_{e \subset \partial K} m_e \phi_e \mathbf{n}_{e,K}. \quad (10)$$

Likewise, if  $(\phi_K)_{K \in \mathbb{T}}$  is rather considered, then  $\phi_e$  is replaced by  $\bar{\phi}_e$  in (10) and the corresponding operator is denoted by  $\nabla_K^c \phi$ .

If  $(\phi_{\mathbf{x}})_{\mathbf{x} \in \mathcal{V}}$  is considered, then  $\phi_e$  is replaced  $\frac{1}{2} \sum_{\mathbf{x} \in \partial e} \phi_{\mathbf{x}}$  and the corresponding operator is denoted by  $\nabla_K^x \phi$ .

In particular, we have the mimetic properties (see [2])

$$\operatorname{div}_K^c(\nabla_K^\perp \varphi) = 0 \quad \text{and} \quad \sum_{K \in \mathbb{T}} m_K \nabla_K^x \varphi \cdot \Phi_K = - \sum_{\mathbf{x} \in \mathcal{V}} m_{\mathbf{x}} \varphi_{\mathbf{x}} \operatorname{div}_x^c \Phi. \quad (11)$$

### 1.1.2. Numerical scheme

In this collocated version, all unknowns are located at the center of the cells except the perturbations ( $\mathbf{q}$  located at the middle of the interfaces and  $\pi$  at the vertices).

With the discrete operators (7-8-9-10) introduced above, the scheme reads, in the semi-discrete case:

$$\begin{cases} \frac{d}{dt} h_K + \operatorname{div}_K(\mathcal{F}) = 0, \\ \frac{d}{dt} (h_K \mathbf{u}_K) + \operatorname{div}_K^{up}(\mathbf{u} \otimes \mathcal{F}) + h_K \nabla_K^c \phi - \nabla_K^x \pi = -\omega \left( h_K \mathbf{u}_K - \frac{1}{2} \sum_{e \subset \partial K} \mathbf{q}_e \right)^\perp, \end{cases} \quad (12)$$

where the interface fluxes are defined at the level of the edge  $e$  as:

$$\mathcal{F}_e = \overline{h \mathbf{u}}_e - \mathbf{q}_e, \quad (13)$$

with

$$\mathbf{q}_e = \gamma \Delta t \bar{h}_e \left( 2 \frac{m_e}{m_K} \delta \phi_e + \omega (\bar{\mathbf{u}}_e^\perp \cdot \mathbf{n}_{e,K}) \mathbf{n}_{e,K} \right), \quad (14)$$

and

$$\pi_{\mathbf{x}} = \nu \Lambda \max\{\Delta x, \Delta y\} \bar{h}_{\mathbf{x}} \operatorname{div}_{\mathbf{x}}^c \mathbf{u}, \quad (15)$$

where  $\gamma$  and  $\nu$  are positive dimensionless constants, and  $\Lambda$  is a positive characteristic velocity.<sup>3</sup>

**Remark 1.1.** *Some remarks must be done on the formulation of the scheme:*

- (1) *The origin of the expression (14) for the perturbation  $q$  will be explained in the proof below. We first mention that it is consistent with the geostrophic equilibrium (it is aimed to vanish at the equilibrium). Second, we only consider the normal component of  $\bar{\mathbf{u}}_e$  in order to have the correct number of constraints. Moreover, as the operator  $\delta \phi_e$  is along the normal vector,  $\mathbf{q}_e = 0$  would imply (without the normal vector) a too restrictive kernel (characterised by  $u_{K,x} = 0$  or  $u_{K,y} = 0$  depending on the edges).*
- (2) *As  $m_K = \Delta x \Delta y$  in this case,  $\mathbf{q}_e$  does not actually depend on  $K$ .*
- (3) *Let us now explain why there is a  $\frac{1}{2}$  coefficient in the correction term of the right hand side of (12). The previous remark combined to the cartesian structure induces that there are only two terms by component in the mean  $\sum_e \mathbf{q}_e$ .*<sup>4</sup>

### 1.1.3. Semi-discrete energy

We now show this scheme ensures a discrete counterpart of (5) through semi-discrete mechanical energy estimates. To obtain such a result, we first need the two following lemmas, describing the evolution of potential and kinetic energies.

<sup>3</sup>Typically, we take  $\Lambda = \max_K \left\{ \|\mathbf{u}_K^0\| + \sqrt{gh_K^0} \right\}$ .

<sup>4</sup>To extend the scheme to a non-cartesian grid, the correction term  $\sum_e c_{e,K} \mathbf{q}_e$  must be consistent with  $\mathbf{q}_K$  with coefficients  $c_{e,K}$  such that  $c_{e,K} m_K$  independent from  $K$  so that  $\sum_K m_K \mathcal{A}_K^c = 0$ , and  $\frac{m_e}{m_K c_{e,K}} \frac{\phi_{K_e} - \phi_K}{2}$  must be consistent with  $\nabla \phi \cdot \mathbf{n}_{e,K}$ .

**Lemma 1.2** (Semi-discrete potential energy identity). *We set  $\mathcal{E}_K = \frac{1}{2}g(h_K)^2$  for  $K \in \mathbb{T}$ . Then:*

$$\frac{d}{dt}\mathcal{E}_K + \operatorname{div}_K(\phi\mathcal{F}) - h_K\mathbf{u}_K \cdot \nabla_K^c\phi + \mathcal{A}_K^\mathcal{E} = -\frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \cdot \mathbf{q}_e, \quad (16)$$

where the interface value for  $\phi\mathcal{F}$  is defined as:

$$(\phi\mathcal{F})_e = \bar{\phi}_e \mathcal{F}_e,$$

and the discrete conservative term:

$$\mathcal{A}_K^\mathcal{E} = \operatorname{div}_K(\delta\phi \delta(h\mathbf{u})) = \frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \delta(h\mathbf{u})_e \cdot \mathbf{n}_{e,K}. \quad (17)$$

*Proof.* Using the definition of  $\mathcal{E}_K$  and the mass conservation law (1), we get:

$$\frac{d}{dt}\mathcal{E}_K = \phi_K \frac{d}{dt}h_K = -\phi_K \operatorname{div}_K(\mathcal{F}).$$

Then, invoking the decomposition  $\phi_K = \bar{\phi}_e + \frac{1}{2}(\phi_K - \phi_{K_e})$  and the expression of the numerical fluxes (13):

$$\frac{d}{dt}\mathcal{E}_K = -\frac{1}{m_K} \sum_{e \subset \partial K} m_e \bar{\phi}_e \mathcal{F}_e \cdot \mathbf{n}_{e,K} + \frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \cdot \bar{h}\mathbf{u}_e - \frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \cdot \mathbf{q}_e.$$

Lastly, using the relation  $\bar{h}\mathbf{u}_e = h_K\mathbf{u}_K + \frac{1}{2}(h_{K_e}\mathbf{u}_{K_e} - h_K\mathbf{u}_K)$  and the discrete Green formula (6), basic computations yield:

$$\frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \cdot \bar{h}\mathbf{u}_e = h_K\mathbf{u}_K \cdot \nabla_K^c\phi + \frac{1}{m_K} \sum_{e \subset \partial K} m_e \delta\phi_e \delta(h\mathbf{u})_e \cdot \mathbf{n}_{e,K},$$

which allows to conclude.  $\square$

**Lemma 1.3** (Semi-discrete kinetic energy identity). *We set  $\mathcal{K}_K = \frac{1}{2}h_K\|\mathbf{u}_K\|^2$  for  $K \in \mathbb{T}$ . Then:*

$$\frac{d}{dt}\mathcal{K}_K + \frac{1}{2}\operatorname{div}_K^{up}(\|\mathbf{u}\|^2\mathcal{F}) + h_K\mathbf{u}_K \cdot \nabla_K^c\phi - \mathcal{A}_K^\mathcal{K} \leq -\frac{\omega}{2} \sum_{e \subset \partial K} \bar{\mathbf{u}}_e^\perp \cdot \mathbf{q}_e + \mathbf{u}_K \cdot \nabla_K^x\pi, \quad (18)$$

with anti-symmetric contributions given by:

$$\mathcal{A}_K^\mathcal{K} = \frac{\omega}{4} \sum_{e \subset \partial K} (\mathbf{u}_K - \mathbf{u}_{K_e}) \cdot \mathbf{q}_e^\perp. \quad (19)$$

*Proof.* We write:

$$\begin{aligned} \frac{d}{dt} \left( \frac{1}{2}h_K\|\mathbf{u}_K\|^2 \right) &= \mathbf{u}_K \cdot \frac{d}{dt}(h_K\mathbf{u}_K) - \frac{1}{2}\|\mathbf{u}_K\|^2 \frac{d}{dt}h_K \\ &= -h_K\mathbf{u}_K \cdot \nabla_K^c\phi - \mathbf{u}_K \cdot \operatorname{div}_K^{up}(\mathbf{u} \otimes \mathcal{F}) + \omega\mathbf{u}_K \cdot \left( \frac{1}{2} \sum_{e \subset \partial K} \mathbf{q}_e^\perp \right) + \frac{1}{2}\|\mathbf{u}_K\|^2 \operatorname{div}_K(\mathcal{F}) + \mathbf{u}_K \cdot \nabla_K^x\pi. \end{aligned} \quad (20)$$

First, we remark that:

$$\mathbf{u}_K \cdot \left( \frac{1}{2} \sum_{e \in \partial K} \mathbf{q}_e^\perp \right) = \frac{1}{2} \sum_{e \in \partial K} \bar{\mathbf{u}}_e \cdot \mathbf{q}_e^\perp + \frac{1}{4} \sum_{e \in \partial K} (\mathbf{u}_K - \mathbf{u}_{K_e}) \cdot \mathbf{q}_e^\perp = -\frac{1}{2} \sum_{e \in \partial K} \bar{\mathbf{u}}_e^\perp \cdot \mathbf{q}_e + \mathcal{A}_K^\mathcal{K}. \quad (21)$$

Then, after some basic computations, we get the relation

$$\frac{1}{2} \|\mathbf{u}_K\|^2 \operatorname{div}_K(\mathcal{F}) - \mathbf{u}_K \cdot \operatorname{div}_K^{up}(\mathbf{u} \otimes \mathcal{F}) = -\frac{1}{2} \operatorname{div}_K^{up}(\|\mathbf{u}\|^2 \mathcal{F}) + \frac{1}{2m_K} \sum_{e \in \partial K} \|\mathbf{u}_K - \mathbf{u}_{K_e}\|^2 (\mathcal{F}_e \cdot \mathbf{n}_{e,k})^- . \quad (22)$$

The second term of the right hand side being non-positive, we get the announced result inserting (21) and (22) into (20).  $\square$

Gathering relations (16) and (18), we obtain the following estimate for the total energy  $E_K = \mathcal{E}_K + \mathcal{K}_K$ :

$$\frac{d}{dt} E_K + \operatorname{div}_K(\phi \mathcal{F}) + \frac{1}{2} \operatorname{div}_K^{up}(\|\mathbf{u}\|^2 \mathcal{F}) + \mathcal{A}_K^\mathcal{E} - \mathcal{A}_K^\mathcal{K} \leq -\frac{1}{2} \sum_{e \in \partial K} \left[ 2 \frac{m_e}{m_K} \delta \phi_e + \omega \bar{\mathbf{u}}_e^\perp \right] \cdot \mathbf{q}_e + \mathbf{u}_K \cdot \nabla_K^x \pi .$$

With the choices (14) and (15), we finally obtain a discrete counterpart of (5) thanks to (11).

**Remark 1.4.** *The contributions  $\mathcal{A}_K^\mathcal{E}$  (17) and  $\mathcal{A}_K^\mathcal{K}$  (19) involve anti-symmetric terms and therefore have no impact on the total energy budget:*

$$\sum_{K \in \mathbb{T}} m_K (\mathcal{A}_K^\mathcal{E} - \mathcal{A}_K^\mathcal{K}) = 0,$$

*up to boundary terms. In fact, these quantities can be interpreted as a bias to the conservative parts of potential and kinetic energies respectively.*

Then:

**Theorem 1.5** (Semi-discrete mechanical energy balance). *The scheme (12, 13, 14, 15) satisfies the following semi-discrete inequality:*

$$\begin{aligned} & \frac{d}{dt} \left( \sum_{K \in \mathbb{T}} m_K E_K \right) \\ & \leq -\frac{\gamma \Delta t}{2} \sum_{K \in \mathbb{T}} m_K \sum_{e \in \partial K} \bar{h}_e \left\| \frac{\phi_{K_e} - \phi_K}{m_K/m_e} + \omega \bar{\mathbf{u}}_e^\perp \cdot \mathbf{n}_{e,K} \right\|^2 - \nu \Lambda \max\{\Delta x, \Delta y\} \sum_{\mathbf{x} \in \mathcal{V}} m_{\mathbf{x}} \bar{h}_{\mathbf{x}} (\operatorname{div}_{\mathbf{x}}^c \mathbf{u})^2. \end{aligned}$$

#### 1.1.4. Steady states

Let us first assess the ability of the scheme to capture the lake at rest, characterized by  $\phi = cte$  and  $\mathbf{u} = 0$ . One observes that the stabilization terms  $\mathbf{q}_e$  and  $\pi_{\mathbf{x}}$  vanish in such a configuration, leading directly to the exact preservation of the steady state in (12).

Then focus on the geostrophic equilibrium. In this case the nonlinear study is particularly complex and we choose to investigate the linear case. Although it leaves aside nonlinear interactions, this strategy may not seem irrelevant since most of low Froude number regimes encountered in practical contexts are governed by linear dynamics. Linearizing the scheme (12) around  $\tilde{\mathbf{u}} = 0$  and  $\tilde{h} = h_0$  for some constant  $h_0 > 0$ , we get:

$$\begin{cases} \frac{d}{dt} h_K + h_0 \operatorname{div}_K^c \mathbf{u} - \operatorname{div}_K \mathbf{q} = 0, \\ \frac{d}{dt} \mathbf{u}_K + \nabla_K^c \phi - \nabla_K^x \tilde{\pi} = -\omega \left( \mathbf{u}_K - \frac{1}{2} \sum_{e \in \partial K} \mathbf{q}_e \right)^\perp, \end{cases}$$



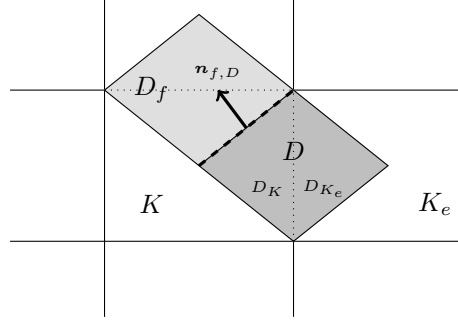


FIGURE 2. Geometric settings: for the primal mesh with cells  $K$  and  $K_e$ , refer to Figs. 1 – for the dual mesh with diamond cells  $D$  and  $D_f$  separated by the interface  $f = \partial D \cap \partial D_f$ ,  $\mathbf{n}_{f,D}$  is the outward normal of  $f$  pointing to  $D_f$ .

where

$$\mathbf{q}_e = \gamma \Delta t \left( 2 \frac{m_e}{m_K} \delta \phi_e + \omega (\bar{\mathbf{u}}_e^\perp \cdot \mathbf{n}_{e,K}) \mathbf{n}_{e,K} \right) \text{ and } \tilde{\pi}_{\mathbf{x}} = -\nu \Lambda \max\{\Delta x, \Delta y\} \operatorname{div}_{\mathbf{x}}^c \mathbf{u}.$$

One should observe here that the discrete geostrophic equilibrium, characterized by:

$$\nabla_K^c \phi + \omega \mathbf{u}_K^\perp = 0 \stackrel{(11)}{\implies} \operatorname{div}_K^c \mathbf{u}_K = 0$$

does not imply  $\mathbf{q}_e = 0$  in the general case. Conversely,  $\mathbf{q}_e = 0 \implies \pi_{\mathbf{x}} = 0$  according to [2] but does not imply  $\nabla_K^c \phi + \omega \mathbf{u}_K^\perp = 0$ .

As exhibited in the numerical validations, although it does not exactly preserve the geostrophic balance, the present scheme behaves much better than standard Godunov-type approaches. However, as detailed in the next section, it is possible to correct such a failure by means of a staggered mesh.

## 1.2. Staggered scheme

### 1.2.1. Mesh and notations

On the basis of the primal mesh  $\mathbb{T}$ , a dual grid  $\mathbb{T}^*$  can be built with respect to the Rannacher-Turek (RT) or Crouzeix-Raviart finite-element spaces, according to the formalism described in [1]. To set ideas, we focus here on the RT case considering a primal mesh  $\mathbb{T}$  made of regular quadrilateral elements like in the previous section, but the following framework also applies to simplicial meshes or mixed triangular/quadrilateral elements.

- A dual cell  $D \in \mathbb{T}^*$  is associated to each internal primal edge  $e$  as the quadrilateral admitting  $e$  as diagonal and the mass centers of the two neighbouring primal cells  $K$  and  $K_e$  as vertices (see Fig. 2).<sup>5</sup>
- This dual element  $D$  is the union of the two triangles  $D_K, D_{K_e}$  separated by the edge  $e$ , whose areas will be denoted  $m_{D_K}$  and  $m_{D_{K_e}}$ , following the notations of the primal mesh.
- Given a dual cell  $D$  and an edge  $e \subset \partial K$ ,  $D_f$  is the neighbouring cell to  $f$  (other than  $D$ ) and  $\mathbf{n}_{f,D}$  the outward normal pointing to  $D_f$ .

In this context, the water height evolves on the primal mesh  $\mathbb{T}$  as for the previous scheme while the velocity field is located on the dual elements  $D \in \mathbb{T}^*$ .

By analogy with the collocated frame, the upwind discrete divergence is defined on a dual element as:

$$\operatorname{div}_D^{up}(\Psi \otimes \Phi) = \frac{1}{m_D} \sum_{f \in \partial D} m_f (\Psi_D (\Phi_f \cdot \mathbf{n}_{f,D})^+ + \Psi_{D_f} (\Phi_f \cdot \mathbf{n}_{f,D})^-),$$

<sup>5</sup>The edge  $e$  and the corresponding diamond cell  $D$  will be used indifferently.

while the dual gradient is defined through the following formula:

$$\nabla_D \varphi = \frac{m_e}{m_D} (\varphi_{K_e} - \varphi_K) \mathbf{n}_{e,K}. \quad (23)$$

Note that this definition ensures the grad/div duality with respect to the  $L^2$  inner product in the present staggered environment (for instance, see [11]):

$$\sum_{K \in \mathbb{T}} m_K \varphi_K \operatorname{div}_K \Psi = - \sum_{D \in \mathbb{T}^*} m_D \nabla_D \varphi \cdot \Psi_D. \quad (24)$$

### 1.2.2. Numerical scheme

The scheme we consider is the following:

$$\begin{cases} \frac{d}{dt} h_K + \operatorname{div}_K(\mathcal{F}) = 0, \\ \frac{d}{dt} (h_D \mathbf{u}_D) + \operatorname{div}_D^{up}(\mathbf{u} \otimes \mathcal{F}) + h_D \nabla_D \phi - \nabla_D \pi = -\omega (h_D \mathbf{u}_D - \mathbf{q}_D)^\perp, \end{cases} \quad (25)$$

where the primal fluxes are defined by interface by:

$$\mathcal{F}_e = h_D \mathbf{u}_D - \mathbf{q}_D, \quad (26)$$

with

$$\mathbf{q}_D = \gamma \Delta t h_D (\nabla_D \phi + \omega \mathbf{u}_D^\perp), \quad (27)$$

and

$$\pi_K = \nu \Lambda \max\{\Delta x, \Delta y\} h_K \operatorname{div}_K \mathbf{u}. \quad (28)$$

As in the collocated frame,  $\gamma$  and  $\nu$  are positive constants and  $\Lambda$  is consistent with a velocity. Considering the momentum equations (25) and the diffusion term (27), it is necessary to define a dual water height  $h_D$ . Following [1], we compute

$$m_D h_D = m_{D_K} h_K + m_{D_{K_e}} h_{K_e}.$$

It is then possible to choose dual fluxes  $\mathcal{F}_f$  (involved in the divergence term of the momentum equations (25)) so that a discrete law of mass conservation also holds on diamond cells:

$$\frac{d}{dt} h_D + \frac{1}{m_D} \sum_{f \in \partial D} m_f \mathcal{F}_f \cdot \mathbf{n}_{f,D} = 0,$$

which can be rewritten in the more compact form:

$$\frac{d}{dt} h_D + \operatorname{div}_D(\mathcal{F}) = 0. \quad (29)$$

As detailed in [1], these dual fluxes are expressed as linear combinations of the primal ones. These relations are purely geometrical and formulated locally, depending on the nature of the staggered mesh. Such a relation is mandatory to derive a kinetic energy budget as shown in the next paragraph.

**Remark 1.6.** *Let us notice that in (27), both components are taken into account in the definition of  $\mathbf{q}_D$  unlike (14) where only the normal component was involved. In the present case,  $\mathbf{q}_D = 0$  implies that the normal component of the velocity field vanishes (since the gradient is normal to the interface) but the value of interest is the reconstructed mean velocity*

$$\mathbf{u}_K = \frac{1}{4} \sum_{e \subset \partial K} \mathbf{u}_D.$$

Here  $\mathbf{u}_D$  is an unknown in itself (unlike in the collocated case where  $\bar{u}_e$  is reconstructed from the unknowns  $\mathbf{u}_K$ ).

### 1.2.3. Semi-discrete energy

Indeed, defining the dual kinetic energy by  $\mathcal{K}_D = \frac{1}{2}h_D\|\mathbf{u}_D\|^2$  and using the relation

$$\frac{d}{dt}\mathcal{K}_D = \mathbf{u}_D \cdot \frac{d}{dt}(h_D\mathbf{u}_D) - \frac{1}{2}\|\mathbf{u}\|^2 \frac{d}{dt}h_D$$

we have, with (29) and applying the relation (22) on the dual mesh:

$$\frac{d}{dt}\mathcal{K}_D + \operatorname{div}_D^{up}(\|\mathbf{u}_D\|^2\mathcal{F}) + h_D\mathbf{u}_D \cdot \nabla_D\phi \leq -\omega\mathbf{u}_D^\perp \cdot \mathbf{q}_D + \mathbf{u}_D \cdot \nabla_D\pi. \quad (30)$$

Then, looking at the potential energy, basic computations give:

$$\frac{d}{dt}\mathcal{E}_K + \operatorname{div}_K(\phi\mathcal{F}) - \frac{1}{m_K} \sum_{e \in \partial K} m_e \delta\phi_e \cdot h_D\mathbf{u}_D = -\frac{1}{m_K} \sum_{e \in \partial K} m_e \delta\phi_e \cdot \mathbf{q}_D. \quad (31)$$

We are now left with the estimates (30) and (31), which are not expressed on the same mesh. A convenient way to understand the global impact of these quantities on the energy budget is to integrate them over their respective meshes, setting:

$$E = \sum_{K \in \mathbb{T}} m_K \mathcal{E}_K + \sum_{D \in \mathbb{T}^*} m_D \mathcal{K}_D,$$

and focus on the residual contributions at the level of dual elements. We first remark that, thanks to the duality relation (24), the second term of the right hand side of (30) furnish a first stabilizing term:

$$\sum_{D \in \mathbb{T}^*} m_D \mathbf{u}_D \cdot \nabla_D\pi = - \sum_{K \in \mathbb{T}} m_K \pi_K \operatorname{div}_K \mathbf{u},$$

according to the definition of  $\pi_K$  (28). Then, considering a generic element  $D \in \mathbb{T}^*$ , the non conservative terms in the kinetic energy balance (30) result in the following quantities on the left and right hand sides:

$$m_D h_D \mathbf{u}_D \cdot \nabla_D\phi \quad (LHS) \quad , \quad -m_D \omega \mathbf{u}_D^\perp \cdot \mathbf{q}_D. \quad (RHS)$$

On the other hand, the term issuing from (31) needs to be counted twice (contributions of the elements  $K$  and  $K_e$  sharing the edge  $e$ ), which yields:

$$-2m_e \delta\phi_e \cdot h_D \mathbf{u}_D \quad (LHS) \quad , \quad -2m_e \delta\phi_e \cdot \mathbf{q}_D. \quad (RHS)$$

Using the definition (23), we easily verify that the terms of the left hand side are in exact balance and the remaining term reads:

$$-m_D \mathbf{q}_D \cdot (\nabla_D\phi + \omega \mathbf{u}_D^\perp).$$

Replacing  $\mathbf{q}_D$  by its expression (27), we have thus established the following result:

**Theorem 1.7** (Semi-discrete mechanical energy balance). *The scheme (25-26-27) satisfies the following semi-discrete inequality:*

$$\frac{d}{dt}E \leq -\gamma \Delta t \sum_{D \in \mathbb{T}^*} m_D h_D \|\nabla_D\phi + \omega \mathbf{u}_D^\perp\|^2 - \nu \Lambda \max\{\Delta x, \Delta y\} \sum_{K \in \mathbb{T}} m_K h_K (\operatorname{div}_K \mathbf{u})^2.$$

As a remark, defining the quantity  $\mathcal{K}_K = \frac{1}{m_K} \sum_{e \in \partial K} m_D \mathcal{K}_D$ , which can be understood as a kinetic energy associated to the cell  $K$ , the previous result can be reformulated under local energy estimates. Note that the present formalism embeds the implicit and explicit approaches recently proposed in [8] without Coriolis force.

#### 1.2.4. Steady states

The linearized version of the scheme (25) is the following:

$$\begin{cases} \frac{d}{dt} h_K + h_0 \operatorname{div}_K \mathbf{u} - \operatorname{div}_K \mathbf{q} = 0, \\ \frac{d}{dt} \mathbf{u}_D + \nabla_D \phi - \nabla_D \tilde{\pi} = -\omega (\mathbf{u}_D - \mathbf{q}_D)^\perp, \end{cases}$$

where

$$\mathbf{q}_D = \gamma \Delta t (\nabla_D \phi + \omega \mathbf{u}_D^\perp) \quad \text{and} \quad \tilde{\pi}_K = \nu \Lambda \max\{\Delta x, \Delta y\} \operatorname{div}_K \mathbf{u}.$$

From this, we immediately see that discrete kernel of the system corresponds to the geostrophic balance expressed on the dual grid  $\mathbb{T}^*$ :

$$\nabla_D \phi + \omega \mathbf{u}_D^\perp = 0,$$

and is consequently exactly preserved by the linearized scheme. We may indeed check that  $\operatorname{div}_K(\nabla_D^\perp \phi) = 0$  by construction which implies  $\operatorname{div}_K \mathbf{u} = 0$  when the geostrophic equilibrium holds.

### 1.3. Time discretisation

For the discretisation in time, fluxes are taken explicit. We only remind the reader of the standard discretisation of the right hand side (Coriolis force). The ODE

$$\frac{du}{dt} = \omega v, \quad \frac{dv}{dt} = -\omega u,$$

is solved by means of the  $\theta$  scheme

$$\frac{u^{n+1} - u^n}{\Delta t} = \omega(\theta_u v^n + (1 - \theta_u)v^{n+1}), \quad \frac{v^{n+1} - v^n}{\Delta t} = -\omega(\theta_v u^n + (1 - \theta_v)u^{n+1})$$

with  $\theta_u + \theta_v \leq 1$ . Here we choose  $\theta_u = 1$  and  $\theta_v = 0$  so that the system is solved explicitly.

The time step is chosen following [2] such that

$$\Delta t^n \leq \min \left\{ \frac{2}{\omega}, \frac{\min(\Delta x, \Delta y)}{\max_D \|\mathbf{u}_D^n\| + \sqrt{gh_D^n}} \right\}.$$

## 2. NUMERICAL ASSESSMENTS OF THE SCHEMES

### 2.1. Setting of parameters

Here the parameters are set as follows:

- $\Delta x = \Delta y = 10^{-2}$ ;
- $\gamma = \nu = 1$ ;
- $g = 9.81$ ;
- $\omega = 1$ ;
- Final time  $T = 5$ .

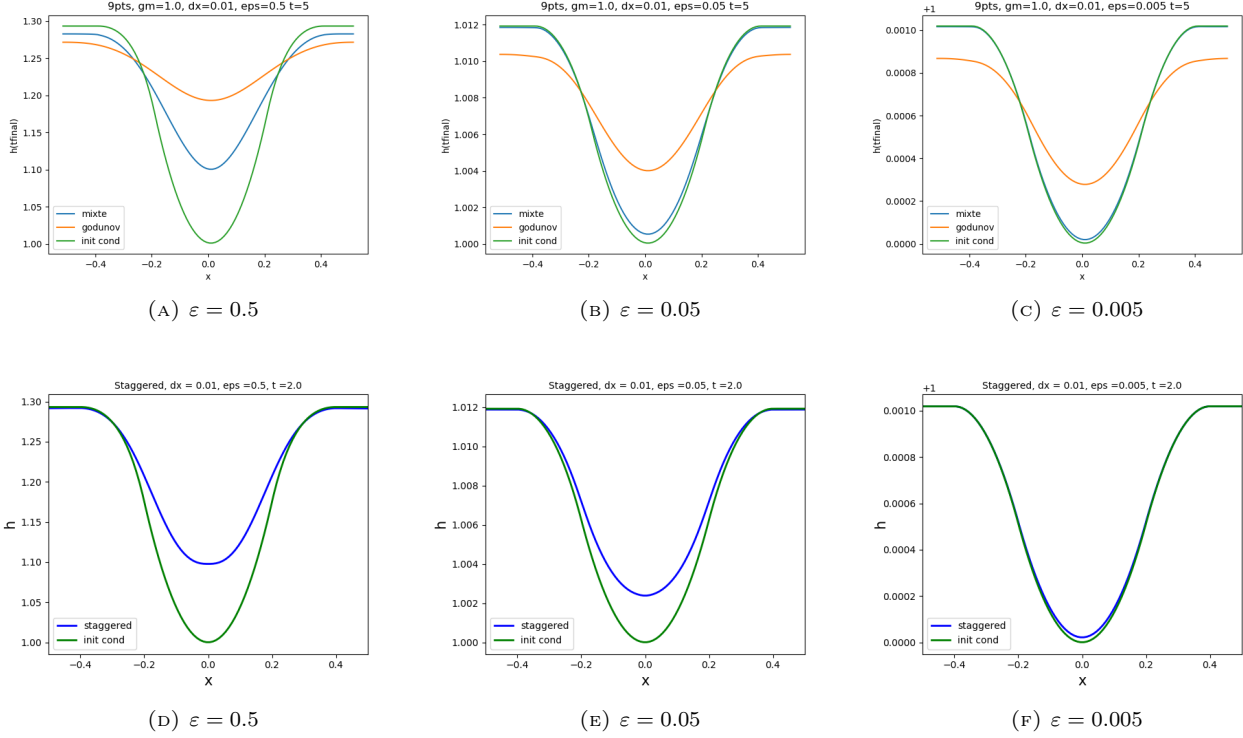


FIGURE 3. Stationary vortex: cross-section of the water height at final time for different values of  $\varepsilon$ ; comparisons between the exact solution, the numerical solution obtained with the classic Godunov scheme, the collocated scheme (12) (A-B-C) and the staggered scheme (25) (D-E-F)

## 2.2. Presentation of the test cases

To assess the numerical schemes designed in this work, we consider the stationary vortex described in [3]. Set in the domain  $\Omega = [-0.5, 0.5] \times [-0.5, 0.5]$ , the equations are supplemented with the initial velocity field given in polar coordinates

$$\mathbf{u}^0(r, \theta) = \nu_\theta(r) \mathbf{e}_\theta \quad \text{where} \quad \nu_\theta(r) = \varepsilon [5r \mathbf{1}_{\{r < 1/5\}} + (2 - 5r) \mathbf{1}_{\{1/5 \leq r < 2/5\}}].$$

The initial water height is then computed so that the steady version of the nonlinear equations is satisfied, *i.e.* such that

$$\partial_r h^0(r, \theta) = \frac{\nu_\theta(r)}{g} \left( \omega + \frac{\nu_\theta(r)}{r} \right).$$

As mentioned in [2], this case induces a Froude number of order  $\mathcal{O}(\varepsilon)$ .

## 2.3. Numerical results

Given the initial condition prescribed in § 2.2, we compare the robustness of several numerical schemes. The data are a steady solution of the nonlinear equations but their approximation on the grid is not necessarily preserved by any numerical scheme. We remind that the present paper was aimed at designing a scheme that is robust in the low Froude number regime by means of diffusion terms based on linear equilibria.

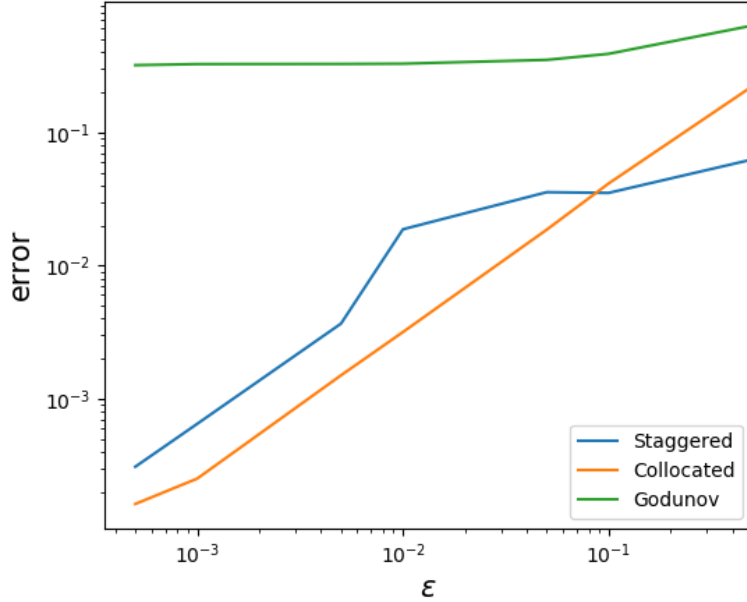


FIGURE 4. Error in  $L^2$  norm between the numerical solution and the exact solution at final time

We observe on Figure 3 that the smaller  $\varepsilon$ , the more accurate the solution obtained with the collocated scheme (12) while the accuracy of the standard Godunov scheme is barely improved.

This is confirmed by Fig. 4 where a kind of relative error is plotted:

$$E(\varepsilon) = \frac{\sum_i (h_K^N - h_K^0)^2}{\sum_i (h_K^0 - \bar{h})^2}$$

where  $N$  is the index of the final time and  $\bar{h} = \max_K h_K^0$ . The Godunov scheme broadly provides results of which the accuracy is independent from  $\varepsilon$  while the modifications proposed in this paper leads to a second order scheme until the accuracy is damaged by the truncation error driven by  $\Delta x$ .

These results tend to show that the target is reached insofar as the resulting scheme is robust in the low Froude number regime where the standard Godunov scheme was inaccurate. We mention that both corrections terms are necessary to obtain stable results which implies  $\gamma > 0$  and  $\nu > 0$ .

### 3. CONCLUSION

We managed to construct some corrections to the standard Godunov scheme so that the resulting collocated and staggered schemes are accurate and energy-decreasing around the geostrophic equilibrium in the low Froude number regime. The stationary vortex test case go to show the improvement due to these new terms.

The proofs are given in the semi-discrete case. In particular, the semi-discrete energy inequality requires  $\gamma \geq 0$  and  $\nu \geq 0$  so that it decreases. But numerical experiments showed that it is necessary to have them nonzero. Future works will thus deal with the stability analysis of the fully discrete version. Moreover, further numerical investigations are needed to ensure the schemes behave well in a large range of applications.

The authors express their deep gratitude to the French National Research Agency project NABUCO, grant ANR-17-CE40-0025 and the french INSU-CNRS (Institut National des Sciences de l'Univers – Centre National de la Recherche Scientifique) program LEFE-MANU (Méthodes Mathématiques et Numériques), project DWAVE for the funding of the project, as well as to the organizing committee of the CEMRACS'19 event. The authors also thank Martin Parisot for fruitful discussions.

## REFERENCES

- [1] G. Ansanay-Alex, F. Babik, J.-C. Latché, and D. Vola. An  $L^2$ -stable approximation of the Navier–Stokes convection operator for low-order non-conforming finite elements. *International Journal for Numerical Methods in Fluids*, 66(5):555–580, 2011.
- [2] E. Audusse, M.H. Do, P. Omnes, and Y. Penel. Analysis of modified Godunov type schemes for the two-dimensional linear wave equation with Coriolis source term on cartesian meshes. *Journal of Computational Physics*, 373:91–129, 2018.
- [3] E. Audusse, R. Klein, and A. Owinoh. Conservative discretization of Coriolis force in a finite volume framework. *Journal of Computational Physics*, 228(8):2934–2950, 2009.
- [4] A. Bermudez and M.E. Vazquez-Cendon. Upwind Methods for Hyperbolic Conservation Laws with Source Terms. *Computers and Fluids*, 23(8):1049–1071, 1994.
- [5] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhauser, 2004.
- [6] F. Bouchut, J. Le Sommer, and V. Zeitlin. Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. Part 2. High-resolution numerical simulations. *Journal of Fluid Mechanics*, 514:35–63, 2004.
- [7] F. Couderc, A. Duran, and J.-P. Vila. An explicit asymptotic preserving low froude scheme for the multilayer shallow water model with density stratification. *Journal of Computational Physics*, 343:235–270, 2017.
- [8] A. Duran, J.-P. Vila, and R. Baraille. Energy-stable staggered schemes for the shallow water equations. <https://hal.archives-ouvertes.fr/hal-01988382/>, January 2019.
- [9] L. Gosse. *Computing qualitatively correct approximations of balance laws*, volume 2. Springer, 2013.
- [10] J.M. Greenberg and A.Y. LeRoux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.
- [11] R. Herbin, W. Kheriji, and J.-C. Latché. On some implicit and semi-implicit staggered schemes for the shallow water equations. *ESAIM:M2AN*, 48:1807–1857, 2014.
- [12] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM Journal on Scientific Computing*, 21(2):441–454, 1999.
- [13] X. Liu, A. Chertock, and A. Kurganov. An asymptotic preserving scheme for the two-dimensional shallow water equations with Coriolis forces. *Journal of Computational Physics*, 391:259–279, 2019.
- [14] M. Parisot and J.-P. Vila. Centered-potential regularization for the advection upstream splitting method. *SIAM Journal on Numerical Analysis*, 54(5):3083–3104, 2015.
- [15] H. Zakerzadeh. The rs-imex scheme for the rotating shallow water equations with the coriolis force. In C. Cancès and P. Omnes, editors, *Finite Volumes for Complex Applications VIII - Methods and Theoretical Aspects*, pages 199–207. Springer, 2017.
- [16] V. Zeitlin. *Geophysical fluid dynamics: understanding (almost) everything with rotating shallow water models*. Oxford University Press, 2018.