



**HAL**  
open science

# Analysis of dissipation operators that damp spurious modes while maintaining discrete approximate geostrophic equilibriums for the B-grid staggered scheme on triangular meshes

Minh-Hieu Do, Van-Thanh Nguyen, Pascal Omnes

► **To cite this version:**

Minh-Hieu Do, Van-Thanh Nguyen, Pascal Omnes. Analysis of dissipation operators that damp spurious modes while maintaining discrete approximate geostrophic equilibriums for the B-grid staggered scheme on triangular meshes. *Journal of Computational Physics*, 2023, 10.1016/j.jcp.2023.112261 . hal-03834155v2

**HAL Id: hal-03834155**

**<https://hal.science/hal-03834155v2>**

Submitted on 14 Jun 2023

**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.

# Analysis of dissipation operators that damp spurious modes while maintaining discrete approximate geostrophic equilibria for the B-grid staggered scheme on triangular meshes

Minh-Hieu Do<sup>a</sup>, Van-Thanh Nguyen<sup>b</sup>, Pascal Omnes<sup>c,d,\*</sup>

<sup>a</sup>CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, 75016 Paris, France

<sup>b</sup>INRIA Paris, Sorbonne Université, Laboratoire Jacques-Louis Lions, équipe ALPINES, France

<sup>c</sup>Université Sorbonne Paris Nord, LAGA, CNRS UMR 7539, Institut Galilée, 99 Avenue J.-B. Clément, 93430 Villetaneuse Cedex, France

<sup>d</sup>Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation, 91191, Gif-sur-Yvette, France.

---

## Abstract

We consider the lowest order - so called B-grid - discretization on triangular cells which is common in the ocean or atmosphere simulation science, with pressure field at the vertices of the triangles and the velocity field at the cell centers. It is known that this discretization presents spurious modes due to a too large number of velocities unknowns and that these modes should be damped; since preservation of approximate geostrophic equilibrium (GE) may be of primary importance in the simulations, we propose and analyze damping operators that, when applied to the linearized constant coefficient configuration, exactly preserve discrete versions of the GEs. We also analyze their effects in more general configurations and their possible coupling with more standard damping operators. These strategies cover triangular variants of the Low Froude (LF) scheme [1], the Apparent Topography (AT) scheme [2, 3], as well as a new modification of it. The LF scheme has no numerical diffusion terms in the pressure equation and damps the normal velocity jumps between adjacent cells in the momentum equation. In the AT scheme, numerical diffusion is added in the pressure equation in such a way that it does not impact the discrete GE. However, since the AT scheme alone cannot be proved to be damping through energy estimates and does not preserve the space which is orthogonal to the discrete GE, we also study what we call the Modified Apparent Topography (MAT) that satisfies these important properties. We also suggest semi-implicit time discretizations that preserve these properties. Extensions to the full non-linear equations are also provided. Numerical experiments illustrate the performance of the new schemes.

*Keywords:* Shallow water equation with Coriolis source term, low Froude number, Hodge decomposition, staggered schemes, spurious mode damping, geostrophic equilibrium

---

## 1. Introduction

Large-scale geophysical flows are those in which the earth's rotation has a significant impact on the dynamics [4]; in such configurations, oceanic and atmospheric circulations are often perturbations of the so-called geostrophic equilibrium (GE) which results from the balance between the pressure gradient and the Coriolis force [4, 5]. When such circulations are investigated by numerical approximations of the corresponding physical model, it is therefore essential that the numerical schemes do not completely damp out this approximate equilibrium. Although the GE is not a perfect steady-state in real dynamics because of forcing and non-linearities, there is some interest in studying whether numerical schemes preserve it in the linear case, for the reasons exposed below.

---

\*Corresponding author

Email addresses: do@ceremade.dauphine.fr (Minh-Hieu Do), van-thanh.nguyen@inria.fr (Van-Thanh Nguyen), pascal.omnes@cea.fr (Pascal Omnes)

We begin with the investigation of the dimensionless shallow water system of equations on the rotating frame, which is given by

$$\begin{cases} St \partial_t h + \nabla \cdot (h \bar{\mathbf{u}}) = 0, \\ St \partial_t (h \bar{\mathbf{u}}) + \nabla \cdot (h \bar{\mathbf{u}} \otimes \bar{\mathbf{u}}) + \frac{1}{Fr^2} \nabla \left( \frac{h^2}{2} \right) = -\frac{1}{Fr^2} h \nabla b - \frac{1}{Ro} h \bar{\mathbf{u}}^\perp. \end{cases} \quad (1)$$

In System (1) unknowns  $h$  and  $\bar{\mathbf{u}}$  respectively denote the water or layer depth and the velocity of the fluid and function  $b(x)$  denotes the topography and is a given function. Dimensionless numbers  $St$ ,  $Fr$  and  $Ro$  respectively stand for the Strouhal, the Froude and the Rossby numbers. As mentioned above, we are interested in large scale flows, so we will focus on the case in which the Froude and Rossby numbers are small. In particular, we shall assume that

$$Ro = \mathcal{O}(M) \quad \text{and} \quad Fr = \mathcal{O}(M)$$

with  $M$  is a small parameter (typical values lead to  $M \sim 10^{-2}$ ). For a Strouhal number of order  $\mathcal{O}(\frac{1}{M})$ , *i.e.* for small time scale, the solution of system (1) with flat topography is  $h = h_0 + Mr + \mathcal{O}(M^2)$  and  $\bar{\mathbf{u}} = \mathbf{u} + \mathcal{O}(M)$ , where  $h_0$  is a space-time constant and where, up to multiplicative constants,  $(r, \mathbf{u})$  satisfies the linear wave equation with Coriolis source term

$$\begin{cases} \partial_t r + a_* \nabla \cdot \mathbf{u} = 0 \\ \partial_t \mathbf{u} + a_* \nabla r = -\omega \mathbf{u}^\perp \end{cases} \quad (2)$$

where  $\mathbf{u} = (u, v)^T$ , and  $\mathbf{u}^\perp = (-v, u)^T$ . The parameters  $a_*$  and  $\omega$  are of order one, respectively related to the wave velocity and the rotating velocity. In case of constant coefficients, System (2) admits as stationary state the GE which is given by

$$a_* \nabla r = -\omega \mathbf{u}^\perp, \quad (3)$$

with incompressibility of the velocity field

$$\nabla \cdot \mathbf{u} = 0. \quad (4)$$

As a consequence, a good reason for analyzing schemes that can preserve a discrete version of the GE in the linearized constant coefficient case, is that this is the minimum requirement that a scheme should have in order to get accurate simulations in more general situations around the GE. We shall then mention how the schemes we propose may be extended to more general cases.

There are few works related to numerical schemes that preserve a discrete version of geostrophic equilibriums. Since Godunov type finite volume collocated schemes (all unknowns are located at the centers of the mesh cells and fluxes are computed using a Riemann solver) have a long and successful history for the numerical approximation of systems of conservation laws such as (1) with  $\omega = 0$ , it is quite natural to first try to extend them to the treatment of the Coriolis source term ( $\omega \neq 0$ ). Some success in this direction on Cartesian grids was previously obtained in [6], but we show in a discussion that is postponed to the Appendix that the direct extension of these schemes to triangular meshes seems to lead to a dead end (we recall that triangular grids may be preferred to Cartesian grids for example because they are much more flexible for accurate discretizations of complex detailed geometries and for local mesh refinement). Instead, we investigate in the present contribution the lowest-order ( $k = 0$ ) staggered arrangement belonging to the well-studied family of schemes with completely discontinuous  $P_k^{\text{disc}}$  velocities and continuous Lagrange  $P_{k+1}$  pressures. As is described in [7], this configuration is very natural for the preservation of a discrete version of (3), since discrete pressure gradients belong to the velocity space. An additional useful property in the lowest order case is that discretely divergence free velocity fields can be characterized as the curls of  $P_1$  Lagrange functions defined by their values at the vertices of the triangular cells. Finally, our choice of the  $P_0 - P_1$  configuration is supported by [8] which stresses that, among a large study of finite element pairs, this is an appealing choice (see the Conclusion of [8] for details).

This, as well as the fact that such a staggered strategy (cell-centered velocities and node based pressures) is proposed and analyzed in the finite volume coastal ocean model (FVCOM) [9]) and actually used for large-scale ocean modeling [10, 11], leads us to the main purpose of this work, which is to propose a family of staggered explicit schemes with damping terms on triangular cells in which pressures and velocities that verify discrete versions of (3)–(4) are not affected by the numerical diffusion of the scheme. This numerical diffusion is necessary for at least

two reasons: it maintains the scheme stability when an explicit (or semi-implicit) discretization in time is used, and it helps damp spurious numerical modes that are associated with this staggered configuration, as analyzed for example in [12] (together with many other finite element pairs) and in [13]. In the latter, discrete velocity Laplacians are used not only to damp the spurious modes, but also to damp variance on small scales. Full velocity Laplacians have the drawback of destroying GEs in circumstances in which the flow takes place around this approximate equilibrium, but may be needed in other circumstances, for example for turbulent forced-dissipative flows. In the present work, the schemes we propose preserve discrete GEs while damping the spurious modes; in order to perform this, they do not use full velocity Laplacians, but only the normal direction of the velocity jumps at the cell interfaces. We indicate in Section 5 of the article how this could be coupled to diffusion of the tangential velocity jumps to recover standard full Laplacians when they are needed. A by-product of the analysis is that we are able to characterize the spurious modes on general triangular grids: they are velocities which have vanishing discrete curl and divergence, while not being zero. Additionally, the family of schemes we propose also verifies points 2 to 6 mentioned in [14]: mass conservation, curl-free pressure gradient, energy conserving pressure terms, energy conserving Coriolis term and steady geostrophic modes. An additional feature which, to our knowledge, has not been discussed previously in the literature at the discrete level, is the rigorous definition of a splitting between the vector subspace of discrete GEs and an orthogonal subspace containing their perturbations. In the linearized case, some of the schemes we present not only preserve discrete GEs, but also the orthogonal subspace: whenever the initial condition is orthogonal to the space of GEs, then so is the solution at any time; this ensures that perturbations around GEs do not exchange energy with the space of GEs. Finally, we also show the stability of two of the proposed schemes.

This work is organized as follows. We propose some semi-discrete (continuous in time) staggered schemes in Section 2:

- The Low Froude staggered scheme (LF),
- The Apparent Topography staggered scheme (AT),
- The Modified Apparent Topography scheme (MAT).

Particularly, we show that the discrete differential operators that appear in the scheme definition satisfy mimetic properties that lead to the desirable properties mentioned above. The mathematical properties mentioned in Sections 2.4 to 2.6 mainly concern the schemes in the linearized, constant coefficient case; this is interesting per se from a mathematical point of view and gives hints on what may happen in more general configurations. Stress is then added on the damping properties of the schemes we propose. We then take into account in Section 3 the time discretization to introduce appropriate fully discrete schemes that still possess the same properties as the semi-discrete schemes. Some numerical test cases are shown in Section 4 to confirm the analysis led in the theoretical part and we discuss an extension to the case of non constant Coriolis parameter. Then, further extensions to non-flat topography and non-linear configurations are discussed and tested in Section 5, in which it is also shown that combining the diffusive operators we propose with some additional velocity diffusion allows to recover usual full velocity Laplacian operators.

Concluding remarks complete the study in Section 6, and in Appendix A, we briefly present the problem encountered with the collocated Godunov scheme on triangular meshes when trying to check whether the scheme steady-states verify discrete versions of the GE.

## 2. Analysis of the semi-discrete staggered schemes in the linearized constant coefficient case

### 2.1. Definition of the discrete operators and the semi-discrete staggered scheme

We first define the discrete version of the gradient and divergence operators. As mentioned above, the discrete gradient of a scalar field will be defined over the cells, from values of the scalar field defined at the vertices. The discrete divergence operator will then be defined such that a discrete integration by parts holds; in this way we shall be able to prove stability of the scheme by energy estimates.

We shall denote by  $T_i$  a generic triangular cell used for the discretization of the computational domain which is supposed to be rectangular, periodic with no internal hole, for the sake of simplicity. Let  $A_{ij}$  be the common edge of the neighboring cells  $T_i$  and  $T_j$ , and  $\mathbf{n}_{ij}$  be the unit normal vector to  $A_{ij}$  pointing from  $T_i$  to  $T_j$ . Moreover, we also denote the area of the cell  $T_i$  by  $|T_i|$  and the length of the edge  $A_{ij}$  by  $|A_{ij}|$ .

Let  $(r_k)_{k \in [1, N_r]}$  be a discrete scalar field defined by its values at the vertices of the mesh, where  $N_r$  is the number of vertices. Let us denote by  $r_h$  the globally continuous function that is piecewise  $P_1$  on each cell and defined by the values  $(r_k)_{k \in [1, N_r]}$  at the vertices of the mesh. We can define the discrete gradient on the (primal) cells by using the following formula

$$(\nabla_h^T r_h)_i = \frac{1}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| \frac{r_h(S_{ij}) + r_h(N_{ij})}{2} \mathbf{n}_{ij}, \quad (5)$$

where for any edge  $A_{ij}$ , we have denoted by  $N_{ij}$  and  $S_{ij}$  its extremities (see Figure 1(b)). It is easily checked that (5) is the gradient of the  $P_1$  Lagrange function  $r_h$ , so that this definition is quite standard.

Let  $(\mathbf{u}_i)_{i \in [1, N]}$  be a discrete vector field defined by its values on the triangular cells. We shall define its divergence on the so-called barycentric dual mesh constructed as follows. Each vertex is associated with a dual cell obtained by joining the barycenters of the cells which share the vertex to the midpoints of the edges (see Figure 1(a)). Let us denote the area of the dual cell  $D_k$  by  $|D_k|$ . Then, we can define the discrete divergence  $(\nabla_h^D \cdot \mathbf{u}_h)$  on the dual cell by the following (also quite standard) formula

$$(\nabla_h^D \cdot \mathbf{u}_h)_k = \frac{1}{|D_k|} \sum_{T_i | T_i \cap D_k \neq \emptyset} \mathbf{u}_i \cdot \frac{1}{2} l_{ik} \mathbf{n}_{ik}, \quad (6)$$

where for a triangle  $T_i$  and a vertex  $k$ ,  $l_{ik}$  is the length of the edge of  $T_i$  that is opposite to vertex  $k$  and  $\mathbf{n}_{ik}$  is the unit normal vector pointing outside  $T_i$  on this edge. This edge will be named  $A_{ik}$  below.

Moreover, we can define the discrete curl of the vector field by

$$(\nabla_h^D \times \mathbf{u}_h)_k = -(\nabla_h^D \cdot \mathbf{u}_h^\perp)_k = -\frac{1}{|D_k|} \sum_{T_i | T_i \cap D_k \neq \emptyset} \mathbf{u}_i^\perp \cdot \frac{1}{2} l_{ik} \mathbf{n}_{ik}. \quad (7)$$

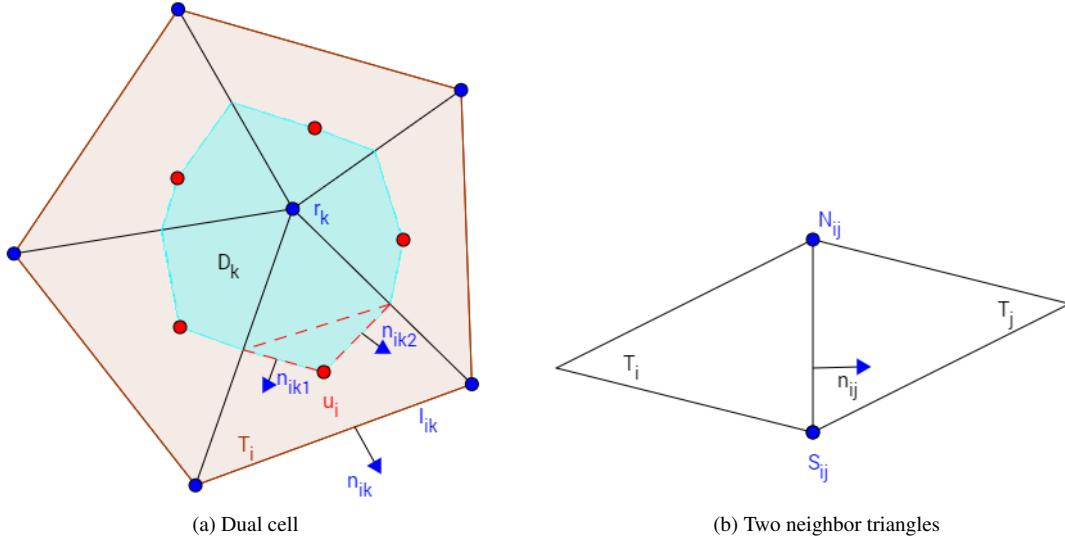


Figure 1: Staggered scheme.

On the other hand, we can define the discrete scalar product between  $q_h^1 = (r_h^1, \mathbf{u}_h^1)$  and  $q_h^2 = (r_h^2, \mathbf{u}_h^2)$  by

$$\langle q_h^1, q_h^2 \rangle := \langle r_h^1, r_h^2 \rangle_{\mathcal{D}} + \langle \mathbf{u}_h^1, \mathbf{u}_h^2 \rangle_{\mathcal{P}} := \sum_{k=1}^{N_r} |D_k| r_k^1 r_k^2 + \sum_{i=1}^N |T_i| \mathbf{u}_i^1 \cdot \mathbf{u}_i^2. \quad (8)$$

With these discrete operators, we propose the following family of semi-discrete staggered schemes which can be applied to the linear wave equation with Coriolis source term:

$$\begin{cases} \frac{d}{dt} r_k(t) + a_\star (\nabla_h^D \cdot \mathbf{u}_h)_k - \nabla_h^D \cdot \left[ \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right]_k = 0, \\ \frac{d}{dt} \mathbf{u}_i(t) + a_\star (\nabla_h^T r_h)_i - \nu_u (\nabla^{\text{CR}} [\mathbf{u}_h \cdot \mathbf{n}])_i = -\omega \mathbf{u}_i^\perp + \frac{\omega}{a_\star} (\gamma_r)_i \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right)_i^\perp, \end{cases} \quad (9)$$

where  $(\gamma_r)_i, (\nu_r)_i = \frac{\kappa_r a_\star h_i}{2}$ , with  $h_i$  being a length related to triangle  $T_i$ , e.g. its circumradius, and  $\nu_u = \frac{\kappa_u a_\star}{2}$  represent the parameters that control the diffusion terms and  $(\nabla^{\text{CR}} [\mathbf{u}_h \cdot \mathbf{n}])_i$  is given by

$$(\nabla^{\text{CR}} [\mathbf{u}_h \cdot \mathbf{n}])_i := \frac{1}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| [(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ij}] \mathbf{n}_{ij}. \quad (10)$$

The classical values for  $\kappa_r$  and  $\kappa_u$  are either 0 (no damping) or 1 (damping). The term defined in (10) is the standard diffusive upwinding term used in the collocated Godunov scheme in the velocity equation. Note that this diffusion term only acts on the jumps of the normal velocities, contrarily to a full velocity Laplacian diffusion term. This difference is essential in the preservation of discretely divergence free velocity fields that have continuous normal components across cell interfaces, and are thus unaffected by this diffusion term. The notation  $\nabla^{\text{CR}}$  is here to recall that the right-hand side in (10) is the gradient of the Crouzeix-Raviart non-conforming  $P_1$  function with value  $[(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ij}]$  at the midpoint of the edge  $A_{ij}$ . We also note that the Low Froude (LF) staggered scheme corresponds to  $\nu_r = \gamma_r = 0$ , the classical Apparent Topography (AT) scheme corresponds to  $\nu_r > 0, \gamma_r = 0$  and the Modified Apparent Topography (MAT) to  $\nu_r = \gamma_r > 0$ .

**Remark 1.** In Scheme (9), it is stressed that the additional terms all vanish when applied to a geostrophic equilibrium. This is obvious for the terms that contain  $\nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp$  because the definition of the GE is exactly that this term vanishes. Moreover, the discrete GE implies that  $\mathbf{u}_h$  is the curl of a  $P_1$  continuous Lagrange function on the triangles; these curls have continuous normal components on the edges of the mesh, which implies that the jump term  $[\mathbf{u}_h \cdot \mathbf{n}]$  also vanishes for such a discrete GE. More details are provided in Section 2.4.

## 2.2. Properties of the discrete operators

**Proposition 1.** With the discrete divergence, curl, gradient and scalar product defined respectively by (6), (7), (5) and (8), we have the following properties for the semi-discrete staggered scheme (9) (These properties are not new findings of this work; they are repeated here without proof for the sake of completeness):

i. Discrete integration by parts (energy conservation for the pressure gradient force)

$$\langle \nabla_h^D \cdot \mathbf{u}_h, r_h \rangle_{\mathcal{D}} = -\langle \nabla_h^T r_h, \mathbf{u}_h \rangle_{\mathcal{P}}. \quad (11)$$

ii. Energy conservation for the Coriolis force

$$\langle \mathbf{u}_h^\perp, \mathbf{u}_h \rangle_{\mathcal{P}} = 0. \quad (12)$$

iii. No vorticity production for the pressure gradient force

$$\nabla_h^D \times (\nabla_h^T r_h) = 0. \quad (13)$$

## 2.3. Evolution of the discrete energy

**Proposition 2.** With  $\nu_r = \gamma_r$  and the discrete energy defined with the following expression

$$E_h(t) = \langle q_h, q_h \rangle = \sum_{k=1}^{N_r} |D_k| r_k^2 + \sum_{i=1}^N |T_i| |\mathbf{u}_i|^2,$$

we obtain

$$\frac{d}{dt} E_h(t) \leq 0$$

which means that the Low Froude and Modified Apparent Topography semi-discrete staggered scheme dissipate the discrete energy.

*Proof.* We take the scalar product of the staggered scheme (9) with  $q_h = (r_h, \mathbf{u}_h)$  in order to obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} E_h(t) &= -a_\star \langle \nabla_h^D \cdot \mathbf{u}_h, r_h \rangle_{\mathcal{D}} - a_\star \langle \nabla_h^T r_h, \mathbf{u}_h \rangle_{\mathcal{P}} - \omega \langle \mathbf{u}_h^\perp, \mathbf{u}_h \rangle_{\mathcal{P}} + \nu_u \langle \nabla^{\text{CR}}[\mathbf{u}_h \cdot \mathbf{n}], \mathbf{u}_h \rangle_{\mathcal{P}} \\ &\quad + \left\langle \nabla_h^D \cdot \left[ \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right], r_h \right\rangle_{\mathcal{D}} + \frac{\omega}{a_\star} \left\langle \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right)^\perp, \mathbf{u}_h \right\rangle_{\mathcal{P}}. \end{aligned}$$

Classically (see [15]), the term  $\langle \nabla^{\text{CR}}[\mathbf{u}_h \cdot \mathbf{n}], \mathbf{u}_h \rangle_{\mathcal{P}}$  is equal to  $-\sum_{A_{ij}} |A_{ij}| |(\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}|^2$ ; then, using (11) and (12), and denoting by  $\|\cdot\|_{\mathcal{P}}$  the norm associated to the scalar product on the primal mesh, we get

$$\frac{d}{dt} E_h(t) = -2\nu_u \sum_{A_{ij}} |A_{ij}| |(\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}|^2 - 2 \left\| \sqrt{\nu_r} \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right\|_{\mathcal{P}}^2, \quad (14)$$

which leads to the conclusion.  $\square$

**Remark 2.** *The fact that  $\gamma_r = \nu_r$  is essential in the calculation above. For this reason, the usual Apparent Topography scheme ( $\nu_r > 0, \gamma_r = 0$ ) cannot be proved to dissipate energy with this kind of energy estimation (which does not mean that it does not).*

#### 2.4. Analysis of the discretized steady-states and their orthogonal subspace

We now define a set of discretized steady-states with staggered variables on triangular meshes by the following expression

$$\mathcal{E}_{\omega \neq 0}^\Delta = \left\{ \hat{q}_h = (\hat{r}_h, \hat{\mathbf{u}}_h) \in \mathbb{R}^{N_r} \times \mathbb{R}^{2N} : a_\star (\nabla_h^T \hat{r}_h)_i = -\omega \hat{\mathbf{u}}_i^\perp \right\} \quad (15)$$

which is a consistent discretization of the GE (3). Defining further a set  $\mathcal{A}_h$  by

$$\mathcal{A}_h := \left\{ q_h = (r_h, \mathbf{u}_h) \in \mathbb{R}^{N_r} \times \mathbb{R}^{2N} : a_\star (\nabla_h^D \times \mathbf{u}_h)_k = \omega r_k \right\}, \quad (16)$$

we have the following discrete orthogonal decomposition:

**Proposition 3.** *The orthogonal space of  $\mathcal{E}_{\omega \neq 0}^\Delta$  is given by*

$$\mathcal{E}_{\omega \neq 0}^{\Delta, \perp} = \mathcal{A}_h. \quad (17)$$

*Proof.* For each  $\hat{q}_h = (\hat{r}_h, \hat{\mathbf{u}}_h) \in \mathcal{E}_{\omega \neq 0}^\Delta$  and arbitrary  $\tilde{q}_h = (\tilde{r}_h, \tilde{\mathbf{u}}_h) \in \mathbb{R}^{N_r} \times \mathbb{R}^{2N}$ , we use the discrete integration by parts formula (11) to obtain

$$\begin{aligned} \langle \hat{q}_h, \tilde{q}_h \rangle &= \langle \hat{r}_h, \tilde{r}_h \rangle_{\mathcal{D}} + \langle \hat{\mathbf{u}}_h, \tilde{\mathbf{u}}_h \rangle_{\mathcal{P}} = \langle \hat{r}_h, \tilde{r}_h \rangle_{\mathcal{D}} + \langle \hat{\mathbf{u}}_h^\perp, \tilde{\mathbf{u}}_h^\perp \rangle_{\mathcal{P}} \\ &= \langle \hat{r}_h, \tilde{r}_h \rangle_{\mathcal{D}} - \frac{a_\star}{\omega} \left\langle \nabla_h^T \hat{r}_h, \tilde{\mathbf{u}}_h^\perp \right\rangle_{\mathcal{P}} = \langle \hat{r}_h, \tilde{r}_h \rangle_{\mathcal{D}} + \frac{a_\star}{\omega} \left\langle \hat{r}_h, \nabla_h^D \cdot \tilde{\mathbf{u}}_h^\perp \right\rangle_{\mathcal{D}} \\ &= \left\langle \hat{r}_h, \tilde{r}_h - \frac{a_\star}{\omega} \nabla_h^D \times \tilde{\mathbf{u}}_h \right\rangle_{\mathcal{D}}. \end{aligned}$$

Hence, if  $\tilde{q}_h \in \mathcal{A}_h$ , we obviously have  $\langle \hat{q}_h, \tilde{q}_h \rangle = 0$  which leads to  $\mathcal{A}_h \subset \mathcal{E}_{\omega \neq 0}^{\Delta, \perp}$ . On the other hand, since  $\hat{r}_h$  can be arbitrary in  $\mathbb{R}^{N_r}$  when  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$ , then the equality  $\langle \hat{r}_h, \tilde{r}_h - \frac{a_\star}{\omega} \nabla_h^D \times \tilde{\mathbf{u}}_h \rangle_{\mathcal{D}} = 0$  for all  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$  implies that  $\tilde{r}_h - \frac{a_\star}{\omega} \nabla_h^D \times \tilde{\mathbf{u}}_h = 0$  and thus  $\tilde{q}_h \in \mathcal{A}_h$ . It follows that  $\mathcal{E}_{\omega \neq 0}^{\Delta, \perp} \subset \mathcal{A}_h$ .  $\square$

**Remark 3.** The discrete orthogonal decomposition allows us to define the following discrete orthogonal projection

$$\mathbb{P}_h : \begin{cases} \mathbb{R}^{N_r+2N} & \longrightarrow \mathcal{E}_{\omega \neq 0}^\Delta \\ q_h & \longmapsto \hat{q}_h \end{cases}$$

and we can construct  $\hat{q}_h$  by what follows:

Let  $q_h = (r_h, \mathbf{u}_h)$  be given in  $\mathbb{R}^{N_r+2N}$ . For all  $(\hat{p}_h, \hat{\mathbf{v}}_h) \in \mathcal{E}_{\omega \neq 0}^\Delta$ , using orthogonality, we have

$$\langle \hat{r}_h, \hat{p}_h \rangle_{\mathcal{D}} + \langle \hat{\mathbf{u}}_h, \hat{\mathbf{v}}_h \rangle_{\mathcal{P}} = \langle r_h, \hat{p}_h \rangle_{\mathcal{D}} + \langle \mathbf{u}_h, \hat{\mathbf{v}}_h \rangle_{\mathcal{P}}.$$

We then use the definition of the discrete steady-states and the discrete integration by parts formula to get

$$\langle \hat{r}_h, \hat{p}_h \rangle_{\mathcal{D}} - \left( \frac{a_\star}{\omega} \right)^2 \langle \nabla_h^D \cdot (\nabla_h^T \hat{r}_h), \hat{p}_h \rangle_{\mathcal{D}} = \langle r_h, \hat{p}_h \rangle_{\mathcal{D}} - \frac{a_\star}{\omega} \langle \nabla_h^D \times \mathbf{u}_h, \hat{p}_h \rangle_{\mathcal{D}}.$$

As a result, since  $\hat{p}_h$  can be arbitrary in  $\mathbb{R}^{N_r}$ , it is possible to find  $\hat{r}_h$  by solving the following linear system

$$\hat{r}_k - \left( \frac{a_\star}{\omega} \right)^2 [\nabla_h^D \cdot (\nabla_h^T \hat{r}_h)]_k = r_k - \frac{a_\star}{\omega} (\nabla_h^D \times \mathbf{u}_h)_k. \quad (18)$$

Then, by definition of the discrete steady-states, the part of the velocity field in  $\mathcal{E}_{\omega \neq 0}^\Delta$  is given by

$$\hat{\mathbf{u}}_i = \frac{a_\star}{\omega} (\nabla_h^T \hat{r}_h)_i^\perp.$$

Finally, the orthogonal component is simply given by  $\tilde{q}_h = q_h - \hat{q}_h$ . Moreover, the linear system (18) defines a unique solution since  $-\nabla_h^D \cdot$  and  $\nabla_h^T$  are adjoint operators as shown by the discrete integration by parts formula (11).

## 2.5. Well-balanced and orthogonality preserving properties

**Definition 1.** A semi-discrete scheme is said to be well-balanced with respect to system (2) if

$$q_h^0 \in \mathcal{E}_{\omega \neq 0}^\Delta \quad \Rightarrow \quad \forall t \geq 0, \quad q_h(t) = q_h^0 \in \mathcal{E}_{\omega \neq 0}^\Delta.$$

**Definition 2.** A semi-discrete scheme is said to be orthogonality preserving with respect to system (2) if

$$q_h^0 \in \mathcal{E}_{\omega \neq 0}^{\Delta, \perp} \quad \Rightarrow \quad \forall t \geq 0, \quad q_h(t) \in \mathcal{E}_{\omega \neq 0}^{\Delta, \perp}.$$

**Proposition 4.** We have:

- (i) The semi-discrete staggered type scheme (9) is a well-balanced scheme in the sense that it can capture the discrete steady states (15).
- (ii) The semi-discrete Low Froude ( $\nu_r = \gamma_r = 0$ ) and Modified Apparent Topography staggered scheme ( $\nu_r = \gamma_r > 0$ ) are orthogonality preserving schemes.

*Proof.* With the discrete steady state (15), the velocity field can be written as

$$\hat{\mathbf{u}}_i = \frac{a_\star}{\omega} (\nabla_h^T \hat{r}_h)_i^\perp. \quad (19)$$

Since we have no vorticity production of the gradient term (see (13)), we get

$$(\nabla_h^D \cdot \hat{\mathbf{u}}_h)_k = \frac{a_\star}{\omega} [\nabla_h^D \cdot (\nabla_h^T \hat{r}_h)_k^\perp] = -\frac{a_\star}{\omega} [\nabla_h^D \times (\nabla_h^T \hat{r}_h)]_k = 0. \quad (20)$$

On the other hand, since  $\hat{\mathbf{u}}_h$  is the curl of a discrete piecewise  $P_1$ , globally continuous function, then its normal jumps through edges vanish:

$$[[\hat{\mathbf{u}} \cdot \mathbf{n}]_{ij}] := (\hat{\mathbf{u}}_i - \hat{\mathbf{u}}_j) \cdot \mathbf{n}_{ij} = 0. \quad (21)$$



Therefore, the definition of the discrete kernel (15), the divergence free property (20) and the vanishing of the jumps of the velocity field through the edges  $A_{ij}$  (21) imply the well-balanced property of the semi-discrete staggered scheme (9). This proves Point (i).

In consideration of the orthogonality preserving property, by taking the discrete scalar product of the semi-discrete staggered scheme with any stationary state  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$ , we obtain when  $\nu_r = \gamma_r$

$$\begin{aligned} \left\langle \frac{d}{dt} q_h(t), \hat{q}_h \right\rangle &= -a_\star \langle \nabla_h^D \cdot \mathbf{u}_h, \hat{r}_h \rangle_{\mathcal{D}} + \left\langle \nabla_h^D \cdot \left[ \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right], \hat{r}_h \right\rangle_{\mathcal{D}} \\ &\quad - a_\star \langle \nabla_h^T r_h, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} + \nu_u \langle \nabla^{\text{CR}} \llbracket \mathbf{u}_h \cdot \mathbf{n} \rrbracket, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} - \omega \langle \mathbf{u}_h^\perp, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} \\ &\quad + \frac{\omega}{a_\star} \left\langle \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right)^\perp, \hat{\mathbf{u}}_h \right\rangle_{\mathcal{P}}. \end{aligned}$$

By using the discrete integration by parts formula and (20), we have

$$\langle \nabla_h^T r_h, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} = - \langle r_h, \nabla_h^D \cdot \hat{\mathbf{u}}_h \rangle_{\mathcal{D}} = 0.$$

Moreover, by simple calculations, we get, thanks to (21):

$$\begin{aligned} - \langle \nabla^{\text{CR}} \llbracket \mathbf{u}_h \cdot \mathbf{n} \rrbracket, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} &= \sum_{A_{ij}} |A_{ij}| [(\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}] \mathbf{n}_{ij} \cdot \hat{\mathbf{u}}_i + \sum_{A_{ij}} |A_{ij}| [(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ji}] \mathbf{n}_{ji} \cdot \hat{\mathbf{u}}_j \\ &= \sum_{A_{ij}} |A_{ij}| [(\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}] [(\hat{\mathbf{u}}_i - \hat{\mathbf{u}}_j) \cdot \mathbf{n}_{ij}] = 0. \end{aligned}$$

Using a final discrete integration by parts formula and the fact that  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$ , we get

$$-a_\star \langle \nabla_h^D \cdot \mathbf{u}_h, \hat{r}_h \rangle_{\mathcal{D}} - \omega \langle \mathbf{u}_h^\perp, \hat{\mathbf{u}}_h \rangle_{\mathcal{P}} = \left\langle a_\star \nabla_h^T \hat{r}_h + \omega \hat{\mathbf{u}}_h^\perp, \mathbf{u}_h \right\rangle_{\mathcal{P}} = 0.$$

On the other hand, using integration by parts,  $\forall \hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$  we get

$$\begin{aligned} \left\langle \nabla_h^D \cdot \left[ \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right], \hat{r}_h \right\rangle_{\mathcal{D}} + \frac{\omega}{a_\star} \left\langle \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right)^\perp, \hat{\mathbf{u}}_h \right\rangle_{\mathcal{P}} = \\ - \left\langle \nu_r \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right), \nabla_h^T \hat{r}_h + \frac{\omega}{a_\star} \hat{\mathbf{u}}_h^\perp \right\rangle = 0. \end{aligned}$$

Therefore, the semi-discrete staggered scheme is orthogonality preserving.  $\square$

**Remark 4.** Proposition 4 tells us that all considered schemes preserve GE because dissipation terms vanish for such steady states. However, the Low Froude and Modified Apparent Topography schemes have an additional advantage over the usual Apparent Topography scheme in that they operate a strict splitting between the GE and their orthogonal subspace. More precisely, by virtue of the orthogonal decomposition of vector fields, it is always possible to decompose the numerical solution into two parts:

$$q_h(t) = \hat{q}_h(t) + \tilde{q}_h(t)$$

with  $\hat{q}_h(t)$  in the GE space  $\mathcal{E}_{\omega \neq 0}^\Delta$  and  $\tilde{q}_h$  in the orthogonal space. Since the Low Froude and Modified Apparent Topography schemes are orthogonality preserving, their orthogonal parts  $\tilde{q}_h(t)$  are equal to the time evolution of  $\tilde{q}_h(0)$  for all time  $t$ . As a result, we can ensure that

$$\hat{q}_h(t) = \mathbb{P}_h q_h^0, \quad \forall t \geq 0$$

for both schemes. This means that the kernel part of their numerical solutions is always equal to the initial projection  $\mathbb{P}_h q_h^0$ , which is also the final state of the numerical solution when the orthogonal part is damped to zero by the

numerical diffusion of the schemes.

On the other hand, the classical Apparent Topography scheme is not orthogonality preserving because of its numerical diffusion on the pressure equation. As a consequence, the orthogonal part not only damps out, but also partly moves into the kernel. Therefore, the kernel part of this scheme may be changed at each time step until there is no energy left in the space orthogonal to the kernel.

## 2.6. Low Froude number accuracy

When both  $a_*$  and  $\omega$  scale like  $\frac{1}{M}$ , with  $M \ll 1$ , then the solution of (2) is typically a GE (an element in the kernel) plus an orthogonal perturbation of order  $M$ . We shall verify under which (sufficient) condition a scheme like (9) mimics this. More precisely, we set

**Definition 3.** Let  $q_{\nu,h}(t)$  be the solution of the semi-discrete scheme (9) with initial condition  $q_h^0$ . This scheme is said to be accurate at low Froude number if

$$\forall C_1 \in \mathbb{R}^+, \text{ if } \|q_h^0 - \mathbb{P}_h(q_h^0)\| = C_1 M, \text{ then } \|q_{\nu,h}(t) - \mathbb{P}_h(q_h^0)\| \leq C_1 M, \forall t \geq 0.$$

**Proposition 5.** With  $\nu_r = \gamma_r$ , the semi-discrete scheme (9) is accurate at low Froude number; this is in particular the case for the LF and the MAT schemes.

*Proof.* By linearity, the solution of semi-discrete staggered scheme  $q_{\nu,h}(t)$  can be written as

$$q_{\nu,h}(t) = q_{\nu,h}^a(t) + q_{\nu,h}^b(t)$$

where  $q_{\nu,h}^a(t)$  and  $q_{\nu,h}^b(t)$  are solutions of (9) with initial conditions respectively given by

$$q_{\nu,h}^a(0) = \mathbb{P}_h(q_h^0) \quad \text{and} \quad q_{\nu,h}^b(0) = q_h^0 - \mathbb{P}_h(q_h^0).$$

Then, we have

$$\|q_{\nu,h}(t) - \mathbb{P}_h(q_h^0)\| = \|q_{\nu,h}^a(t) + q_{\nu,h}^b(t) - \mathbb{P}_h(q_h^0)\| \leq \|q_{\nu,h}^a(t) - \mathbb{P}_h(q_h^0)\| + \|q_{\nu,h}^b(t)\|.$$

Moreover, when  $\nu_r = \gamma_r$ , the dissipation of the semi-discrete staggered scheme proved in Lemma 2 leads to the conclusion that  $\|q_{\nu,h}^b(t)\| \leq \|q_{\nu,h}^b(0)\|$ . For this reason, the accuracy of the scheme is linked to the behavior of  $q_{\nu,h}^a(t)$ . Since the semi-discrete scheme (9) is a well-balanced scheme, we obviously have  $q_{\nu,h}^a(t) = \mathbb{P}_h(q_h^0)$ . Therefore, we obtain

$$\forall t \geq 0, \quad \|q_{\nu,h}(t) - \mathbb{P}_h(q_h^0)\| \leq C_1 M.$$

□

**Remark 5.** Since it is difficult to prove dissipation of energy for the classical Apparent Topography scheme (see Remark 2), we do not have enough evidence to conclude that this well-balanced scheme is accurate at low Froude number.

## 2.7. Analysis and damping of spurious modes

One of the drawbacks of the triangular B-grid scheme is that the standard scheme without damping (i.e.  $\nu_r = \gamma_r = \nu_u = 0$  in (9)) supports spurious numerical modes: the analysis of the numerical dispersion relation on uniform equilateral meshes performed in [13] shows that (non trivially constant) divergence free velocity fields (with vanishing pressure) can oscillate with pulsation  $\omega$ . This means that there are non trivial velocity fields verifying

$$(\nabla_h^D \cdot \mathbf{u}_h)_k = 0, \forall k, \tag{22}$$

$$\frac{d}{dt} \mathbf{u}_i(t) = -\omega \mathbf{u}_i^\perp, \forall i. \tag{23}$$

We analyse in what follows these fields on general triangular grids: taking the  $\nabla_h^D \cdot$  operator on both sides of (23) and using (22), we have that

$$\frac{d}{dt}(\nabla_h^D \cdot \mathbf{u}_h)_k(t) = -\omega(\nabla_h^D \cdot \mathbf{u}_h^\perp)_k = 0, \forall k.$$

With relationship (7), we conclude that for  $\omega \neq 0$ , it also holds that

$$(\nabla_h^D \times \mathbf{u}_h)_k = 0, \forall k. \quad (24)$$

If  $\mathbf{u}_h$  were a continuous velocity field in a periodic rectangular domain with no holes as is considered here, a vanishing curl and a vanishing divergence would be sufficient to infer that  $\mathbf{u}_h$  is a constant vector field. However, at the discrete level, (22) and (24) do not imply that  $\mathbf{u}_h$  is a constant vector field: indeed, with  $N$  the number of triangles and  $N_r$  the number of vertices, a cell-centered velocity field has  $2N$  degrees of freedom, while (22) and (24) sum up to only  $2N_r = N$  constraints (this equality holds on a periodic triangular mesh of a periodic rectangular domain with no holes). Therefore there are many such spurious modes and a general construction is given in the associated numerical test in Section 4.3. We shall build on (14) to show that these modes are damped to zero by the family of schemes we propose as soon as  $\nu_r = \gamma_r > 0$ . More precisely, with an initial condition  $r_h(0) = 0$  and  $\mathbf{u}_h(0)$  such that (24) is verified, then by definition (16) and property (17), it holds that  $(r_h(0), \mathbf{u}_h(0)) \in \mathcal{E}_{\omega \neq 0}^{\Delta, \perp}$ . By virtue of Lemma 4 (ii), we infer that  $(r_h(t), \mathbf{u}_h(t)) \in \mathcal{E}_{\omega \neq 0}^{\Delta, \perp}$  for all  $t \geq 0$ . This implies by (16)–(17) that

$$a_\star \nabla_h^D \times \mathbf{u}_h(t) = \omega r_h(t), \forall t \geq 0. \quad (25)$$

Now, we use a general decomposition of discrete velocity fields on periodic triangular meshes with no holes (see [16, Lemma 3.5] with  $\alpha \equiv 1$ ): there exists two time-dependent discrete fields  $\psi_h(t)$  and  $\phi_h(t)$ , with  $\psi_h$  defined by its values at the vertices of the mesh and  $\phi_h$  defined by its values at the edges of the mesh, and a uniform vector field  $\bar{\mathbf{u}}(t)$  such that

$$\mathbf{u}_i(t) = (\nabla_h^T \times \psi_h(t))_i + (\nabla^{\text{CR}} \phi_h(t))_i + \bar{\mathbf{u}}(t) \forall i. \quad (26)$$

Moreover, in this decomposition, each of the three components is orthogonal to the other two with respect to the scalar product  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$ . Now, computing the second term in the right-hand side of (14), we obtain, thanks to (26) and the orthogonality of  $\nabla_h^T$  with  $\nabla^{\text{CR}} \times$  and uniform fields

$$\left\| \sqrt{\nu_r} \left( \nabla_h^T r_h + \frac{\omega}{a_\star} \mathbf{u}_h^\perp \right) \right\|_{\mathcal{P}}^2 \geq \nu_{\min} \left( \left\| \nabla_h^T (r_h + \frac{\omega}{a_\star} \psi_h) \right\|_{\mathcal{P}}^2 + \left( \frac{\omega}{a_\star} \right)^2 \left( \|\nabla^{\text{CR}} \times \phi_h\|_{\mathcal{P}}^2 + \|\bar{\mathbf{u}}^\perp\|_{\mathcal{P}}^2 \right) \right), \quad (27)$$

where  $\nu_{\min}$  is the minimum value of  $\nu_r$  over the mesh. Developing the first term in the right-hand side of (27) we obtain

$$\left\| \nabla_h^T (r_h + \frac{\omega}{a_\star} \psi_h) \right\|_{\mathcal{P}}^2 = \|\nabla_h^T r_h\|_{\mathcal{P}}^2 + 2 \frac{\omega}{a_\star} \langle \nabla_h^T r_h, \nabla_h^T \psi_h \rangle_{\mathcal{P}} + \left( \frac{\omega}{a_\star} \right)^2 \|\nabla_h^T \psi_h\|_{\mathcal{P}}^2. \quad (28)$$

In order to evaluate the second term in the right-hand side of (28), we first use (25), then a discrete integration by part for the curl operators (that can be proved just like (11) was proved), then (26) and the fact that this decomposition is orthogonal, to obtain

$$\frac{\omega}{a_\star} \|r_h\|_{\mathcal{D}}^2 = \langle \nabla_h^D \times \mathbf{u}_h, r_h \rangle_{\mathcal{D}} = \langle \mathbf{u}_h, \nabla_h^T \times r_h \rangle_{\mathcal{P}} = \langle \nabla_h^T \times \psi_h, \nabla_h^T \times r_h \rangle_{\mathcal{P}} = \langle \nabla_h^T \psi_h, \nabla_h^T r_h \rangle_{\mathcal{P}}. \quad (29)$$

Plugging (29) into (28) and then into (27), and using (14) we obtain the following inequalities

$$\begin{aligned} \frac{d}{dt} E_h(t) + 2\nu_{\min} \left[ \|\nabla_h^T r_h\|_{\mathcal{P}}^2 + 2 \left( \frac{\omega}{a_\star} \right)^2 \|r_h\|_{\mathcal{D}}^2 + \left( \frac{\omega}{a_\star} \right)^2 \left( \|\nabla_h^T \psi_h\|_{\mathcal{P}}^2 + \|\nabla^{\text{CR}} \times \phi_h\|_{\mathcal{P}}^2 + \|\bar{\mathbf{u}}^\perp\|_{\mathcal{P}}^2 \right) \right] &\leq 0, \\ \frac{d}{dt} E_h(t) + 2\nu_{\min} \left( \frac{\omega}{a_\star} \right)^2 \left( \|r_h\|_{\mathcal{D}}^2 + \|\mathbf{u}_h^\perp\|_{\mathcal{P}}^2 \right) &\leq 0 \\ \frac{d}{dt} E_h(t) + 2\nu_{\min} \left( \frac{\omega}{a_\star} \right)^2 E_h(t) &\leq 0, \end{aligned}$$

which implies

$$E_h(t) \leq \exp \left[ -2\nu_{\min} \left( \frac{\omega}{a_\star} \right)^2 t \right] E_h(0),$$

and thus the exponential damping of the spurious modes with a damping rate (at least) proportional to  $\nu_{\min}$  and thus to  $h_{\min}$ , where  $h_{\min}$  is the minimum of the  $h_i$  over the mesh.

### 3. Analysis of fully discrete staggered schemes

We present in this section a time discretization for scheme (9). Of course, one may first think about the splitting scheme where we deal with the homogeneous system in the first step and take the source term into account in the second step. However, this naive splitting scheme is unable to capture the steady states. Therefore, in this section, we focus on the analysis of the so-called one step scheme.

#### 3.1. The fully discrete one step scheme

Since a totally explicit discretization of the Coriolis term ( $\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \omega \Delta t \mathbf{u}_i^{n,\perp}$ ) would be unstable [1], we need to discretize it implicitly enough. In order to achieve this, let us introduce two new parameters  $\theta_1$  and  $\theta_2$  corresponding to the time discretization of the Coriolis source term. Let us also introduce  $\tau_1$  and  $\tau_2$  corresponding to the time discretization of the divergence term in the pressure equation. For convenience, let us denote

$$\mathbf{u}^\theta = \begin{pmatrix} \theta_1 u^n + (1 - \theta_1) u^{n+1} \\ \theta_2 v^n + (1 - \theta_2) v^{n+1} \end{pmatrix} \text{ and } \mathbf{u}^\tau = \begin{pmatrix} \tau_1 u^n + (1 - \tau_1) u^{n+1} \\ \tau_2 v^n + (1 - \tau_2) v^{n+1} \end{pmatrix}.$$

Then, we introduce the  $\theta$ - $\tau$  scheme which is given by

$$\begin{cases} r_k^{n+1} = r_k^n - a_\star \Delta t (\nabla_h^D \cdot \mathbf{u}_h^\tau)_k + \Delta t \nabla_h^D \cdot \left[ \nu_r \left( \nabla_h^T r_h^n + \frac{\omega}{a_\star} \mathbf{u}_h^{n,\perp} \right) \right]_k \\ \mathbf{u}_i^{n+1} = \mathbf{u}_i^n - a_\star \Delta t (\nabla_h^T r_h^n)_i + \nu_u \Delta t (\nabla^{\text{CR}} [\mathbf{u}_h^n \cdot \mathbf{n}])_i - \omega \Delta t \mathbf{u}_i^{\theta,\perp} + \frac{\omega}{a_\star} \Delta t (\gamma_r)_i \left( \nabla_h^T r_h^n + \frac{\omega}{a_\star} \mathbf{u}_h^{n,\perp} \right)_i^\perp \end{cases} \quad (30)$$

**Remark 6.** The  $\theta$ - $\tau$  scheme (30) is not totally explicit but we can still calculate the velocity and pressure updates without having to solve any linear system, although the velocity field  $\mathbf{u}^{n+1}$  appears in the pressure equation: we first compute the updated velocity field from the momentum equation, and then use it to compute  $\mathbf{u}^\tau$  in the pressure equation.

#### 3.2. Well-balanced and orthogonality preserving scheme

**Proposition 6.** It holds that

- (i) The fully discrete one step scheme (30) is a well-balanced scheme.
- (ii) When  $\nu_r = \gamma_r$ , the fully discrete one step scheme (30) is an orthogonality preserving scheme if

$$\tau_1 = \theta_1 \quad \text{and} \quad \tau_2 = \theta_2,$$

which means that the velocity field used in the Coriolis source term in the velocity equation and that used to compute the divergence term in the pressure equation should be the same.

*Proof.* Let us consider at time  $t^n$  a numerical solution  $\hat{q}_h^n = (\hat{r}_h^n, \hat{\mathbf{u}}_h^n)$  which is in the discrete kernel, which reads (19). We also recall that this implies that  $\hat{\mathbf{u}}_h^n$  also verifies (20) and (21). We shall show that the numerical solution at time  $t^{n+1}$  is still equal to  $\hat{q}_h^n$ . First, developing  $\mathbf{u}^\theta$  and taking (19) into account to express the pressure gradient, as well as the fact that the jump term vanishes due to (21), the velocity equation of (30) leads to

$$\begin{pmatrix} 1 & -\omega \Delta t (1 - \theta_2) \\ \omega \Delta t (1 - \theta_1) & 1 \end{pmatrix} \begin{pmatrix} u_i^{n+1} \\ v_i^{n+1} \end{pmatrix} = \begin{pmatrix} 1 & -\omega \Delta t (1 - \theta_2) \\ \omega \Delta t (1 - \theta_1) & 1 \end{pmatrix} \begin{pmatrix} u_i^n \\ v_i^n \end{pmatrix},$$

which implies that  $u_i^{n+1} = \hat{u}_i^n$  and  $v_i^{n+1} = \hat{v}_i^n$ . Next, this implies that  $\mathbf{u}^\tau = \hat{\mathbf{u}}^n$  and  $\nabla_h^D \cdot \mathbf{u}^\tau = 0$  by (20). Therefore, the pressure equation of the fully discrete one step scheme (30) reduces to  $r_k^{n+1} = \hat{r}_k^n$ . This proves (i).

Next, let us consider a numerical solution  $q_h^n$  belonging to  $\mathcal{E}_{\omega \neq 0}^{\Delta, \perp}$ , so that  $\langle q_h^n, \hat{q}_h \rangle = 0$  for any steady state  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$ . Taking the scalar product of (30) with any  $\hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta$ , and following the proof of Proposition 4, the only terms that remain are

$$\langle q_h^{n+1}, \hat{q}_h \rangle = a_* \Delta t \langle \mathbf{u}_h^\tau, \nabla_h^T \hat{r}_h \rangle_{\mathcal{P}} + \omega \Delta t \langle \mathbf{u}_h^\theta, \hat{\mathbf{u}}_h^\perp \rangle_{\mathcal{P}}.$$

We realize that when  $\tau_1 = \theta_1$  and  $\tau_2 = \theta_2$ , we get  $\mathbf{u}_h^\tau = \mathbf{u}_h^\theta$ , from which it follows that

$$\langle q_h^{n+1}, \hat{q}_h \rangle = 0, \forall \hat{q}_h \in \mathcal{E}_{\omega \neq 0}^\Delta.$$

Therefore, we conclude that  $q_h^{n+1} \in \mathcal{E}_{\omega \neq 0}^{\Delta, \perp}$  and so the scheme is orthogonality preserving.  $\square$

#### 4. Numerical test cases

In this section, we perform several test cases for all proposed schemes: LF, AT and MAT. Moreover, for the sake of comparison, we also run the scheme with the Laplacian diffusion in the pressure equation and with the normal velocity jump diffusion for the velocity equations (PL-VJ scheme): in (9), the diffusion in the pressure equation is replaced by  $-\nabla_h^D \cdot (\nu_r \nabla_h^T r_h)$  and  $\gamma_r$  is set to 0. The classical scheme with the full Laplacian diffusion for the velocity equations and no diffusion in the pressure equation (VL scheme) is also tested in this section: in (9), we set  $\nu_r = \gamma_r = 0$  and then replace the diffusion term (10) in the velocity equation by  $\frac{1}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| (\mathbf{u}_j - \mathbf{u}_i)$ . The crucial difference with (10) is that (10) only involves the normal jump of the velocity field, and this normal jump vanishes on discretely divergence free velocities, as recalled in Remark 1 and Proposition 4.

##### 4.1. Well-balanced test case

In this test case, we investigate the behavior of the staggered schemes with a GE as an initial condition. Particularly, we consider the stationary vortex in the square periodic domain  $\mathbb{T}_2 = [-0.5, 0.5] \times [-0.5, 0.5]$  with initial pressure  $r^0$  given by

$$r(x, y, t = 0) = 1 - \exp \left[ - \left( \frac{3x}{0.5} \right)^2 - \left( \frac{3y}{0.5} \right)^2 \right],$$

and we construct the discrete initial pressure by interpolating this pressure field at the mesh vertices. Then, we construct the initial velocity field  $\mathbf{u}^0$  by using the definition of the discrete kernel (15) so that we can obtain a discrete stationary state.

Figure 2 indicates that the classical VL staggered scheme is not well-balanced since it quickly produces some spurious waves in the orthogonal subspace (Fig. 2b); it also strongly damps the kernel part (Fig. 2a). Similarly, The PL-VJ scheme also destroys the geostrophic equilibrium. However, compared to the VL scheme, it produces waves with a smaller amplitude in the orthogonal part and the damping rate of the kernel part is much slower.

On the contrary, the LF, AT and MAT staggered schemes are well-balanced as expected. This can be checked because the orthogonal parts of their solutions remain equal to zero (Figure 2b) during the computation and their kernel parts remain constant (Figure 2a). Additionally, Figure 3 shows that the pressure contours of the VL and, to a much smaller extent, PL-VJ schemes are different from those of the other schemes. This is another evidence that they are unable to capture steady-states, contrarily to the LF, AT and MAT schemes.

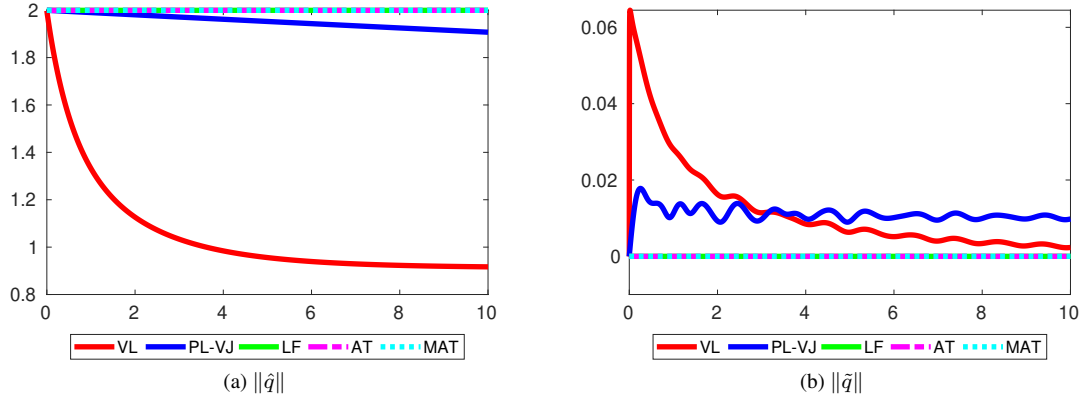


Figure 2: Vortex test case: evolution of the kernel and orthogonal parts.

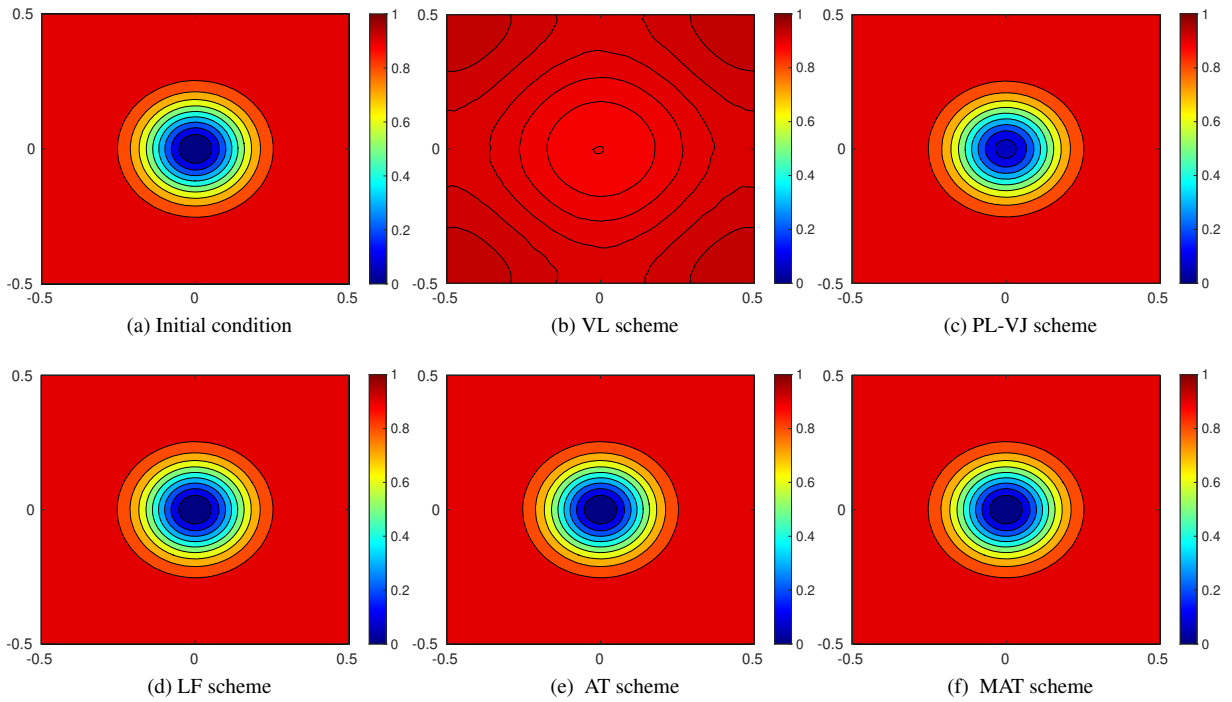


Figure 3: Pressure contours  $r(x, y, t)$  at time  $t = 10$  obtained from staggered type schemes and initial condition (top left).

#### 4.2. Orthogonality preserving test case

In this test case, we consider periodic boundary conditions and an initial vector field given by

$$u(x, y, t = 0) = \frac{1}{2} \exp \left[ - \left( \frac{4x}{0.4} \right)^2 - \left( \frac{4y}{0.8} \right)^2 \right] \quad \text{and} \quad v(x, y, t = 0) = \frac{1}{2} \exp \left[ - \left( \frac{4x}{0.8} \right)^2 - \left( \frac{4y}{0.4} \right)^2 \right]$$

in the domain  $\mathbb{T}_2 = [-0.5, 0.5] \times [-0.5, 0.5]$  and we construct the discrete initial velocity by interpolating this velocity field at the cell centers. Then the initial pressure  $r(x, y, t = 0)$  is constructed at the mesh vertices by using the definition of the discrete orthogonal subspace (16)–(17). In all cases, we choose  $\theta_1 = \theta_2 = \frac{1}{2}$  for the time discretization of the Coriolis force in (30) and the value of  $(\tau_1, \tau_2)$  may vary from test to test.

Figure 4a shows that the classical VL, PL-VJ and AT schemes are not orthogonality preserving since the kernel components of these schemes are not equal to zero. Moreover, since the kernel part is updated at each time step, it is not a constant in time. Figure 4b shows that the damping rates of the orthogonal parts of the VL, PL-VJ, AT and MAT schemes are larger than that of the Low Froude scheme. Although the PL-VJ and the AT schemes are not orthogonality preserving, these strategies on triangular grid create waves in the kernel that have much smaller amplitudes than those on Cartesian grids (see [6]). Next, Figure 4a was obtained with  $\tau_1 = \tau_2 = 1$  for the fully discrete LF and MAT schemes, and this generates waves with very small amplitudes in the kernel part; this underlines that in order to obtain schemes that are really orthogonality preserving, we need to use a velocity field in the Coriolis source term that is equal to that used for the computation of the divergence term in the pressure equation. Indeed, Figures 5a and 5b clearly show that the LF- $\theta$ - $\tau$  and the MAT- $\theta$ - $\tau$  schemes with  $\theta_1 = \theta_2 = \tau_1 = \tau_2 = \frac{1}{2}$  are perfectly orthogonality preserving since these strategies do not create waves in the kernel subspace during the computational process, as proved by Lemma 6.

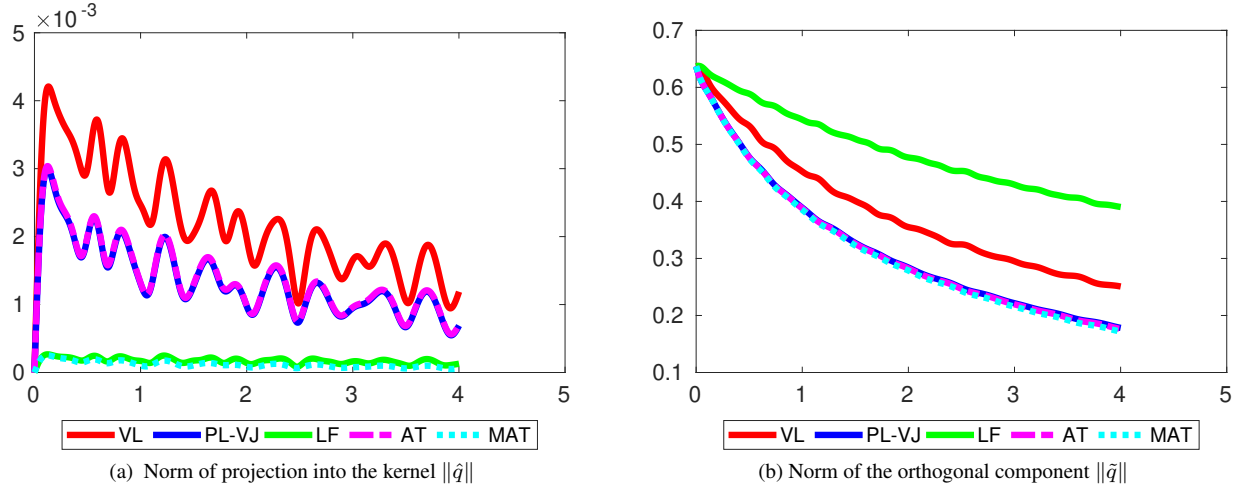


Figure 4: Orthogonality preserving test case: evolution of the kernel and orthogonal parts of different staggered type schemes with  $\theta_1 = \theta_2 = \frac{1}{2}$  and  $\tau_1 = \tau_2 = 1$  for the time discretization of the Coriolis force.

#### 4.3. Spurious mode test case

In this test case, we consider an initial condition given by a spurious mode of the triangular B-grid when no damping is used (i.e. (9) with  $\nu_r = \gamma_r = \nu_u = 0$ ), and we shall observe that this mode is now correctly damped to 0. As described in Section 2.7, this initial condition satisfies

$$r = 0 \quad \text{and} \quad \nabla_h^D \cdot \mathbf{u}_h = \nabla_h^D \times \mathbf{u}_h = 0.$$

In order to construct such a spurious mode, we recall the discrete Hodge decomposition on periodic triangular grids that is given by (26) (taken at time  $t = 0$ ):

$$\mathbf{u}_h = \nabla_h^T \times \psi_h + \nabla_h^{\text{CR}} \phi_h + \bar{\mathbf{u}}.$$

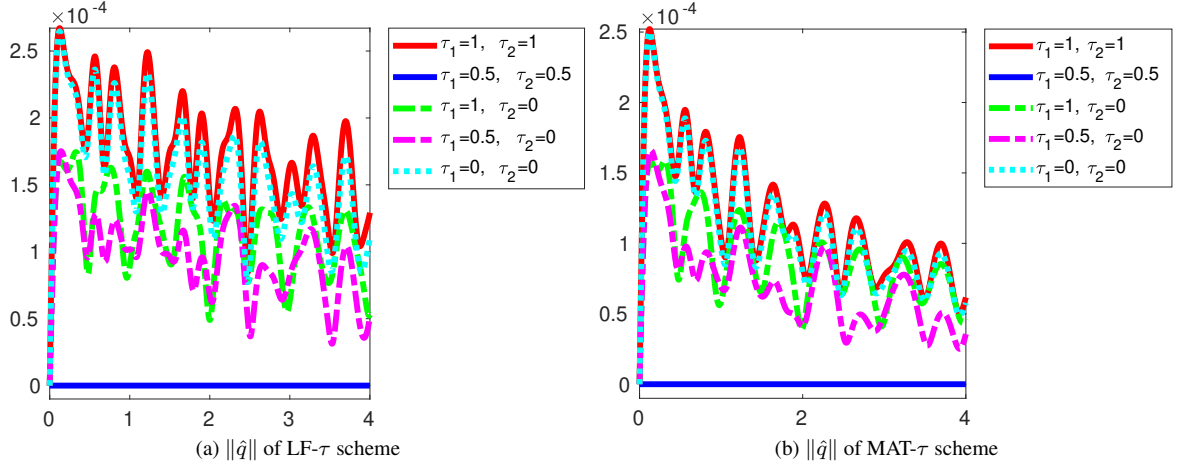


Figure 5: Orthogonality preserving test case: evolution of the kernel part of the LF- $\theta$ - $\tau$  and MAT- $\theta$ - $\tau$  scheme with  $\theta_1 = \theta_2 = \frac{1}{2}$  and various values of  $(\tau_1, \tau_2)$ . Orthogonality preservation is only achieved when  $\tau = \theta$ .

But the role of the curl and gradient operators can be exchanged so that there also exist  $(\mu_h, \lambda_h)$  defined respectively by their values at the vertices and at the edges of the mesh such that

$$\mathbf{u}_h = \nabla_h^T \mu_h + \nabla^{\text{CR}} \times \lambda_h + \bar{\mathbf{u}}.$$

Now  $\nabla_h^D \cdot \mathbf{u}_h = 0$  implies that (up to the constant  $\bar{\mathbf{u}}$ )  $\mathbf{u}_h = \nabla^{\text{CR}} \times \lambda_h$  and  $\nabla_h^D \times \mathbf{u}_h = 0$  implies that  $\mathbf{u}_h = \nabla^{\text{CR}} \phi_h$ . In order to have a velocity field that is at the same time a non-conforming gradient and a non-conforming curl, a possibility is that one picks any random (periodic)  $\alpha_h$  defined at the edges of the mesh and first set  $\mathbf{v}_h = \nabla^{\text{CR}} \alpha_h$ . Then, one computes the second Hodge decomposition of  $\mathbf{v}_h$ :

$$\mathbf{v}_h = \nabla_h^T \zeta_h + \nabla^{\text{CR}} \times \lambda_h.$$

Then, if we set  $\mathbf{u}_h = \mathbf{v}_h - \nabla_h^T \zeta_h$ , we obtain that  $\mathbf{u}_h$  is both a non-conforming gradient and a non-conforming curl:

$$\mathbf{u}_h = \nabla^{\text{CR}} \times \lambda_h \quad \text{and} \quad \mathbf{u}_h = \nabla_h^{\text{CR}} \alpha_h - \nabla_h^T \zeta_h = \nabla^{\text{CR}} (\alpha_h - \tilde{\zeta}_h),$$

where  $\tilde{\zeta}_h$  is defined at the edge midpoints as the half-sum of the values of  $\zeta_h$  at the vertices of the edges.

Note that, in order to decompose  $\mathbf{v}_h$ , we have to solve a  $P_1$  Lagrange finite element solution of the Laplace equation:

$$(\nabla_h^T \zeta_h, \nabla_h^T \theta_h) = (\mathbf{v}_h, \nabla_h^T \theta_h), \quad \forall \theta_h.$$

Having constructed a spurious mode, we use it as an initial value for the various schemes developed in this article, as well as the "Inertial Oscillation" (IO scheme) which is simply the  $\theta$ -scheme (with  $\theta = \frac{1}{2}$ ) applied to (23), since this corresponds to the standard B-grid scheme applied to the spurious mode ( $r = 0$  and  $\nabla_h^D \cdot \mathbf{u}_h = 0$ ) when no damping is used. We recall that these spurious modes belong to the orthogonal space (16)–(17) since both their pressure components and the curls of their velocity components vanish. So their initial projections in the kernel space (15) also vanish. According to Figure 6a, we can see that the three non-orthogonality preserving schemes VL, PL-VJ and AT immediately create some waves in the kernel part. Moreover, concerning the long time behavior, the kernel part of the VL scheme is damped by the Laplacian diffusion while it seems to be a constant for the PL-VJ and AT schemes. On the other hand, the LF and MAT schemes do not create any wave in the kernel. Figure 6b shows that the orthogonal part is damped by all the numerical schemes with numerical diffusion. As a final remark, we can stress that Figure 6c illustrates the fact that damping spurious modes supported by the standard B-grid scheme (the IO scheme) can be achieved by schemes that preserve GEs and their orthogonal space (the LF and MAT schemes), which is a major improvement over the usual damping solution (the VL scheme).



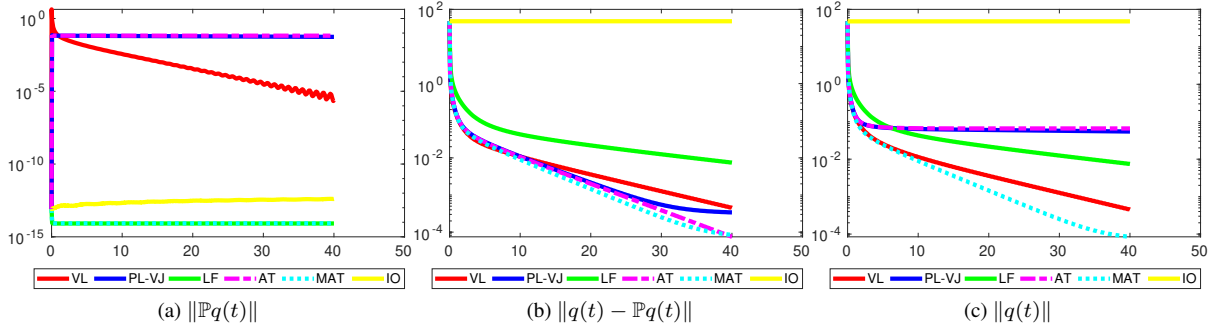


Figure 6: Energy evolution of the kernel part (a), of the orthogonal part (b) and of the full solution (c).

#### 4.4. Accuracy at low Froude number test case

We now consider an initial condition close to the discrete kernel, up to a perturbation of size  $M$ . This initial condition is simply given by

$$q_h^0 = \hat{q}_h^0 + M \frac{\tilde{q}_h^0}{\|\tilde{q}_h^0\|},$$

where  $\hat{q}_h^0$  stands for the kernel part given in Section 4.1 and  $\tilde{q}_h^0$  is the orthogonal part considered in Section 4.2. In that case, the exact solution remains close to  $\hat{q}_h^0$ , in the sense that their difference remains of size  $M$  for all time (see Lemma 5). So we expect efficient schemes to follow this behaviour. However, Figures 7a and 7b indicate that the classical VL and PL-VJ schemes are not accurate at low Froude number because the norm of  $\|q(t) - \mathbb{P}q^0\|$  has a size which is much larger than  $M$  and independent of the value of  $M$ . These figures also illustrate that the solutions of the VL and PL-VJ schemes quickly move far from the kernel without regard to the size of parameter  $M$ . By contrast, the proposed LF, AT and MAT schemes are accurate at low Froude number because the norm of the total deviation remains of order  $\mathcal{O}(M)$ .

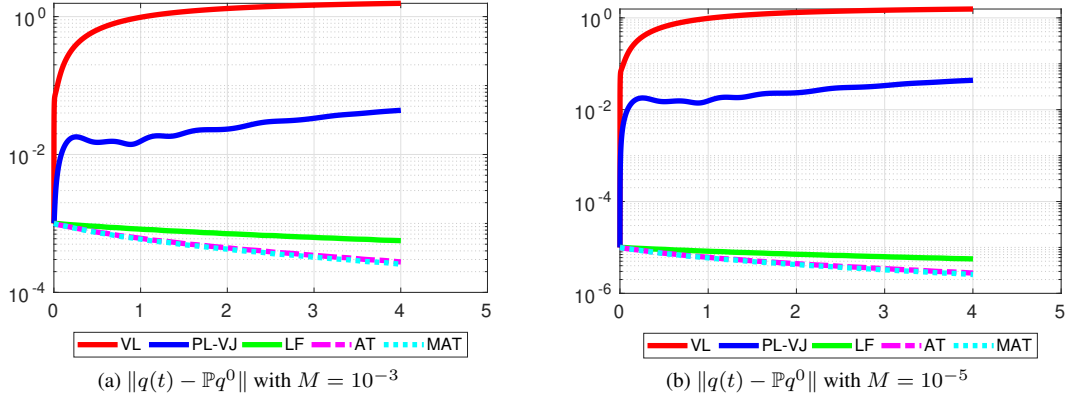


Figure 7: Evolution of the total deviation from the initial condition, when the initial condition is close to the discrete kernel.

#### 4.5. Circular dam-break test case

In this test case, we consider the initial condition given by

$$\begin{cases} r(x, y, t = 0) = \begin{cases} 2, & \text{if } x^2 + y^2 \leq 1 \\ 1, & \text{if } x^2 + y^2 > 1. \end{cases} \\ u(x, y, t = 0) = 0, \\ v(x, y, t = 0) = 0. \end{cases}$$

with periodic boundary condition,  $a_* = \omega = 1$  and the domain is the square  $[-5, 5] \times [-5, 5]$ . In this test case, we aim to investigate the long time behavior of the numerical solution when the initial condition is far from the GE. Figure 8a presents the projection into the kernel of the initial condition. This is the target solution for the simulation in long time since the numerical diffusion will damp the orthogonal part.

Figure 8 clearly shows that in long time, the solution of the VL and PL-VJ schemes tend to a trivial steady state which consists of only a constant pressure field. On the contrary, the LF, AT and MAT schemes tend to the GE, since the final state is similar to projection  $\mathbb{P}q^0$  of the initial condition shown in Figure 8a.

On the other hand, Figure 9 shows that the kernel components of the classical VL or PL-VJ schemes are damped during the simulation while it is a constant for the LF and MAT schemes, and nearly a constant for the classical AT scheme. Moreover, the total deviation (distance between the numerical solution and the initial projection) of VL or PL-VJ scheme decreases in early times, but eventually increases in long time, while it tends to zero with the other schemes. This is another evidence to conclude that the VL and PL-VJ schemes will tend to a wrong component in the kernel, while the other schemes have solutions that, as expected, tend to the correct GE. The fact that the classical AT scheme does not capture a projection in the kernel which is constant in time is due to the fact that it is not orthogonality preserving, and so there is an exchange of energy between the orthogonal space and the kernel.

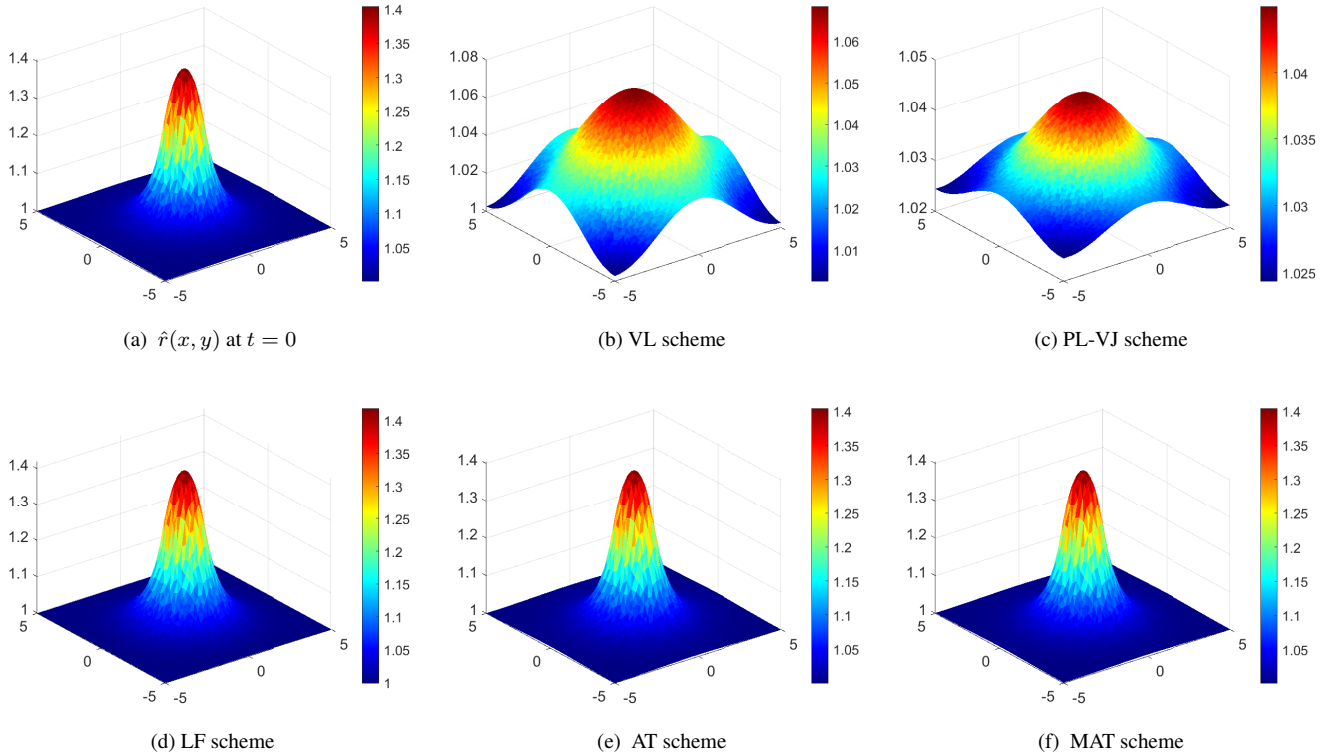


Figure 8: Pressure  $r(x, y, t)$  at time  $t = 150$  obtained from staggered type schemes and the initial projection (top left).

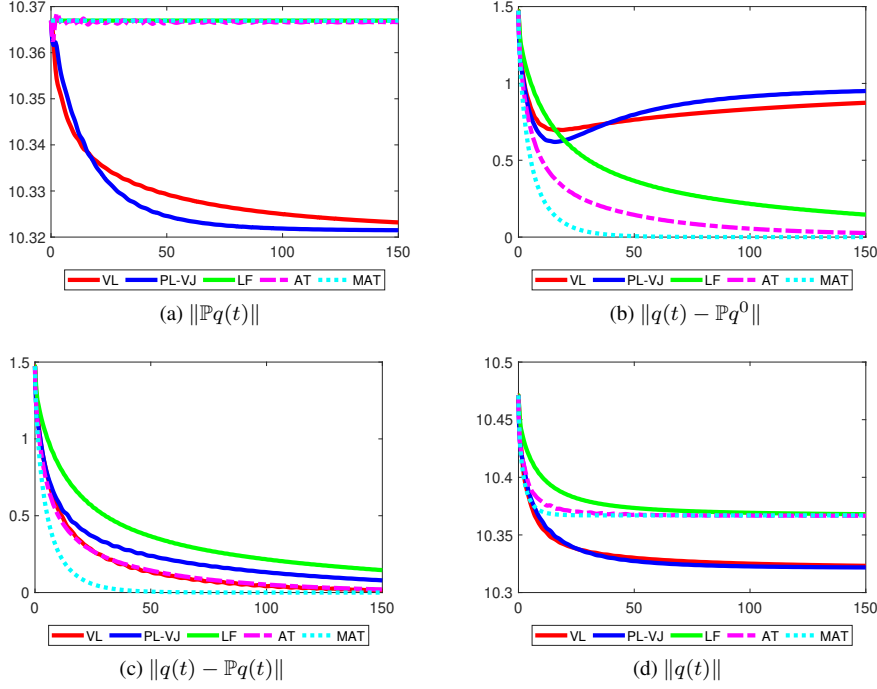


Figure 9: The evolution of the kernel part (a), deviation from the initial projection (b), orthogonal component (c) and norm of the full solution (d).

#### 4.6. Anticyclonic eddy propagation test case for Rossby waves

As a first extension of what was presented in the simplified case of the constant coefficient linearized system (2), we consider in this test case the Rossby waves generated by the variation of the Coriolis force with latitude. In particular, we focus on the  $\beta$ -plane approximation for the Coriolis source term given by  $\omega = \omega_0 + \beta y$ . We aim to investigate how the above schemes influence the numerical propagation of the Rossby waves.

This test case is motivated by the second test in [17]. Here, all parameters are scaled according to our setting. Particularly, the computational domain is a rectangular  $[-1.0, 1.0] \times [-0.6, 0.6]$  with periodic boundary conditions. The initial pressure is a Gaussian distribution at the center of the domain and the initial velocity field is chosen so as to satisfy the geostrophic balance  $a_* \nabla r = -(\omega_0 + \beta y) \mathbf{u}^\perp$ . Therefore, the initial condition is given by

$$\begin{cases} r(x, y, t = 0) = A \exp\left(-\frac{x^2 + y^2}{B^2}\right), \\ u(x, y, t = 0) = 2A \frac{a_*}{\omega_0 + \beta y} \frac{y}{B^2} \exp\left(-\frac{x^2 + y^2}{B^2}\right), \\ v(x, y, t = 0) = -2A \frac{a_*}{\omega_0 + \beta y} \frac{x}{B^2} \exp\left(-\frac{x^2 + y^2}{B^2}\right). \end{cases}$$

We choose  $A = 0.6$ ,  $B = 0.13$ ,  $a_* = 1$  as well as the Coriolis parameters  $\omega_0 = 4$ ,  $\beta = 4/3$ . According to Figure 10, the anticyclonic vortex evolve purely westward. Qualitatively, the results with the MAT scheme presented here are in good accordance with those published in [17], and much less diffused than those obtained with the VL and PL-VJ schemes.

### 5. Extension to non-linear equations, non-flat topography and hybrid velocity dissipation

We propose the following extension of the presented schemes for the inclusion of the non-linear and topography terms; moreover, we discuss how to recover full Laplacian velocity damping, with a possible tuning of the respective

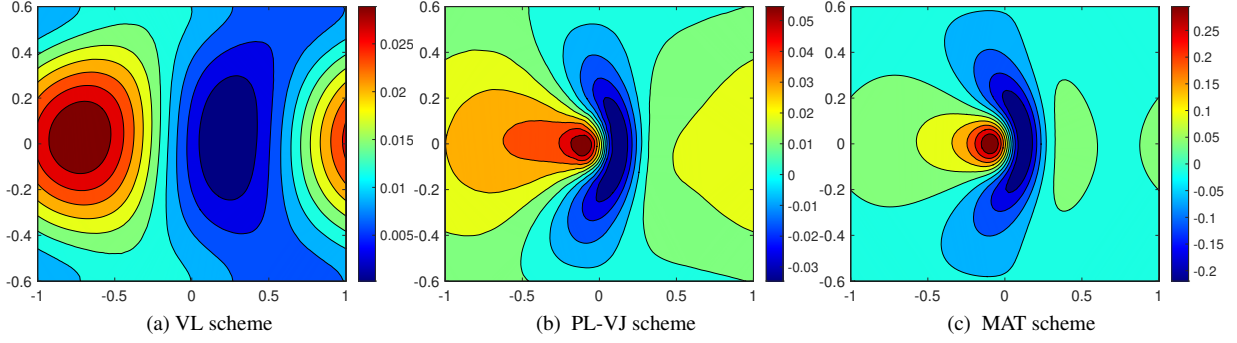


Figure 10: The contours of  $h(x, y, t)$  at time  $t = 20$  obtained from staggered type schemes for the case  $\beta = 4/3$ .

weights of damping the normal and tangential components of the velocity jumps at the cell interfaces. We first define a further dissipation operator acting on the tangential components of the velocity jumps:

$$(\nabla^{\text{CR}} \times \llbracket \mathbf{u}_h \cdot \mathbf{n}^\perp \rrbracket)_i := \frac{1}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| [(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ij}^\perp] \mathbf{n}_{ij}^\perp, \quad (31)$$

where  $\mathbf{n}^\perp$  is a unit vector orthogonal to  $\mathbf{n}$ .

Then, using  $\mathbf{q} = h\mathbf{u}$  as main unknown, the scheme reads: Once  $(h^n, \mathbf{q}^n)$  is known, then compute

$$\begin{cases} \mathbf{q}_i^{n+1} = \mathbf{q}_i^n - \Delta t \left[ \tilde{\nabla}_h^T \cdot \left( \frac{1}{h_c^n} \mathbf{q}^n \otimes \mathbf{q}^n \right) \right]_i - \Delta t g(h_c)_i (\nabla_h^T h^n)_i - \omega \Delta t \mathbf{q}_i^{\theta, \perp} - \Delta t g(h_c)_i (\nabla_h^T b)_i \\ + \nu_u \Delta t \left( \nabla^{\text{CR}} \llbracket \frac{\mathbf{q}^n}{h_c} \cdot \mathbf{n} \rrbracket \right)_i + \nu_\perp \Delta t \left( \nabla^{\text{CR}} \times \llbracket \frac{\mathbf{q}^n}{h_c} \cdot \mathbf{n}^\perp \rrbracket \right)_i + \omega \Delta t (\gamma_r)_i (gh_c^n \nabla_h^T h^n + \omega \mathbf{q}^{n, \perp} + gh_c^n \nabla_h^T b)_i^\perp, \\ h_k^{n+1} = h_k^n - \Delta t (\nabla_h^D \cdot \mathbf{q}^\tau)_k + \Delta t \nabla_h^D \cdot [\nu_r (gh_c^n \nabla_h^T h^n + \omega \mathbf{q}^{n, \perp} + gh_c^n \nabla_h^T b)]_k, \end{cases} \quad (32)$$

in which, for each cell  $T_i$ , the interpolated value  $(h_c)_i$  is equal to one third of the sum of the values of  $h$  at the vertices of  $T_i$ . There are several options for the definition of the  $\tilde{\nabla}_h^T \cdot$  operator which we do not detail here. In practice, we chose to define this divergence on the primal cells as the mean value of the three divergences  $\nabla_h^D \cdot$  that can be computed at the vertices of each cell. Note that now, the diffusion terms vanish on fields that verify a discrete GE taking into account the topography and the jump terms still vanish for discretely divergence free velocities when the additional tuning parameter  $\nu_\perp$  vanishes. On the contrary, we recover the full Laplacian velocity damping when  $\nu_\perp = \nu_u$ . Indeed in that case

$$[(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ij}^\perp] \mathbf{n}_{ij}^\perp + [(\mathbf{u}_j - \mathbf{u}_i) \cdot \mathbf{n}_{ij}] \mathbf{n}_{ij} = [(\mathbf{u}_j - \mathbf{u}_i)].$$

Scheme (32) is the fully discrete version of a semi-discrete scheme (continuous in time, discrete in space) devised so that, apart from the advection terms, it dissipates energy. Indeed, when we take the scalar product on primal cells of the first equation of (32) with  $\frac{\mathbf{q}}{h_c}$  and then sum with the scalar product of the second equation with  $g(h+b)$ , we obtain (omitting the time derivatives and advection terms)

$$\begin{aligned} & - \Delta t g \langle \nabla_h^T (h+b), \mathbf{q} \rangle_{\mathcal{P}} - \Delta t g \langle \nabla_h^D \cdot \mathbf{q}, (h+b) \rangle_{\mathcal{D}} \\ & - \nu_u \Delta t \sum_{A_{ij}} |A_{ij}| \left| \left[ \left( \frac{\mathbf{q}}{h_c} \right)_i - \left( \frac{\mathbf{q}}{h_c} \right)_j \right] \cdot \mathbf{n}_{ij} \right|^2 - \nu_\perp \Delta t \sum_{A_{ij}} |A_{ij}| \left| \left[ \left( \frac{\mathbf{q}}{h_c} \right)_i - \left( \frac{\mathbf{q}}{h_c} \right)_j \right] \cdot \mathbf{n}_{ij}^\perp \right|^2 \\ & - \omega \Delta t \left\langle \gamma_r \left( g \nabla_h^T (h+b) + \omega \frac{\mathbf{q}^\perp}{h_c} \right), \mathbf{q}^\perp \right\rangle_{\mathcal{P}} + \Delta t \left\langle \nabla_h^D \cdot \left[ \nu_r h_c \left( g \nabla_h^T (h+b) + \omega \frac{\mathbf{q}^\perp}{h_c} \right) \right], g(h+b) \right\rangle_{\mathcal{D}}. \end{aligned}$$

The sum of the first two terms vanishes like in the linear case. The third and fourth terms are always negative. In addition, when  $\nu_r = \gamma_r$  and thanks to a discrete integration by part, the last two terms sum up to

$$-\Delta t \gamma_r \left\langle g \nabla_h^T (h + b) + \omega \frac{\mathbf{q}^\perp}{h_c}, h_c \left( g \nabla_h^T (h + b) + \omega \frac{\mathbf{q}^\perp}{h_c} \right) \right\rangle_{\mathcal{P}},$$

which is negative as long as all values of  $(h_c)$  remain positive. This energy dissipation ensures in practice the stability of the scheme.

As a test case, we consider the stationary vortex of the nonlinear rotating shallow water equation studied in [18, 19] on the domain  $[-0.5, 0.5] \times [-0.5, 0.5]$  with periodic boundary condition. The initial velocity field is given by

$$\mathbf{u}^0(r, \theta) = \nu_\theta(r) \bar{\mathbf{e}}_\theta, \quad \nu_\theta(r) = \varepsilon \left[ 5r \chi \left( r < \frac{1}{5} \right) + (2 - 5r) \chi \left( \frac{1}{5} \leq r < \frac{2}{5} \right) \right], \quad (33)$$

where  $r$  stands for the distance to the center of the domain and  $\chi$  is the characteristic function of a domain. The corresponding initial water depth is the solution of the following ODE

$$\partial_r (h^0 + b) = \frac{1}{g} \left( \omega \nu_\theta + \frac{\nu_\theta^2}{r} \right). \quad (34)$$

Let us note that when the coefficient  $\varepsilon$  is small, the nonlinear term in the definition of the water depth (34) is much smaller than the linear term and the stationary solution is very close to a geostrophic equilibrium. Note that we do not claim that the scheme presented here preserves this initial condition; however, we show that for small values of  $\varepsilon$ , it behaves much better than other schemes which do not preserve geostrophic equilibria. In this test, in addition to the non-linear terms, we also consider the topography  $b(x)$  given by

$$b(x, y) = \frac{1}{2} e^{-40(x^2 + y^2)} \quad (35)$$

in order to show that the presented schemes also behaves correctly in presence of a non flat topography. Figure 11 (for  $\varepsilon = 0.01$ ) shows the contours of  $(h + b)$  after some time of simulation and it is shown that the solution of the MAT scheme is closer to the initial condition than the PL-VJ scheme, which produces more diffused solutions. We used  $\nu_\perp = 0$  in this test.

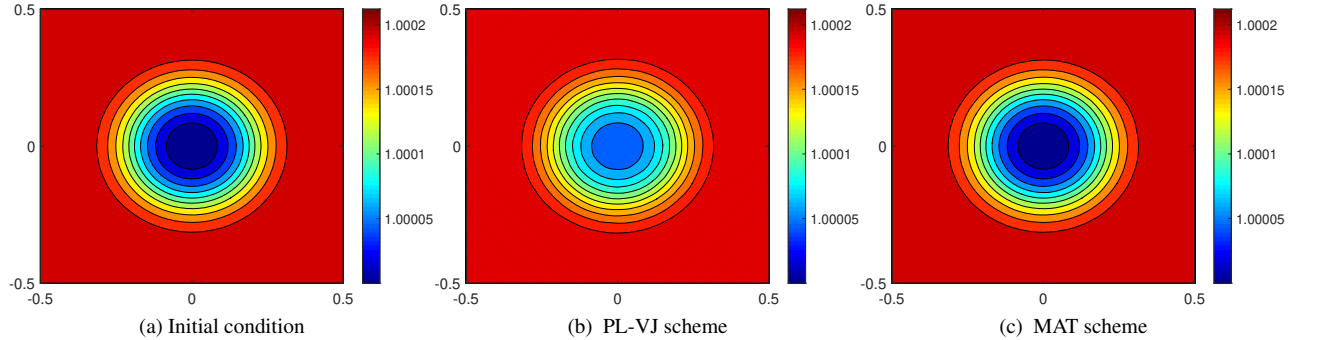


Figure 11: The contours of  $h(x, y, t) + b(x, y)$  at time  $t = 20$  obtained from staggered type schemes for the case  $\varepsilon = 0.01$ .

## 6. Conclusion

In this work, we propose new staggered type schemes on triangular cells that are accurate for the simulation of flows near the GE. The construction of these schemes is based on the adaptation of the Low Froude and Apparent Topography strategies to triangular grids. Theoretical analysis and numerical results in the linearized setting with

constant Coriolis parameter show that the new schemes are well-balanced. Additionally, unlike the classical VL, PL-VJ and Apparent Topography method, the Low Froude and the Modified Apparent Topography schemes are orthogonality preserving under an appropriate discretization in time. This seems particularly important to really separate the GE from the orthogonal perturbations around it. An important feature of the proposed schemes is that they are able to damp the spurious modes associated to the staggered triangular unknown arrangement without destroying the GE. Encouraging preliminary results show that the schemes can be used in and extended to more realistic and complex situations, like non-constant Coriolis parameters and non-linear dynamics; this needs to be further explored in future works. Coupling with more dissipative schemes was also shown possible via a tuning parameter that can be adjusted according to the expected physical regime. Additional effort could also be dedicated to the optimal time step choice for these schemes and to the incorporation of ideas drawn from [20, 21] for the energy control in fully discrete extensions to the nonlinear case. Another possible extension would be to consider higher-order schemes like those in the  $P_k^{\text{disc}}-P_{k+1}$  family discussed in the introduction.

## Appendix A. Analysis of the behavior of collocated schemes

Since collocated schemes have a long and successful history for the numerical approximation of systems of conservation laws such as (1) with  $\omega = 0$ , one could have the natural idea to try to extend them to the treatment of the Coriolis source term ( $\omega \neq 0$ ). The aim of this Appendix is to show the difficulties one encounters in trying to devise collocated schemes that would verify a discrete version of the GEs on triangular meshes.

First, as far as Cartesian meshes are concerned, it is well-known that the classical Godunov scheme applied to (2) will fail to capture a discrete equivalent of the GE (3) [3]. Without going into details, the reason is that the numerical diffusion of that scheme, that is necessary for stability reasons, also destroys GEs and, as a consequence, generates spurious modes at a rate which is proportional to  $\frac{h}{M}$ , with  $h$  being the mesh spacing. In the 1D case, the only reason for the accuracy problem is linked to the numerical diffusion in the pressure equation [1]. To remedy this issue, the authors of [3] introduce the so-called Apparent Topography method by adapting the hydrostatic reconstruction of [22]; this works well in the 1D case and extensions to the 2D Cartesian case were considered in [23, 24, 25, 26]. Other relevant works on this issue can be found in [27, 28, 29, 30, 31].

However, in the 2D Cartesian case, the numerical diffusion in the velocity equations also comes into play since it does not vanish on velocity fields verifying a discrete version of (4). This issue was discussed in [6], in which modified Godunov schemes on Cartesian meshes were investigated that correctly capture discrete GEs. Extensions to staggered schemes on Cartesian meshes are investigated in [32].

We now investigate what happens if the use of triangular grids is preferred to Cartesian grids, for example because triangular grids are much more flexible than Cartesian ones for accurate discretizations of complex detailed geometries and for local mesh refinement. An essential characteristic of triangular meshes is that the numerical diffusion of Godunov schemes on the velocity equation is different in nature (as compared to Cartesian grids) and does not affect velocity fields that verify a discrete version of (4). References [15, 16] give a detailed explanation of this, which may be summarized by the fact that such discretely divergence free velocity fields can be characterized as the curls of  $P_1$  Lagrange functions defined by their values at the vertices of the triangular cells. This has two major consequences: we will first show that this is hardly compatible with a discrete version of (3) when the pressure gradient is defined like in the Godunov scheme with cell-centered pressures; on the other hand, we shall secondly show that this is obviously compatible with a discrete version of (3) when the pressure gradient is defined with pressures located at the vertices.

To begin with, let us note that the semi-discrete collocated scheme on triangular mesh can be written as

$$\frac{d}{dt}q_h + L_{\kappa,h}q_h = 0 \quad (\text{A.1})$$

where

$$q_h := \begin{pmatrix} r_h \\ \mathbf{u}_h \end{pmatrix} \in \mathbb{R}^{3N} \quad \text{and} \quad L_{\kappa,h}^i q_h = \begin{pmatrix} \frac{a_*}{2|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| [(\mathbf{u}_i + \mathbf{u}_j) \cdot \mathbf{n}_{ij} + \kappa_r (r_i - r_j)] \\ \frac{a_*}{2|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| [(r_i + r_j) + \kappa_u (\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}] \mathbf{n}_{ij} + \omega \mathbf{u}_i^\perp \end{pmatrix}. \quad (\text{A.2})$$

In this scheme,  $\kappa_r$  and  $\kappa_u$  represent parameters that can be used to tune the diffusion terms. The classical Godunov scheme corresponds to the case  $\kappa_r = \kappa_u = 1$ . We refer to, e.g., [15] for the construction of the scheme applied to the homogeneous wave equation. Here, we take into account the effect of the Coriolis force in the right-hand side of the semi-discrete scheme and we analyze whether steady-states of (A.1)–(A.2), i.e., fields in the kernel of the operator  $L_{\kappa,h}$ , are able to correctly represent discrete versions of the GE.

**Proposition 7.** *On a triangular mesh, with the collocated Godunov scheme, we have:*

$$\text{Ker}L_{\kappa_r \neq 0, h} = \left\{ q_h := \begin{pmatrix} r_h \\ \mathbf{u}_h \end{pmatrix} \in \mathbb{R}^{3N} \text{ such that } \exists a \in \mathbb{R} : r_i = a \text{ and } \mathbf{u} = 0 \right\}. \quad (\text{A.3})$$

Moreover, we also have

$$\text{Ker}L_{\kappa_r = 0, h} = C_h^1 \cap C_h^2 \quad (\text{A.4})$$

with the following definitions for  $C_h^1$  and  $C_h^2$ :

$$C_h^1 := \left\{ q_h := \begin{pmatrix} r_h \\ \mathbf{u}_h \end{pmatrix} \in \mathbb{R}^{3N} \text{ such that } \frac{a_\star}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| \frac{r_i + r_j}{2} \mathbf{n}_{ij} = -\omega \mathbf{u}_i^\perp \right\},$$

$$C_h^2 := \left\{ q_h := \begin{pmatrix} r_h \\ \mathbf{u}_h \end{pmatrix} \in \mathbb{R}^{3N} \text{ such that } \exists \phi_h^L \in C_\#^0(\bar{\Omega}), (\phi_h^L)|_{T_i} \in P_1(T_i), \mathbf{u}_i = (\nabla \times \phi_h^L)|_{T_i}, \forall i \in [1, N] \right\},$$

where  $C_\#^0(\bar{\Omega})$  denotes the space of periodic continuous functions over  $\bar{\Omega}$  and  $\nabla \times \phi_h^L = (\partial_y \phi_h^L, -\partial_x \phi_h^L)^T$ .

*Proof.* Denoting by  $\langle \cdot, \cdot \rangle$  the scalar product weighted by the cell areas  $|T_i|$  and following the properties of the collocated scheme [15] and using the energy conservation for the Coriolis force  $\langle \mathbf{u}_h^\perp, \mathbf{u}_h \rangle = 0$ , we obtain

$$\langle L_{\kappa,h} q_h, q_h \rangle = \frac{a_\star \kappa_r}{2} \sum_{A_{ij}} |A_{ij}| |r_i - r_j|^2 + \frac{a_\star \kappa_u}{2} \sum_{A_{ij}} |A_{ij}| |(\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij}|^2. \quad (\text{A.5})$$

We now suppose that  $\kappa_u \neq 0$  and we will consider the influence of  $\kappa_r$  on the structure of the kernel  $\text{Ker}L_{\kappa,h}$ . From (A.5), a necessary condition for a velocity field to be in the kernel is

$$\forall (i, j) : (\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij} = 0. \quad (\text{A.6})$$

This obviously implies that (using, for the second equality below, the fact that  $\sum_{A_{ij} \subset \partial T_i} |A_{ij}| \mathbf{n}_{ij} = 0$ )

$$\sum_{A_{ij} \subset \partial T_i} |A_{ij}| (\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{n}_{ij} = 0 \quad \text{and} \quad \sum_{A_{ij} \subset \partial T_i} |A_{ij}| (\mathbf{u}_i + \mathbf{u}_j) \cdot \mathbf{n}_{ij} = 0. \quad (\text{A.7})$$

When  $\kappa_r \neq 0$ , equation (A.5) also implies that  $\forall (i, j), r_i = r_j$  and it follows that  $r$  must be a constant:  $\exists a \in \mathbb{R}$  such that  $r_i = a$  for all  $i \in [1, N]$ . Then, the velocity equation of  $L_{\kappa,h} q_h = 0$  leads to  $\omega \mathbf{u}_i^\perp = 0 \forall i$ . Therefore, we get (A.3).

In the other case, when  $\kappa_r = 0$ , the second equality in (A.7) implies that the pressure equality in  $L_{\kappa,h} q_h = 0$  is verified. Moreover, we recall from [15, Lemma 5.1] that (A.6) is equivalent to the fact that there exists a Lagrange piecewise  $P_1$  conforming function, denoted by  $\phi_h^L$ , defined by its values at the vertices of the mesh and periodic over the computational domain, and constants  $(b, c) \in \mathbb{R}^2$  such that on each triangle  $T_i$  there holds

$$\mathbf{u}_i = (b, c)^T + (\nabla \times \phi_h^L)|_{T_i}. \quad (\text{A.8})$$

Moreover, from the velocity equation of  $L_{\kappa,h} q_h = 0$  and (A.6), we get that

$$\frac{a_\star}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| \frac{(r_i + r_j)}{2} \mathbf{n}_{ij} = -\omega \mathbf{u}_i^\perp, \quad (\text{A.9})$$

which means that  $q_h \in C_h^1$ . In addition, by multiplying (A.9) by  $|T_i|$  and summing over all  $i \in [1, N]$ , we obtain by periodicity that  $\sum_{i=1, N} |T_i| \mathbf{u}_i = 0$ , which implies that  $(b, c)^T$  in (A.8) vanishes. Therefore, (A.8) with  $b = c = 0$  implies that  $q_h \in C_h^2$ . In conclusion, we obtain

$$\text{Ker} L_{\kappa_r=0, h} \subset C_h^1 \cap C_h^2$$

and the converse inclusion is also true.  $\square$

**Remark 7.** *The kernel given by (A.3) clearly shows that the standard Godunov scheme ( $\kappa_r \neq 0$ ) is not able to preserve general geostrophic equilibriums. On the other hand, the situation is less clear when the diffusion in the pressure equation vanishes. Indeed, let us consider (A.4). An important property is that*

$$\frac{1}{|T_i|} \sum_{A_{ij} \subset \partial T_i} |A_{ij}| \frac{r_i + r_j}{2} \mathbf{n}_{ij}$$

is exactly the gradient on the cell  $T_i$  of the non-conforming  $P_1$  finite element function  $\tilde{r}_h^{\text{CR}}$  which is defined by the values  $\frac{r_i + r_j}{2}$  at the midpoints of the mesh edges  $A_{ij}$ . So  $C_h^1$  tells us that  $\mathbf{u}_h$  is, up to a multiplicative constant, the curl of the non-conforming  $P_1$  function  $\tilde{r}_h^{\text{CR}}$ , while  $C_h^2$  tells us that  $\mathbf{u}_h$  is the curl of a conforming  $P_1$  function  $\phi_h^{\text{L}}$ . The equality

$$(\text{curl } \phi_h^{\text{L}})|_{T_i} = \frac{\alpha_*}{\omega} (\text{curl } \tilde{r}_h^{\text{CR}})|_{T_i} \quad (\text{A.10})$$

for all  $i \in [1, N]$  provides  $2N$  constraints while  $r$  has  $N$  degrees of freedom but  $\phi_h^{\text{L}}$  only has a number of degrees of freedom equal to the number of vertices of the mesh. In a periodic triangular mesh with no holes, the number of vertices is half the number of cells, so that (A.10) is actually probably too constraining ( $2N$  constraints for only  $\frac{3N}{2}$  degrees of freedom) to admit non trivial solutions.

With the previous remark, it becomes obvious that one possibility is to define the pressure field at the vertices and replace the gradient of  $\tilde{r}_h^{\text{CR}}$  by the gradient of the conforming Lagrange  $P_1$  function defined by these values at the vertices. Of course then, we have to change the pressure evolution equation accordingly. This staggered scheme, and several variations of it, is what we explore in the body of this article.

## References

- [1] E. Audusse, S. Dellacherie, M. H. Do, P. Omnes, Y. Penel, Godunov type scheme for the linear wave equation with Coriolis source term, ESAIM: Proceedings and Surveys 58 (2017) 1–26.
- [2] F. Bouchut, Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2004.
- [3] F. Bouchut, J. Le Sommer, V. Zeitlin, Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. II. High-resolution numerical simulations, J. Fluid Mech. 514 (2004) 35–63.
- [4] J. Pedlosky, Geophysical fluid dynamics, Springer Science & Business Media, 2013.
- [5] G. K. Vallis, Atmospheric and oceanic fluid dynamics. Fundamentals and large-scale circulation, Cambridge: Cambridge University Press, 2017.
- [6] E. Audusse, M. H. Do, P. Omnes, 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 (2018) 91–129.
- [7] C. J. Cotter, D. A. Ham, Numerical wave propagation for the triangular  $P1_{DG} - P2$  finite element pair, J. Comput. Phys. 230 (8) (2011) 2806–2820.
- [8] V. Rostand, D. Y. Le Roux, G. Carey, Kernel analysis of the discretized finite difference and finite element shallow-water models, SIAM Journal on Scientific Computing 31 (1) (2008) 531–556.
- [9] C. Chen, H. Liu, R. C. Beardsley, An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries, Journal of atmospheric and oceanic technology 20 (1) (2003) 159–186.
- [10] S. Danilov, Two finite-volume unstructured mesh models for large-scale ocean modeling, Ocean Modelling 47 (2012) 14–25.
- [11] S. Danilov, Ocean modeling on unstructured meshes, Ocean Modelling 69 (2013) 195–210.
- [12] D. Le Roux, Spurious inertial oscillations in shallow-water models, J. Comput. Phys. 231 (24) (2012) 7959–7987.
- [13] S. Danilov, A. Kutsenko, On the geometric origin of spurious waves in finite-volume discretizations of shallow water equations on triangular meshes, J. Comput. Phys. 398 (2019) 108891.
- [14] J. Thuburn, C. J. Cotter, A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids, SIAM J. Sci. Comput. 34 (3) (2012) B203–B225.



- [15] S. Dellacherie, P. Omnes, F. Rieper, The influence of cell geometry on the Godunov scheme applied to the linear wave equation, *J. Comput. Phys.* 229 (14) (2010) 5315–5338.
- [16] S. Dellacherie, J. Jung, P. Omnes, Construction of a low Mach finite volume scheme for the isentropic Euler system with porosity, *ESAIM, Math. Model. Numer. Anal.* 55 (3) (2021) 1199–1237.
- [17] D. Y. Le Roux, B. Pouliot, Analysis of numerically induced oscillations in two-dimensional finite-element shallow-water models part II: Free planetary waves, *SIAM journal on scientific computing* 30 (4) (2008) 1971–1991.
- [18] E. Audusse, R. Klein, D. Nguyen, S. Vater, Preservation of the discrete geostrophic equilibrium in shallow water flows, in: *Finite Volumes for Complex Applications VI Problems & Perspectives*, Springer, 2011, pp. 59–67.
- [19] A. Kurganov, Y. Liu, M. Lukáčová-Medvid'ová, A well-balanced asymptotic preserving scheme for the two-dimensional rotating shallow water equations with nonflat bottom topography, *SIAM Journal on Scientific Computing* 44 (3) (2022) A1655–A1680.
- [20] M. Parisot, J.-P. Vila, Centered-potential regularization for the advection upstream splitting method, *SIAM J. Numer. Anal.* 54 (5) (2016) 3083–3104.
- [21] F. Couderc, A. Duran, J.-P. Vila, An explicit asymptotic preserving low Froude scheme for the multilayer shallow water model with density stratification, *J. Comput. Phys.* 343 (2017) 235–270.
- [22] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, *SIAM J. Sci. Comput.* 25 (6) (2004) 2050–2065.
- [23] V. Zeitlin, *Nonlinear dynamics of rotating shallow water: Methods and advances*, Vol. 2, Elsevier Science, 2007.
- [24] M. Castro, J. López, C. Parés, Finite volume simulation of the geostrophic adjustment in a rotating shallow-water system, *SIAM J. Sci. Comput.* 31 (1) (2008) 444–477.
- [25] A. Chertock, M. Dudzinski, A. Kurganov, M. Lukáčová-Medvid'ová, Well-balanced schemes for the shallow water equations with Coriolis forces, *Numer. Math.* 138 (4) (2018) 939–973.
- [26] A. Navas-Montilla, J. Murillo, 2D well-balanced augmented ADER schemes for the shallow water equations with bed elevation and extension to the rotating frame, *Journal of Computational Physics* 372 (2018) 316–348.
- [27] E. Audusse, R. Klein, A. Owinoh, Conservative discretization of Coriolis force in a finite volume framework, *J. Comput. Phys.* 228 (8) (2009) 2934–2950.
- [28] X. Liu, A. Chertock, A. Kurganov, An asymptotic preserving scheme for the two-dimensional shallow water equations with Coriolis forces, *Journal of Computational Physics* 391 (2019) 259–279.
- [29] Audusse, E., Dubos, V., Duran, A., Gaveau, N., Nasser, Y., Penel, Y., Numerical approximation of the shallow water equations with Coriolis source term, *ESAIM: ProcS* 70 (2021) 31–44.
- [30] V. Desveaux, A. Masset, A fully well-balanced scheme for shallow water equations with Coriolis force, *arXiv preprint arXiv:2105.08357*.
- [31] H. Zakerzadeh, The rs-imex scheme for the rotating shallow water equations with the Coriolis force, in: *International Conference on Finite Volumes for Complex Applications*, Springer, 2017, pp. 199–207.
- [32] M. H. Do, Analyse mathématique de schémas volume finis pour la simulation des écoulements quasi-géostrophiques à bas nombre de Froude, Ph.D. thesis, Sorbonne Paris Cité (2017).