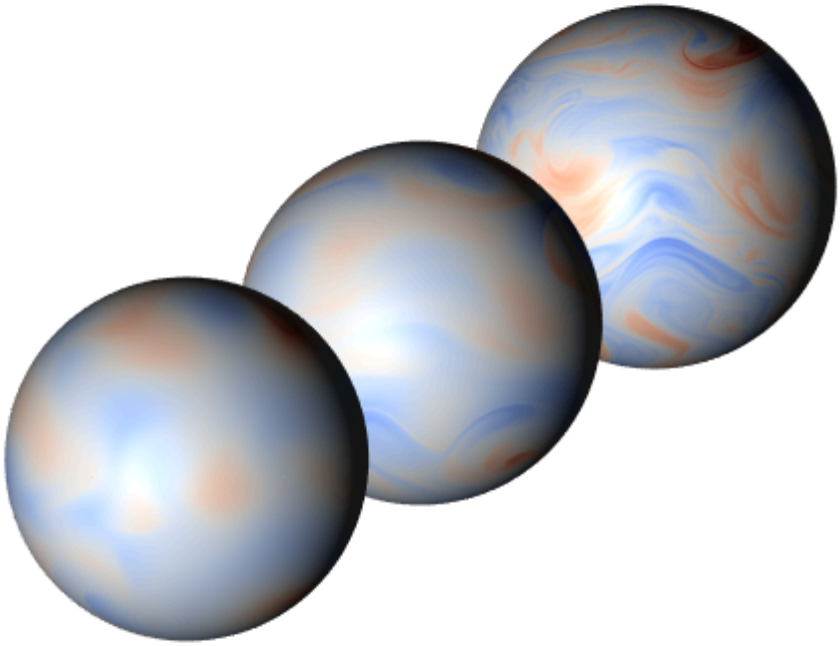
Variational integrators for the rotating shallow water equations



A thesis submitted to the School of Graduate Studies in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Scientific Computing
Memorial University of Newfoundland

August 2021

Abstract

The numerical simulation of the Earth's atmosphere plays an important role in developing our understanding of climate change. The atmosphere and ocean can be seen as a shallow fluid on the globe; here, we use the shallow water equations as a first step to approximate these geophysical flows. Then, the numerical model can only be accurate if it has good conservation properties, e.g. without conserving mass the simulation can not be physical. Obtaining such a numerical model can be achieved using numerical variational integration.

Here, we have derived a numerical variational integrator for the rotating shallow water equations on the sphere using the Euler–Poincaré framework. First, the continuous Lagrangian is discretized; then, the numerical scheme is obtained by computing the discrete variational principle. The conservational properties and accuracy of the model are verified with standard test cases.

However, in order to obtain more realistic simulations, the shallow water equations need to include physical parametrizations. Thus, we introduce a new representation of the rotating shallow water equations based on a stochastic transport principle. Then, benchmarks are carried out to demonstrate that the spatial part of the stochastic scheme preserves the total energy. The proposed random model better captures the structure of a large-scale flow than a comparable deterministic model.

Furthermore, to be able to carry out long term simulations we extend the discrete Euler–Poincaré framework with a selective decay. The selective decay dissipates an otherwise conserved quantity while conserving energy. We apply the new framework to the shallow water equations to dissipate the potential enstrophy. Then, we carry out standard benchmarks to demonstrate the conservation properties. We show that the selective decay resolves more small scales compared to a standard dissipation.

Lay summary

To have accurate weather and climate forecasts, we need to model the motion of a fluid (such as air or water) on the globe. Simulations of the shallow water equations provide a first step towards these forecasts. The shallow water equations simulate a thin layer of fluid, which is a reasonable approximation for the ocean and atmosphere. To have an accurate forecast, the simulation needs to have certain conservation properties. For example, a simulation of a lake without mass conservation may result in an empty lake. Thus, it is important to investigate numerical methods with good conservation properties.

In general, when we talk about a mathematical model, we think about its equations. However, it is equivalent to define the model using the total energy of the system. Once we define the total energy, we can calculate the model equations. To simulate the shallow water equations on the sphere, we follow the same strategy: we define a numerical energy and then compute the numerical model. Then, we analyse this model and demonstrate its conservation properties.

The computational costs of simulating a fluid flow in full detail on the globe are beyond reach today. To overcome this problem, we can use stochastic forecasting. This means that, instead of having one simulation, we carry out multiple different simulations, which enables estimating likely scenarios. Here, we derive a new stochastic representation of the shallow water equations. Then, we demonstrate that the stochastic shallow water model better captures the flow structure on the sphere compared to a non-stochastic model.

Every numerical model accumulates errors over time, which need to be reduced so that the model is stable for long term simulations. Here, we develop a numerical stabilization which is physically meaningful. In particular, the stabilization is designed to conserve the total energy and reduce numerical errors. We adopt this new stabilization for the shallow water equations and demonstrate the conservation of energy. Moreover, we show that this stabilization behaves more physically compared to standard stabilizations.

Acknowledgements

I would like to thank all the people who supported me during my time as a PhD. student. In particular, I would like to thank my supervisors Alexander Bihlo and Scott MacLachlan for their professional advice and constant feedback. Then, I am thankful for Werner Bauer for being an unofficial supervisor and hosting me for several research stays. Moreover, I would like to thank Etienne Mémin for hosting me for a research stay at the Inria Institute in Rennes, also partially financed by the Mitacs Globalink Research Award. I acknowledge the financial support of the School of Graduate Studies and the Scientific Computing Program. Finally, I thank my family for their support. Especially, I want to thank Joany for her love and support.

Statement of contributions

The work presented in Chapter 3 and 5 is the result of collaborative research between Rüdiger Brecht, Werner Bauer, Alex Bihlo, François Gay-Balmaz and Scott MacLachlan, with its intellectual property equally co-owned by all.

The work presented in Chapter 4 is the result of collaborative research between Rüdiger Brecht, Long Li, Werner Bauer and Etienne Mémin, with its intellectual property equally co-owned by all.

Contents

1. Introduction

In the context of weather prediction and climate dynamics, the motion of a fluid on a rotating spheroid under the influence of gravity needs to be modelled. This is done using numerical models, where approximations of the fully compressible Euler equations are solved. The rotating shallow water equations (SWE) are used as a first step towards developing numerical models for geophysical fluid dynamics (GFD), since they are similar to the fully compressible Euler equations and have similar conservation properties. To develop accurate numerical models, it is important to have at least some of these conservation properties preserved numerically. This is especially true for the modelling of climate dynamics, for which it is of great interest to correctly capture the long-term statistical behaviour, which can only be achieved by a numerical scheme with appropriate conservation properties. The field of geometric numerical integration, a branch of numerical analysis, focuses on the development of discretization schemes that preserve geometric structures, such as conservation laws. Here, we will use numerical variational integration, a sub-field of geometric numerical integration, to discretize the rotating SWE on the sphere.

Weather forecasts need to be precise and computationally affordable. However, the computational expense for simulating the full range of dynamical scales of a geophysical flow is beyond reach today. Thus, the effect of unresolved scales has to be modelled or parametrized. Here, we will incorporate a stochastic representation of the small scale processes called location uncertainty that relies on stochastic

transport.

Then, to be able to simulate climate dynamics, long term simulations need to be carried out. This can only be achieved with a numerical scheme that diffuses numerical errors and has coherent physical structures. Here, we will present a new variational discrete selective dissipation framework, which dissipates one quantity while conserving others.

In the following sections, we use common terms from GFD. We give a brief explanation of these terms in Appendix A.

1.1. Shallow water equations

The SWE are a model for thin layer fluids, where the vertical velocity is negligible compared to the horizontal scale. The ocean and atmosphere have a horizontal scale of some 10,000 km, and a mean depth layer of 4,000 m for the ocean or about 10 km for the atmosphere; thus, they can be modelled in a simplified way by the SWE. In addition, by using different approximations of the Coriolis force, they can simulate different geophysical waves, such as Poincaré, Rossby and Kelvin waves (see [64]). Further, the SWE have conserved quantities such as energy, mass, potential vorticity and potential enstrophy, which are of interest for the simulation of ocean and atmosphere models. All these properties make the SWE a good starting point to compare different numerical methods and their properties. Furthermore, while the SWE are still a toy model, they are an excellent (and common) first step to test a discretization methodology when moving towards models used at operational centres. For example, the SWE are used for tsunami propagation models at prewarning centres to predict the arrival of tsunami waves at potential impact zones along the coastline.

1.2. Important conserved quantities

A continuous model and its discrete analogue should have similar conserved quantities. However, the continuous SWE possess an infinite number of such conservation laws, whereas the numerical model can only reflect a finite number. In the context of numerical weather prediction and climate forecasting, the most desirable conservation properties are (see [81] for details):

1. Mass

   The dry mass in the atmosphere (or the amount of water in the ocean) does not significantly change. Therefore, in any numerical simulation a non-conservation of mass will lead to spurious behaviour of the system after a sufficiently long integration.

2. Momentum and angular momentum

   The vertical momentum is important for the hydro-static balance of the fluid flow and is on the order of the inverse of the buoyancy frequency (some minutes). The horizontal momentum is important for the geostrophic balance and change on the order of a few hours. For angular momentum, the timescale is 12 days, which corresponds to one rotation of the fluid around the globe.

3. Potential enstrophy

   The timescale for potential enstrophy is comparable to the turnover of an eddy, which is about 10 days. In a realistic flow, there is a downscale transfer of potential enstrophy, with its spectra shifted towards bigger wave numbers. For a numerical simulation there is a transfer between resolved and unresolved scales.

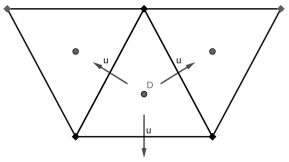
4. Energy

The energy of the system is the sum of potential and kinetic energy. There is no cascade between scales for the potential energy, but for the kinetic energy there is. If the cascade is insignificant compared to the total energy, then its conservation is indeed justified, otherwise the cascade needs to be fed back into the resolved scales.

1.3. Desirable properties for discretization schemes

Above, we stated properties which come from a physical background. Some additional desirable properties for a numerical scheme are (see [79, 83] for details):

1. Grid staggering
   It is difficult to get stable simulations when all variables are defined on the same position on the grid. For this reason, we use C-grid staggering, which has the normal velocity component (u) at the cell edges and the geopotential (D) at the cell centre. Other grid staggerings are well known to lead to worse phase speeds and group velocities (see [68]).



2. Minimal grid imprinting

   The phenomena of grid imprinting happens when geometric grid properties become visible in numerical solutions. E.g. generating finer grids of the sphere with the bisection algorithm leads to an inconsistent approximation of finite difference operators and grid imprinting, see [44]. Thus, it is desirable to use an optimized grid with a minimal grid imprinting.

3. Curl-free pressure gradient

   Vector calculus tells us that, for a continuous scalar function, h, we have $\nabla \times \nabla h \equiv 0$. In the GFD context, this means that the pressure gradient

should not change the total vorticity. We need the same property at the discrete level: $\mathrm{curl} \cdot \mathrm{grad} = 0$, for the discrete operator matrices.

4. Energy conserving pressure and Coriolis term

   As noted above, there can be a cascade in scales, such that it is debatable whether conserving the resolved total energy is desirable. However, both the pressure gradient and the Coriolis terms should be energy conserving. For the pressure gradient this means that we have a discrete version of the vector calculus identity $v \cdot \nabla h + h \nabla \cdot v = \nabla \cdot (vh)$, where h is a continuous function and v a vector field. For the Coriolis term $v \cdot v^{\perp} = 0$. This seems to be obvious, but in a discrete scheme $v^{\perp}$ and v need to be (directly or indirectly) reconstructed from $v \cdot n$ (or $v \cdot n^{\perp}$), where n is the normal vector of an edge.

5. Geostrophic balance

   The linearized SWE on an f-plane approximation maintain steady solutions in geostrophic balance. Then, the geopotential is constant in time, such that the divergence of the velocity is zero. A discrete scheme needs to have this property as well.

1.4. Importance of potential vorticity

The potential vorticity is one of the most important quantities in understanding large-scale dynamics. A geostrophically balanced flow can be entirely described by its potential vorticity field. This is stated by the invertibility principle for potential vorticity. Further, there is no flux of potential vorticity concentration across an isentropic surface (a surface of constant potential temperature) and it is conserved following the flow. That is why a numerical model of the atmosphere needs to conserve potential vorticity.

1.5. Modelling

1.5.1. Formulation

A number of GFD equations have been derived, such as the shallow water equations or quasi-geostrophic model. These have been collected in [2] which is referred to as the "ABN list" (Allen–Barth–Newberger list). A GFD model from the ABN list is considered to be good if it has a Kelvin circulation theorem and accurate energy balance.

We can classify the models from the ABN list using the Euler–Poincaré framework because it provides theorems for the energy balance and circulation. Then, the exact equations and many of their approximations can be systematically generated by changing the kinetic energy in the Lagrangian for the Hamiltonian principle and imposing constraints. As a consequence, the Lagrangian changes, but not the Euler–Poincaré form of the equations. The equations can then be implemented numerically in a consistent way.

1.5.2. Spatial discretization

There are many different approaches to discretize partial differential equations, examples of different schemes in the GFD context are:

- Spectral methods: STSWM (NCAR Spectral Transform Shallow Water Model). It is based on spherical harmonics and uses a latitudinal-longitudinal grid. Spectral models are often used as a high-order reference solution.
- Finite differences: The Arkawa and Lamb 1981 scheme [5]. First scheme having multiple conservation properties.
- Finite volumes: TRiSK [74], based on a general polygonal or spherical polygonal grid. It has potential enstrophy or total energy conservation.

- Finite elements: Finite Element Exterior Calculus (FEEC) [24], which allows high-order simulations, and conserves energy and potential enstrophy.

One of the first schemes having multiple conservative properties is the Arakawa and Lamb 1981 scheme, which can be interpreted as a Hamiltonian-based scheme. However, it has some disadvantages: it is restricted to a square grid, it is low-order (otherwise it loses its conservative properties) and, for small wavelength/Δx, it has poor wave representation. Discretizations like TRiSK and FEEC tried to overcome these issues.

Most, if not all, of the models up to now discretized the equations derived from the Lagrangian through the Hamiltonian principle and use standard time integrators for the temporal discretization. Here, we follow a different approach: first, the Lagrangian itself is discretized, then the numerical equations are calculated, and finally we use a variational time integrator. By doing this, we obtain a fully variational integrator, which is a new approach for discrete schemes and has not been done before in the GFD context.

1.6. Sub-grid model

Even with today's computational power it is not possible to simulate the whole dynamical scale of a geophysical flow. Thus, the unresolved sub-grid component of the velocity is unknown or uncertain. Uncertainty quantification also deals with modelling the unresolved scales by introducing randomness (noise) to the model. This can be introduced to the model in many different ways. Nevertheless, the noise needs to have certain properties to capture the flow dynamics. When carrying out a forecast with multiple ensemble members, the spread (based on the noise) of the ensemble has to be efficient. An efficient spread means that we only need a few ensemble members, which are distributed around the true flow dynam-

ics.

A natural way to introduce noise in the simulation is to perturb the initial conditions (PIC). The European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP), among others, are using this method. They use different techniques to perturb the initial conditions to maximise the ensemble variance. But, the PIC method is known to underestimate the error of the state variable dynamics, meaning it is under-dispersive, see [13, 40].

An alternative method tries to simulate the unresolved scales and introduce noise in the small scales during the simulation. Introducing noise in the small scales leads to a back scattering of energy to the large scales. Formalizing this idea can be done by introducing stochastic transport. Here, we follow the Location Uncertainty (LU) framework, where the flow is assumed to be stochastic and the velocity can be decomposed into large scale and small scale uncertainty components. In this framework, the stochastic transport theorem describes the time differential of a scalar integrated over a volume. Consequently, the energy is shown to be conserved. The energy conserving property of the LU model motivates a structure preserving discretization.

1.7. Dissipation

In general, CFD models need dissipation for many purposes, such as cleaning up numerical noise generated by the discretization, Gibbs ringing in spectral modes, poor initializations, grid-scale forcing from the physical parametrization or weak computational stability. Moreover, dissipation also has a physical meaning. In

2D turbulence, enstrophy cascades from large scale to small scales, resulting in an accumulation of enstrophy at the smallest grid scale, see [17, 73]. Thus, enstrophy has to be dissipated to carry out stable simulations.

A main motivation for geometric integrators, such as variational integrators, is their excellent long term behaviour. Nevertheless, while carrying out long term simulations, numerical errors still accumulate. This motivates the design of a variational stabilization. Here, we extend the variational discretization framework with a selective decay. We then apply this new framework and obtain a discretization of the RSW with enstrophy dissipation and energy conservation, such that physically meaningful long term simulations can be carried out.

1.8. Literature overview

In this section, we give a short overview of what has been done in the field of variational integration, location uncertainty ,and selective Casimir dissipation.

Variational integration

Geometric numerical integrators are discrete schemes that preserve important geometric properties of differential equations. Such properties can be conservation laws, Lie group symmetries, and Hamiltonian or variational formulations. Examples of geometric integrators preserving at least one geometric property are given in [8, 9, 12, 15, 16, 21, 42, 55, 58, 65, 77, 87] Preserving geometric properties can guarantee long term stability, consistency in statistical properties and the prevention of drift in stationary solutions, see e.g. [42, 55, 88]. For atmospheric dynamics, the accurate representation of the statistical properties is crucial, thus geometric integrators are receiving more attention in recent years.

Here, we focus on a variational integrator that possess a number of desirable

properties, such as a discrete form of Noether's theorem. This theorem guarantees the preservation of conserved quantities related to the variational symmetries and stability for exponentially long time periods, see e.g. [42, 58].

A relevant variational integrator for incompressible fluids was developed in [65]. The discretization relies on first discretizing the variational principle and then deriving the numerical scheme. This particular approach was extended to various equations of incompressible fluid dynamics with advected quantities in [36]. In [26], the method was extended to rotating and stratified fluids, and in [9] to anelastic and pseudo-incompressible fluids. The extension to the shallow water equations was done in [8].

However, all of these extension have been simulated on the plane. The framework is naturally coordinate independent and this feature has not been utilized. Moreover, for a GFD dynamical core the model needs to run on spherical geometry. Here, as a next step we extend the scheme to the sphere.

Location uncertainty (Stochastic sub-grid model)

Despite the increasing power of computational resources and the availability of high quality observations, a precise description of geophysical flows over their whole dynamical scales is completely beyond reach today. To face these challenges, numerous efforts are taking place to build an ever-increasing quality, quantity, duration, and integration of all observations. Yet, for ocean models, the unresolved small scales and associated fluxes are always accounted for by simple mathematical models, i.e. parametrizations. Although the development of more efficient sub-grid representations remains a very active research area, the possible separation between relatively low-frequency, large scale patterns and transient, small-scale fluctuations, strongly invites to consider stochastic representations of the geophysical dynamics, see [31, 39, 41, 57].

Here, we consider a specific stochastic model, called Location Uncertainty, de-

rived in [62]. This random model uses a stochastic transport operator [69] which involves a multiplicative random forcing, a heterogeneous diffusion and a corrected advection resulting from the inhomogeneity of the random flow. The resulting equations then also conserve the total energy. With this stochastic transport principle, stochastic representations under Location Uncertainty of large-scale geophysical dynamics have been derived, see e.g. [10, 22, 69]. The Location Uncertainty representations for oceanic quasi-geostrophic flows performed very well, see e.g. [10, 11, 70, 71].

However, none of these models have yet demonstrated numerical energy conservation. In this thesis, we use this stochastic transport principle to derive the shallow water equations under Location Uncertainty which conserve energy. Then, to obtain a numerical scheme with good energy conservation we extend the variational discretization of the shallow water equations with approximations of the new stochastic terms.

Selective Casimir dissipation

The shallow water equations simulate thin layered fluid flows, which are approximately two dimensional. The principles of two dimensional turbulence give insight into their flow dynamics. For two dimensional turbulence, enstrophy cascades to higher wave numbers and energy to smaller wave numbers, see [52, 56]. This led to the selective decay hypothesis [59], meaning that the enstrophy should be dissipated (otherwise it would accumulate at the grid scale) and energy should be conserved. Therefore, enstrophy dissipating and energy conserving numerical schemes were developed for the shallow water equations, see e.g. [4, 76].

A more general method to derive models that conserve energy while dissipating another otherwise conserved quantity was introduced in the Lie-Poisson framework in [37, 38]. Since this selective dissipation framework has not been applied to discrete models of geophysical fluid dynamics, here, as a first step, we use

the framework to selectively dissipate enstrophy for the shallow water equations. Then, we apply the variational discretization framework to obtain the discrete scheme.

1.9. Thesis overview

This thesis provides a first step to answer the question whether the variational discretization framework can be used to obtain competitive next generation GFD dynamical cores. This framework [8, 9, 26, 36, 65] has not been explored on the sphere, which is of great interest when modelling the atmosphere and ocean. Moreover, this framework has not been combined with any sub-grid model nor dissipation terms. Here, we present a new stochastic representation [62, 69] of the SWE and combine it with the discrete variational integrator. Furthermore, we extend the framework with a new discrete selective decay [37, 38].

In Chapter 2, we review the mathematical concepts used in the following chapters. First, we introduce the assumptions for the Euler–Poincaré framework and the mathematical background from differential geometry. Finally, we present the Euler–Poincaré framework. Then, we give an example to explain the Euler–Poincaré reduction.

In Chapter 3, we extend the discrete variational integrator for the rotating SWE to the sphere. We carry out standard benchmarks to analyse the conservation properties.

In Chapter 4, we derive a new stochastic representation of the rotating SWE using the location uncertainty principle. Then, we carry out a simulation to analyse the energy conservation. Furthermore, we carry out ensemble forecasts to analyse

the spread compared to a deterministic model.

In Chapter 5, we present a new discrete selective decay scheme in the Euler–Poincaré framework. We first review the continuous selective decay scheme, and then derive the discrete selective decay scheme, mimicking the continuous level. We apply the framework to the SWE to dissipate enstrophy while conserving energy.

In Chapter 6, we summarize the results and discuss possible projects for future work.

2. Background

A natural way to obtain a discretization of a system of differential equations is to approximate the derivatives. Usually this is done using finite differences, finite volume and finite element methods, which are based on a local approximation of the derivatives using values around a given point. However, a main drawback of a local approximation is that differential equations may possess geometric properties of a global nature, such as conservation laws or an underlying variational structure. Geometric numerical integrators (such as variational integrators) provide alternative numerical schemes that preserve these geometric properties on a discrete level. The variational discretization is described in Chapters 3 and 5. Nevertheless, a key ingredient of this variational discretization is the Euler–Poincaré reduction: the reduction of Hamilton's Principle from the Lagrangian description of fluid dynamics to the Eulerian description. This reduction relies on the relation between Lie groups and Lie algebras, as well as their representations. Thus, to introduce the Euler–Poincaré reduction on the continuous level, in this chapter we first review some aspects of Lie groups and Lie algebras. Then, to better understand the reduction, we give an example in $\mathbb{R}^3$ for the rotation of a rigid body.

2.1. Lie group and Lie algebra

Definition 1. A real **Lie group** G is a finite dimensional manifold with a group structure, such that the group action and group inversion are smooth maps.

Definition 2. A **Lie algebra** is a vector space $\mathfrak{g}$ endowed with a skew-symmetric bilinear form $[.,.]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which satisfies the Jacobi identity: $[x,[y,z]] + [z,[x,y]] + [y,[z,x]] = 0,\ \forall x,y,z \in \mathfrak{g}$.

Now we will relate the Lie group to a Lie algebra. Let G be a Lie group and TG its tangent bundle. Consider $a \in G$ then $L_a\colon G \to G, g \mapsto ag$ and $R_a\colon G \to G, g \mapsto ga$ are the left and right translations. Their differentials at the identity define maps dL_a and dR_a on the tangent space.

Let $\mathrm{Vect}(G)$ be the set of all vector fields on G. A vector field $u \in \mathrm{Vect}(G)$ is said to be left/right invariant if

$$dL_a(u(g)) = u(L_a(g)) = u(ga), \qquad dR_a(u(g)) = u(R_a(g)) = u(ag),$$

which means that the following diagram commutes (same holds for R_a instead of L_a):

$$\begin{array}{ccc} TG & \xrightarrow{dL_a} & TG \\ {\scriptstyle u}\uparrow & & \uparrow{\scriptstyle u} \\ G & \xrightarrow{L_a} & G \end{array}$$

Now, we define $\mathfrak{g} := T_eG$ as the tangent space at the identity, and $\mathrm{Vect}(G)^L$ as the left and $\mathrm{Vect}(G)^R$ as the right invariant vector fields on G.

Proposition 3. *The space of all left/right invariant vector fields is isomorphic to $\mathfrak{g} = T_eG$, by the map $\Phi\colon \mathrm{Vect}(G)^L \to \mathfrak{g}, u \mapsto u(e)$. The same map works for $\mathrm{Vect}(G)^R$.*

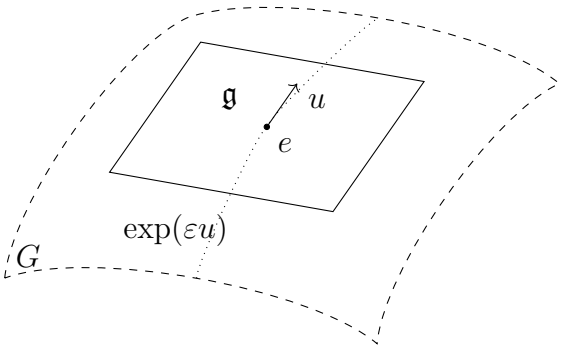


Figure 2.1.: Relation between elements in $\mathfrak{g}$ and G via the exponential map.

One can check that the Lie bracket of left/right invariant vector fields is left/right invariant. Thus, $\mathfrak{g} = T_eG$ is indeed a Lie algebra. This shows that the associated Lie algebra $\mathfrak{g}$ of a Lie Group G are the left/right invariant vector fields on G.

The map $\exp\colon \mathfrak{g} \to G$ relates the elements of $\mathfrak{g}$ and G. Details on how to define such a map are given in e.g. [63]. The **exponential map** has the properties

- $\exp((\varepsilon_1 + \varepsilon_2)u) = \exp(\varepsilon_1 u)\exp(\varepsilon_2 u)$,
- $\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \exp(\varepsilon u) = u$.

Now, $\exp(\varepsilon u)$ for $\varepsilon \in \mathbb{R}$ defines a flow in G. By computing $u_g = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \exp(\varepsilon u)g$ we obtain the elements in $\mathfrak{g}$. In Fig. 2.1, we visualize the connection between a Lie group and Lie algebra by the exponential map.

2.2. Adjoint representation

Definition 4. The representation of a Lie group G on a vector space V is a group homomorphism $\phi\colon G \to GL(V)$. Similarly, a representation of a Lie algebra $\mathfrak{g}$ on V is an algebra homomorphism $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

For a Lie group G, the *group conjugation* $K_g(h) = ghg^{-1}$ defines a diffeomorphism on G, and its differential $\mathrm{Ad}_g := dK_g\colon TG_{|_h} \to TG_{|_{K_{g(h)}}}$ is an invertible

map. Thus, $\mathrm{Ad}\colon G \to GL(\mathfrak{g}), g \mapsto \mathrm{Ad}_g$ defines a representation, and it is called the **adjoint representation**. Then, for the Lie algebra $\mathfrak{g}$ the adjoint representation is the map $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}), u \mapsto (d\,\mathrm{Ad})(u)$.

We can compute the adjoint representation on the Lie algebra by:

$$(\mathrm{ad}(u))(v) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathrm{Ad}\Big(\exp(\varepsilon u)\Big)v, \qquad v \in \mathfrak{g}.$$

Example 5.

Let us consider the case of a matrix Lie group G, here we have $K_A(B) = ABA^{-1}$. We compute Ad by:

$$\mathrm{Ad}_g(u) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} K_g(\exp(\varepsilon u)) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} g\exp(\varepsilon u)g^{-1} = gug^{-1}.$$

And now we compute the adjoint representation on the Lie algebra:

$$(\mathrm{ad}(u))(v) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathrm{Ad}(\exp(\varepsilon u))v = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \exp(\varepsilon u)v\exp(-\varepsilon u) = uv - vu = [u, v].$$

The last relation between the adjoint representation and the Lie bracket generalizes to:

Proposition 6. *For any $u, v \in \mathfrak{g}$ we have:*

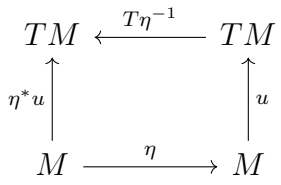
$$(\mathrm{ad}\, u)(v) = [u, v].$$

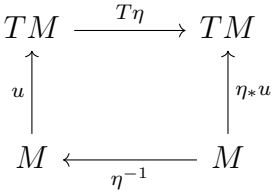
For Ad and ad we denote their dual (**coadjoint**) by Ad^* and ad^*, and they are defined by the identity:

$$\langle \mathrm{Ad}_g^* u, v\rangle = \langle u, \mathrm{Ad}_g v\rangle, \qquad\qquad \langle \mathrm{ad}_w^* u, v\rangle = \langle u, \mathrm{ad}_w v\rangle. \tag{2.1}$$

2.3. Lie derivative

Let $\eta\colon M \to M$ be a diffeomorphism. The **pull-back** $\eta^* u\colon M \to TM$ and the **push-forward** $\eta_* u\colon M \to TM$ of η for a vector field $u\colon M \to TM$ are defined by





For an infinite dimensional manifold M, the **Lie derivative** of v along u is defined by

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \eta^* v =: \mathcal{L}_u v,$$

where $\eta(\varepsilon) = \exp(\varepsilon u)$ is the flow of the vector field u.

The adjoint representation was defined by $K_\eta(h) = \eta h \eta^{-1}$. Then, we have for $h(\varepsilon) = \exp(\varepsilon v)$, using the chain rule,

$$\mathrm{Ad}_\eta(v) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \eta\Big(h(\varepsilon)(\eta^{-1})\Big) = T\eta \cdot v_{|_{\eta^{-1}}} = \eta_*(v).$$

For the infinitesimal adjoint action, we have

$$[u, v] = \mathrm{ad}_u(v) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \eta^*(v) = \mathcal{L}_u v.$$

2.4. Notation and assumptions

We consider the motion of a fluid in terms of the movement of particles and, then, we reduce this to its velocities. Let M be a smooth compact manifold (e.g., the sphere), on which the fluid moves. The elements $X \in M$ are fluid particles and their coordinates are the particle labels. We define $\mathrm{Diff}(M) := \{\eta \mid \eta\colon M \to M \text{ is a diffeomorphism}\}$, and call $\eta \in \mathrm{Diff}(M)$ a configuration. Now, at each

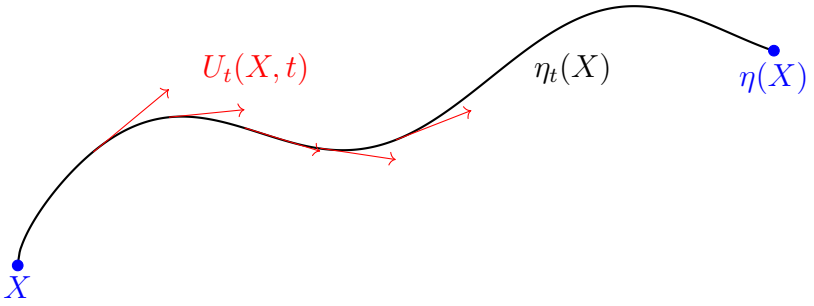


Figure 2.2.: Visualization of a motion η_t and its Lagrangian velocity U_t.

time t the configuration of the fluid is described by η and the motion of the fluid is defined by the path $\eta_t \in \mathrm{Diff}(M)$ (see Fig. 2.2). The motion relates to the Lagrangian velocity, $U(X,t)$, by:

$$U(X,t) := \frac{\partial}{\partial t}\Big|_X \eta_t(X).$$

The points $x(X,t) := \eta_t(X)$ form a path in M, they are called Eulerian points and define the Eulerian velocity:

$$u(x,t) = u(\eta_t(X),t) := U(X,t) = \frac{\partial}{\partial t}\Big|_X \eta_t(X).$$

The Eulerian velocity is a time dependent vector field $u(x,t) =: u_t(x) \in \mathrm{Vect}(M) := \{u \mid u\colon M \to TM \text{ is a vector field}\}$, and we have $U_t = u_t \circ \eta_t$. If we suppress "t" and write $\dot{\eta}$ for U we have $u = \dot{\eta} \circ \eta^{-1}$.

The quantities transported by the flow of the fluid, such as the mass, are described by elements of $(\mathrm{Vect}(M))^*$. The elements of $(\mathrm{Vect}(M))^*$ (the dual space of $\mathrm{Vect}(M)$) can be understood as one-forms. $\mathrm{Diff}(M)$ with the concatenation of functions resembles a Lie group (but it is infinite dimensional), and $\mathrm{Vect}(M)$ has the structure of a Lie algebra.

Definition 7. We define the **Lagrangian** of a mechanical system as the function

$$L\colon TG \times V^* \to \mathbb{R},$$

which is right invariant, such that it induces a function, the **reduced Lagrangian**:

$$\ell\colon \mathfrak{g} \times V^* \to \mathbb{R}, \qquad \ell(u,a) = L(id, \dot{\eta}\circ\eta^{-1}, a_0\eta^{-1}) = L(\eta,\dot{\eta},a_0). \tag{2.2}$$

The evolution of a is given by the action $\dot{a} = -a\cdot u$.

If the Lagrangian L is left invariant, then we have the action of η^{-1} on the left.

2.5. Euler–Poincaré reduction

First, we define the variational derivative $\frac{\delta L}{\delta u}$ for a functional $L\colon M \to \mathbb{R}$ and $v, u \in M$:

$$\int_M \frac{\delta L}{\delta u}\cdot v\ dx = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} L(u+\varepsilon v), \qquad v\equiv 0 \text{ on } \partial M, \varepsilon\in\mathbb{R}$$

Example 8. Let us consider $L = \frac{1}{2}\int_M \langle u,u\rangle$ as the kinetic energy, then we have:

$$\int_M \frac{\delta L}{\delta u}\cdot v\ dx = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} L(u+\varepsilon v),$$

expanding the right hand side gives:

$$\begin{aligned}
&= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\frac{1}{2}\int_M \langle u+\varepsilon v, u+\varepsilon v\rangle\ dx \\
&= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\frac{1}{2}\int_M \langle u,u\rangle + 2\varepsilon\langle u,v\rangle + \varepsilon^2\langle v,v\rangle\ dx \\
&= \frac{1}{2}\int_M 2\langle u,v\rangle\ dx = \int_M \langle u,v\rangle\ dx.
\end{aligned}$$

Thus, we see that $\frac{\delta L}{\delta u} = \langle u, \cdot \rangle =: u^\flat$, here we use the shorthand notation $u^\flat$ for the one-form associated with u.

Theorem 9 (Euler–Poincaré Reduction)**.** *Let L be a Lagrangian and ℓ the reduced Lagrangian, see Def. 7. The following statements are equivalent:*

1. $\delta \int_a^b L \, dt = 0$ *holds on G for variations $\delta\eta_t$ with $0 = \delta\eta(a) = \delta\eta(b)$.*
2. *η_t satisfies the Euler–Lagrange equations for L, for details regarding the equations see [48].*
3. $\delta \int_a^b \ell(u, a) \, dt = 0$ *holds on $\mathfrak{g} \times V^*$, using variations of the form $\delta u = \dot{u} \pm [v, u]$ and $\delta a = -a \cdot v$, where $v(t) \in \mathfrak{g}$ with $v(a) = v(b) = 0$.*
4. *The Euler–Poincaré equations hold:*

$$\frac{d}{dt}\frac{\delta \ell}{\delta u} = \pm \mathrm{ad}^*_u \frac{\delta \ell}{\delta u} + \frac{\delta \ell}{\delta a} \diamond a \qquad \text{and} \qquad (\partial_t + \mathcal{L}_u)a = 0.$$

For $b \in V, a \in V^*$ and $w \in \mathfrak{g}$, the $\diamond \colon V^* \times V \to \mathfrak{g}^*$ operator is defined by:

$$\langle b, a \cdot w \rangle = -\langle b \diamond a, w \rangle. \tag{2.3}$$

Note that $\pm$ depends on L being left or right invariant.

For details and the proof see [47].

The above theorem shows that formulating the Lagrangian in terms of the velocity and calculating the variational principle is equivalent to solving the Euler–Lagrange equations for the Lagrangian depending on the particles. Thus, we will use the reduced Lagrangian to calculate the equations of motion.

2.6. Rigid body example

To better understand the Euler–Poincaré reduction, we give an example in $\mathbb{R}^3$ without advected parameters. Let $R \in SO(3)$ give the configuration of a body and $L(R, \dot{R})$ be the associated left invariant Lagrangian. Then, a reference label point X is mapped to $x = R\,X$. For a motion, meaning that R is time dependent, we have $u = \dot{R}\,X = \dot{R}R^{-1}\,x$. Since $\dot{R}R^{-1}$ is skew symmetric, there is a vector (the angular velocity vector) ω such that $u = \dot{R}R^{-1}\,x = \omega \times x$. Now we have

$$R^{-1}\dot{R}\,X = R^{-1}\dot{R}R^{-1}\,x = R^{-1}(\omega \times x) = \underbrace{R^{-1}\omega}_{=:\Omega} \times \underbrace{R^{-1}x}_{=X}.$$

The vector Ω relates to a matrix

$$\tilde{\Omega} := \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix}$$

such that $\tilde{\Omega}X = \Omega \times X = R^{-1}\dot{R}\,X$ and we have $\tilde{\Omega} = R^{-1}\dot{R}$.

Thus, the curve $R(t)$ satisfying the Euler-Lagrange equations for $L(R, \dot{R})$ is equivalent to $\tilde{\Omega}$ satisfying the Euler–Poincaré equations for $\ell(\tilde{\Omega})$.

Let us now compute the variations $\delta\tilde{\Omega}$. We define $S = R^{-1}(\delta R)$ this gives

$$\dot{S} = R^{-1}(\delta\dot{R}) - R^{-1}\dot{R}R^{-1}(\delta R) = R^{-1}(\delta\dot{R}) - \tilde{\Omega}S.$$

Then, we obtain the variation of $\tilde{\Omega}$ by differentiating with respect to R

$$\begin{aligned} \delta\tilde{\Omega} &= R^{-1}(\delta\dot{R}) - R^{-1}(\delta R)R^{-1}\dot{R} = \dot{S} + \tilde{\Omega}S - S\tilde{\Omega} \\ &= \dot{S} + [\tilde{\Omega}, S]. \end{aligned}$$

To compute the Euler–Poincaré equations, we consider the reduced Lagrangian

$\ell(\tilde{\Omega})$ and compute:

$$0 = \delta \int_0^T \ell(\tilde{\Omega})\ dt = \int_0^T \left\langle \frac{\delta \ell}{\delta \tilde{\Omega}}, \delta\tilde{\Omega} \right\rangle\ dt = \int_0^T \left\langle \frac{\delta \ell}{\delta \tilde{\Omega}}, \dot{S} + [\tilde{\Omega}, S] \right\rangle\ dt$$
$$= \int_0^T \left\langle -\frac{d}{dt}\frac{\delta \ell}{\delta \tilde{\Omega}}, S \right\rangle + \left\langle \frac{\delta \ell}{\delta \tilde{\Omega}}, [\tilde{\Omega}, S] \right\rangle\ dt = \int_0^T \left\langle -\frac{d}{dt}\frac{\delta \ell}{\delta \tilde{\Omega}}, S \right\rangle + \left\langle \mathrm{ad}^*_{\tilde{\Omega}}\frac{\delta \ell}{\delta \tilde{\Omega}}, S \right\rangle\ dt, \tag{2.4}$$

where we used integration by parts and Eq. 2.1. The above calculation holds for any S and we obtain the Euler–Poincaré equations

$$\frac{d}{dt}\frac{\delta \ell}{\delta \tilde{\Omega}} = \mathrm{ad}^*_{\tilde{\Omega}}\frac{\delta \ell}{\delta \tilde{\Omega}}.$$

To obtain the rigid body equations in concrete terms, we define the reduced Lagrangian to be the kinetic energy

$$\ell(\tilde{\Omega}) = \frac{1}{2} Tr((\mathbb{I}\tilde{\Omega})^\top \tilde{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)$$

where I_1, I_2, I_3 are the moments of inertia and $\mathbb{I} = diag(I_1, I_2, I_3)$. Then, similar to example 8, we have $\frac{\delta \ell}{\delta \tilde{\Omega}} = \mathbb{I}\tilde{\Omega}$. In the case that the inner product in Eq. (2.4) is defined by the trace operator Tr the adjoint operator reduces to the commutator $\mathrm{ad}^* = [.,.]$. Thus, the rigid body motion is given by

$$\frac{d}{dt}\mathbb{I}\tilde{\Omega} = [\mathbb{I}\tilde{\Omega}, \tilde{\Omega}].$$

3. Variational integrator for the rotating shallow-water equations on the sphere

Published in Quarterly Journal of the Royal Meteorological Society, see [20]

Rüdiger Brecht[†], Werner Bauer[‡], Alexander Bihlo[†], François Gay-Balmaz[§] and Scott MacLachlan[†]

[†] Department of Mathematics and Statistics, Memorial University of Newfoundland,
St. John's (NL) A1C 5S7, Canada

[‡] Imperial College London, Department of Mathematics, 180 Queen's Gate, London SW7 2AZ, United Kingdom.

[§] École Normale Supérieure/CNRS, Laboratoire de Météorologie Dynamique, Paris, France.

E-mail: rbrecht@mun.ca, w.bauer@imperial.ac.uk, abihlo@mun.ca, gaybalma@lmd.ens.fr, smaclachlan@mun.ca

Abstract

We develop a variational integrator for the shallow-water equations on a rotating sphere. The variational integrator is built around a discretization of the continuous Euler–Poincaré reduction framework for Eulerian hydrodynamics. We describe the discretization of the continuous Euler–Poincaré equations on arbitrary simplicial meshes. Standard numerical tests are carried out to verify the accuracy and the excellent conservational properties of the discrete variational integrator.

3.1. Introduction

Geometric numerical integration is the branch of numerical analysis devoted to the development of discretization schemes that preserve important geometric properties of differential equations. Examples of geometric properties of practical interest include symplectic forms and Hamiltonian formulations, Lie group symmetries and conservation laws, volume forms, variational formulations, maximum principles and blow-up properties. Numerical integrators that discretely preserve one or more of the aforementioned geometric properties are presented e.g. in [8, 9, 12, 15, 16, 21, 42, 55, 58, 65, 77, 87].

A main motivation behind the development of integrators capable of preserving geometric properties of differential equations is their, in general, superior long-term behaviour. Preserving geometric properties can guarantee arbitrarily long-term stability, consistency in statistical properties and the prevention of systematic drift in stationary or periodic solutions, see e.g. [42, 55, 88] for further details.

Recent years have seen an increased interest in geometric numerical integration for models of atmospheric dynamics. This is natural since long time integrations and the accurate representation of the statistical properties of these models lie at the heart of climate prediction and turbulence modeling. A particular model that

has received considerable attention is the rotating shallow-water (RSW) equations, both in the plane and on the sphere.

Energy and enstrophy preserving integrators for the shallow-water equations were developed as early as in the seminal paper of [5]. There it was recognized that preserving energy and enstrophy in a finite difference discretization of the shallow-water equations is crucial to guaranteeing the numerical stability of typical flow regimes. In recent years, considerable effort was devoted to the development of structure-preserving integrators on general structured and unstructured grids, see e.g. [8, 14, 27, 30, 73, 82, 83].

In [8], a variational integrator for the rotating shallow-water equations in the plane was proposed. Variational integrators rest on first discretizing the continuous variational principle underlying the governing equations of interest, and then deriving the numerical scheme as a discrete system of Euler–Lagrange equations [58]. Variational integrators possess a number of desirable properties, including compatibility with a discrete form of Noether's theorem that guarantees the exact numerical preservation of those conserved quantities related to the variational symmetries of the discretized governing equations, as well as stability for exponentially long time periods [42, 58].

While most of the work on variational integration was devoted to ordinary differential equations, recent years have seen an increased interest in the partial differential equation case, see e.g. [65] and [8, 9, 26] for some applications of the variational methodology to important models of geophysical fluid dynamics. Variational integrators for the partial differential equations of fluid dynamics are designed by replacing the continuous configuration space of the model equation, represented as an appropriate infinite-dimensional Lie group, by a suitable finite dimensional matrix Lie group on which the variational principle can be applied in both its Lagrangian and Eulerian versions, thanks to an application of the Euler–Poincaré reduction theorem. The purpose of the present paper is to extend the

variational integrator proposed in [8] for the shallow-water equations in the plane to the shallow-water equations on the rotating sphere.

The further organization of the paper is as follows. In Section 3.2, we give a brief summary of the variational description of the shallow-water equations using the Euler–Poincaré formulation. Section 3.3 is devoted to the description of the discretization of the continuous Euler–Poincaré formulation, originally presented in [8], and the representation of the variational integrator on the icosahedral mesh geometry used to approximate the sphere. In Section 3.4, we verify the consistency of the corresponding approximations of the standard differential operators. Test cases and numerical benchmarks showcasing the behaviour of the variational integrator for the shallow-water equations are given in Sections 3.5 and 3.6. The conclusions and thoughts for future research within this field of geometric numerical integration are found in Section 3.7.

3.2. Variational principle for the rotating shallow-water equations

In absence of irreversible processes, the equations of motion of fluid dynamics can be derived via the Hamilton principle

$$\delta \int_0^T L(\varphi, \dot{\varphi}) dt = 0,$$

with respect to variations $\delta\varphi$ vanishing at $t = 0$ and $t = T$. Here L is the Lagrangian function of the fluid, given by the kinetic energy minus the internal energy, and expressed in terms of the Lagrangian fluid trajectory φ and Lagrangian fluid velocity $\dot{\varphi}$. Following this point of view, the configuration space for fluid dynamics, away from shocks, is the group $\mathrm{Diff}(D)$ of diffeomorphisms φ of the fluid domain D. While in the Lagrangian (or material) description this principle is

a straightforward extension of the Hamilton principle of classical mechanics, in the Eulerian (or spatial) description the variational principle is more involved, since it uses constrained variations. It can be rigorously justified by applying the process of Euler–Poincaré reduction [48], which directly gives the general form of the Eulerian variational principle induced by the Hamilton principle of fluid mechanics. We refer to [46] for an application to the equations of geophysical fluid dynamics (GFD). Since this principle is central to the derivation of the numerical scheme, we quickly review it below for the shallow-water case. A more detailed review of Euler–Poincaré reduction and its application to the shallow water equations as well as to simpler examples is presented in §3.8.3 in the Appendix.

Let us consider the rotating shallow-water (RSW) dynamics on the sphere $\mathcal{S}$ of radius R. The sphere is naturally endowed with a Riemannian metric γ induced from the standard Euclidian metric on $\mathbb{R}^3$ and with a volume $d\sigma$ associated to γ, see §3.8.1. In terms of latitude (θ) and longitude (λ), we have $\gamma = R^2 d\theta^2 + R^2 \cos\theta d\lambda^2$ and $d\sigma = R^2 \cos\theta d\theta \wedge d\lambda$, but our approach is geometrically intrinsic, i.e., independent of any choice of coordinates on $\mathcal{S}$. In the Eulerian description, the variables are the fluid velocity $\mathbf{u}$ and the fluid depth h, defined in terms of the Lagrangian variables φ and $\dot{\varphi}$, with $\varphi \in \mathrm{Diff}(\mathcal{S})$, as

$$\mathbf{u} = \dot{\varphi} \circ \varphi^{-1} \quad \text{and} \quad h = h_0 \bullet \varphi^{-1} := (h_0 \circ \varphi^{-1}) J\varphi^{-1}, \tag{3.1}$$

where h_0 is the initial fluid depth and $J\varphi^{-1}$ is the Jacobian of the diffeomorphism φ^{-1} with respect to the metric. The second relation in (3.1) is the *natural action* of diffeomorphisms on densities, that we have denoted using $\bullet$, see also §3.8.1 for more details.

The Lagrangian for rotating shallow-water fluids in Eulerian coordinates is given by

$$\ell(\mathbf{u}, h) = \int_{\mathcal{S}} \Big[\frac{1}{2} h\, \gamma(\mathbf{u}, \mathbf{u}) + h\, \gamma(\mathbf{R}, \mathbf{u}) - \frac{1}{2} g (h + B)^2 \Big] d\sigma, \tag{3.2}$$

where B is the bottom topography, $h + B$ describes the free surface elevation of the fluid, g is the gravitational acceleration and $\mathbf{R}$ is the vector potential of the angular velocity of the Earth. We recall that γ is the Riemannian metric on $\mathcal{S}$ induced from the Euclidean metric on $\mathbb{R}^3$.

Given this Lagrangian, the equations of motion follow from the Euler–Poincaré variational principle given by

$$\delta \int_0^T \ell(\mathbf{u}, h)dt = 0, \tag{3.3}$$

with respect to constrained variations of the form $\delta\mathbf{u} = \partial_t\mathbf{v} + [\mathbf{u}, \mathbf{v}]$ and $\delta h = -\operatorname{div}(h\mathbf{v})$, where $\mathbf{v}$ is an arbitrary time dependent vector field on $\mathcal{S}$, vanishing at the endpoints $t = 0$ and $t = T$, and where div denotes the divergence on the sphere, associated to the metric, see §3.8.1. The form of the constrained variations is obtained by using the relations (3.1) and computing the variations $\delta\mathbf{u}$ and δh induced by free variations $\delta\varphi$ vanishing at $t = 0$ and $t = T$. This principle yields the Euler–Poincaré equations in the general form

$$\partial_t \frac{\delta\ell}{\delta\mathbf{u}} + \mathsf{L}_{\mathbf{u}} \frac{\delta\ell}{\delta\mathbf{u}} = h\,\mathbf{d}\frac{\delta\ell}{\delta h}, \tag{3.4}$$

where the second term denotes the Lie derivative of the fluid momentum density (a one-form density) along the vector field $\mathbf{u}$, explicitly given by $\mathsf{L}_{\mathbf{u}}\alpha = \mathbf{i}_{\mathbf{u}}\mathbf{d}\alpha + \mathbf{d}\mathbf{i}_{\mathbf{u}}\alpha + \alpha \operatorname{div}\mathbf{u}$, with $\mathbf{d}$ being the exterior derivative and $\mathbf{i}_{\mathbf{u}}\omega$ the inner product of the k-form ω with the vector field $\mathbf{u}$. We refer to §3.8.1 for a review of exterior derivatives and Lie derivatives and to §3.8.3 for a detailed derivation of (3.4). The functional derivative $\frac{\delta\ell}{\delta\mathbf{u}}$, resp. $\frac{\delta\ell}{\delta h}$, is the one-form, resp., the function, defined by

$$\int_{\mathcal{S}} \frac{\delta\ell}{\delta\mathbf{u}} \cdot \mathbf{v}\, d\sigma = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \ell(\mathbf{u} + \varepsilon\mathbf{v}, h), \quad \text{resp.} \quad \int_{\mathcal{S}} \frac{\delta\ell}{\delta h} v\, d\sigma = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \ell(\mathbf{u}, h + \varepsilon v)$$

for any vector field $\mathbf{v}$ or function v. Equation (3.4) is accompanied with the

advection equation $\partial_t h + \mathrm{div}(h\mathbf{u}) = 0$ for the fluid depth which, in the Euler–Poincaré approach, follows from the second relation in (3.1).

For the RSW Lagrangian (3.2), we have

$$\frac{\delta \ell}{\delta \mathbf{u}} = h(\mathbf{u}^\flat + \mathbf{R}^\flat) \quad \text{and} \quad \frac{\delta \ell}{\delta h} = \frac{1}{2}\gamma(\mathbf{u},\mathbf{u}) + \gamma(\mathbf{R},\mathbf{u}) - g(h+B), \tag{3.5}$$

where $\mathbf{u}^\flat$ and $\mathbf{R}^\flat$ are the one-forms associated to the vector fields $\mathbf{u}$ and $\mathbf{R}$ via the flat operator $\flat$ of the Riemannian metric γ, see §3.8.1. Using these expressions, the Euler–Poincaré equations (3.4) lead to the momentum RSW equations, written in the space of one-forms as

$$h\,\partial_t \mathbf{u}^\flat + \mathbf{i}_{h\mathbf{u}}\mathbf{d}(\mathbf{u}^\flat + \mathbf{R}^\flat) = -h\,\mathbf{d}\Big(\frac{1}{2}\gamma(\mathbf{u},\mathbf{u}) + g(h+B)\Big), \tag{3.6}$$

see [8] and §3.8.3 for details. It is this expression of the RSW equations that appears in a discretized form in the variational discretization later in (3.23).

3.3. Discrete variational principle for the RSW equations

The variational discretization of the rotating shallow-water equations mimics the continuous variational method; in particular, each step of the continuous theory is translated to the discrete level. We provide here a review of the discrete Euler–Poincaré theory for the RSW, and refer the reader to [8] for full details.

3.3.1. Definition of the appropriate discrete configuration space for RSW

Discrete diffeomorphism group. The discretization procedure starts with the choice of a discrete version of the configuration group $\mathrm{Diff}(\mathcal{S})$ of a shallow-water

fluid. Following [65], [8], this group is obtained by first discretizing the space of functions $\mathcal{F}(\mathcal{S})$ on which $\mathrm{Diff}(\mathcal{S})$ acts by composition on the right, and then identifying a finite dimensional group acting by matrix multiplication on the finite dimensional space of discrete functions, while preserving some properties of the action by diffeomorphisms. In the compressible case, we shall only retain the property that constant functions are preserved under composition by a diffeomorphism. Given a mesh $\mathbb{M}$ of the sphere and choosing as discrete functions the space $\mathbb{R}^N$ of piecewise constant functions on $\mathbb{M}$, this results in the discrete diffeomorphism group

$$\mathsf{D}(\mathbb{M}) = \left\{ q \in \mathrm{GL}(N)^+ \mid q \cdot \mathbf{1} = \mathbf{1} \right\} \tag{3.7}$$

of dimension $N^2 - N$, where N is the number of cells in $\mathbb{M}$, the vector $\mathbf{1} \in \mathbb{R}^N$ is defined by $\mathbf{1} = (1, ..., 1)^\mathsf{T}$, and $\mathrm{GL}(N)^+ = \{q \in \mathrm{Mat}(N) \mid \det q > 0\}$ is the general linear group of orientation-preserving $N \times N$ matrices. The condition $q \cdot \mathbf{1} = \mathbf{1}$ encodes, at the discrete level, the fact that constant functions are preserved under composition by a diffeomorphism. The action of the group $\mathsf{D}(\mathbb{M})$ by matrix multiplication on discrete functions $F \in \mathbb{R}^N$ is denoted as

$$F \in \mathbb{R}^N \mapsto qF = F \circ q^{-1} \in \mathbb{R}^N, \quad q \in \mathsf{D}(\mathbb{M}),$$

where the suggestive notation $F \circ q^{-1}$ for the multiplication of the vector F by the matrix q is introduced to indicate that this action is understood as a discrete version of the action of $\mathrm{Diff}(\mathcal{S})$ by composition on the space $\mathcal{F}(\mathcal{S})$ of functions on $\mathcal{S}$, namely $f \in \mathcal{F}(\mathcal{S}) \mapsto f \circ \varphi^{-1} \in \mathcal{F}(\mathcal{S})$, see [65], [8] for details. The situation is formally illustrated by the diagram

$$\begin{array}{ccc} f \in \mathcal{F}(\mathcal{S}) & \xrightarrow{\mathrm{Diff}(\mathcal{S}} & f \circ \varphi^{-1} \in \mathcal{F}(\mathcal{S}) \\ \downarrow & & \downarrow \\ f \circ \varphi^{-1} \in \mathcal{F}(\mathcal{S}) & \xrightarrow{\mathsf{D}(\mathbb{M}} & F = F \circ q^{-1} \in \mathbb{R}^N \end{array}$$

The Lie algebra of the Lie group $\mathsf{D}(\mathbb{M})$ is the space of row-null $N \times N$ matrices

$$\mathfrak{d}(\mathbb{M}) = \{A \in \mathrm{Mat}(N) \mid A \cdot \mathbf{1} = 0\}, \tag{3.8}$$

endowed with the Lie bracket $[A, B] = AB - BA$. This Lie algebra is a discrete version of the Lie algebra of $\mathrm{Diff}(\mathcal{S})$ given by vector fields on $\mathcal{S}$. It is of particular interest for our derivations as it will allow us to formulate the discrete spatial Lagrangian required to derive the Euler–Poincaré equations from variational principles. By taking the derivative of continuous and discrete actions at the identity, we get $\frac{d}{dt}\Big|_{t=0} f \circ \varphi_t^{-1} = -\mathbf{d}f \cdot \mathbf{u}$ and $\frac{d}{dt}\Big|_{t=0} F \circ q_t^{-1} = AF$, where $\frac{d}{dt}\Big|_{t=0} \varphi_t = \mathbf{u}$ and $\frac{d}{dt}\Big|_{t=0} q_t = A$. This suggests that AF, with A an element of the Lie algebra $\mathfrak{d}(\mathbb{M})$, may play the role of a discrete version of the derivative of f in the direction of a continuous vector field $\mathbf{u}$. As we will recall below, not all $A \in \mathfrak{d}(\mathbb{M})$ can be interpreted as a discrete vector field. This induces nonholonomic constraints on the Lie algebra (3.8), which have to be appropriately taken into account in the variational principle.

Nonholonomic constraints. It can be shown that if a matrix $A \in \mathfrak{d}(\mathbb{M})$ approximates a vector field $\mathbf{u}$, then the matrix elements A_{ij} satisfy

$$\begin{aligned} A_{ij} &\simeq -\frac{1}{2\Omega_{ii}} \int_{D_{ij}} (\mathbf{u} \cdot \mathbf{n}_{ij}) dS, \quad \text{for all } j \in N(i),\ j \neq i, \\ A_{ii} &\simeq \frac{1}{2\Omega_{ii}} \int_{C_i} (\operatorname{div} \mathbf{u}) d\mathbf{x}, \end{aligned} \tag{3.9}$$

where $N(i)$ denotes the set of all indices (including i) of cells sharing a face with cell C_i, D_{ij} denotes the face common to cells C_i and C_j with unit normal $\mathbf{n}_{ij}$ pointing from C_i to C_j, and Ω_{ii} is the volume of cell C_i. We refer to [8] for the precise statement of these approximations. This identification imposes several constraints on the matrices in $\mathfrak{d}(\mathbb{M})$ to ensure that they represent a velocity vector field $\mathbf{u}$. First it is required that fluxes are nonzero only between neighboring cells,

hence we have the linear constraint

$$\mathcal{S} = \Big\{ A \in \mathfrak{d}(\mathbb{M}) \mid A_{ij} = 0, \ \text{ for all } j \notin N(i) \Big\}. \tag{3.10}$$

Second, we have the constraint $\Omega_{ii} A_{ij} = -\Omega_{jj} A_{ji}$, for all $j \neq i$, i.e., $A^\mathsf{T}\Omega + \Omega A$ is a diagonal matrix, with Ω being the $N \times N$ diagonal matrix with elements Ω_{ii}. This gives the additional linear constraint

$$\mathcal{R} = \Big\{ A \in \mathfrak{d}(\mathbb{M}) \mid A^\mathsf{T}\Omega + \Omega A \ \text{ is diagonal} \Big\}. \tag{3.11}$$

For the subsequent application of the variational principle, it is important to note that the constraints (3.10) and (3.11) are nonholonomic. Such constraints are taken into account by using the *Euler–Poincaré–d'Alembert principle*, which is the nonholonomic version of the Euler–Poincaré principle. As we recall in §3.8.3, the application of this principle requires the choice of an appropriate dual space and duality pairing.

Duality pairing and projector. We recall, see [65], that in the context of the discrete diffeomorphism group, the space of discrete one-forms $\Omega^1_d(\mathbb{M})$ is identified with the space of skew-symmetric $N \times N$ matrices. The discrete version of the L^2-pairing between discrete one-forms and discrete vector fields is given by

$$\langle L, A \rangle = \operatorname{Tr}(L^\mathsf{T}\Omega A), \quad \text{for} \quad A, L \in \operatorname{Mat}(N). \tag{3.12}$$

The application of this variational principle makes crucial use of Proposition 10 of [8] recalled below, which identifies a projector onto the dual space to the constraint $\mathcal{R}$ with respect to the pairing (3.12). The role of this proposition will become clear below when stating the Euler–Poincaré–d'Alembert principle.

Proposition 10 ([8]). *Given an $N \times N$ matrix L, we have the equivalence*

$$\langle L, A\rangle = 0, \;\; \textit{for all } A \in \mathcal{R} \quad \Leftrightarrow \quad \mathbf{P}(L) = 0,$$

where $\mathbf{P} : \mathrm{Mat}(N) \to \Omega^1_d(\mathbb{M})$ is the projector defined by $\mathbf{P}(L) := (L - \widehat{L})^{(A)}$, with $\widehat{L}_{ij} := L_{ii}$ and $L^{(A)} := \frac{1}{2}(L - L^{\mathsf{T}})$.

3.3.2. Euler–Poincaré–d'Alembert variational principle

We shall now reproduce, at the discrete level, the variational formulation for the RSW recalled earlier in Section 3.2, with the goal of obtaining the semi-discrete RSW equations via the Euler–Poincaré–d'Alembert principle.

As a first step, we need to identify the action of the group $\mathsf{D}(\mathbb{M})$ on the variables $D \in \mathbb{R}^N$ representing the discrete fluid depth. As in the continuous case, see the second equation in (3.1), this action, also denoted by $D \mapsto D \bullet q$, is dual to the action on discrete functions, namely

$$\langle D \bullet q, F\rangle_0 = \langle D, F \circ q^{-1}\rangle_0, \;\; \text{for all } F \in \mathbb{R}^N,\, q \in \mathsf{D}(\mathbb{M}), \tag{3.13}$$

with respect to the discrete L^2-pairing $\langle D, F\rangle_0 = D^{\mathsf{T}} \Omega F$. A direct application of (3.13) shows that this action and the associated Lie algebra action are given by the formulas

$$D \bullet q = \Omega^{-1} q^{\mathsf{T}} \Omega D \quad \text{and} \quad D \bullet A = \Omega^{-1} A^{\mathsf{T}} \Omega D, \tag{3.14}$$

for $q \in \mathsf{D}(\mathbb{M})$ and $A \in \mathfrak{d}(\mathbb{M})$.

Given a semi-discrete Lagrangian $\ell(A, D)$, the discrete version of the variational principle (3.3) reads as follows

$$\delta \int_0^T \ell(A, D) dt = 0, \tag{3.15}$$

with respect to variations $\delta A = \partial_t B + [B, A]$ and $\delta D = -D \bullet B$, with $A, B \in \mathcal{S} \cap \mathcal{R}$ and $B(0) = B(T) = 0$. As in the continuous case, the expression of the variations follows from the definition of A and D in terms of the Lagrangian variables q, $\dot{q}$, namely $A = \dot{q}q^{-1}$ and $D = D_0 \bullet q^{-1}$, which are discrete versions of (3.1). This principle, which incorporates the nonholonomic constraint $\mathcal{S} \cap \mathcal{R}$ in the Euler–Poincaré approach, is referred to as the Euler–Poincaré–d'Alembert principle. Note that both $A = \dot{q}q^{-1}$ and $B = \delta q q^{-1}$ have to satisfy the nonholonomic constraint $\mathcal{S} \cap \mathcal{R}$.

In the next theorem, this principle is applied to yield the general semi-discrete form of compressible fluid equations.

Theorem 11 (Discrete variational equations, [8])**.** *For a semi-discrete Lagrangian $\ell = \ell(A, D) : \mathfrak{d}(\mathbb{M}) \times \mathbb{R}^N \to \mathbb{R}$, the curves $A(t), D(t)$ are critical for the variational principle* (3.15) *if and only if they satisfy the equations*

$$\mathbf{P}\left(\frac{d}{dt}\frac{\delta \ell}{\delta A} + \Omega^{-1}\left[A^{\mathsf{T}}, \Omega \frac{\delta \ell}{\delta A} \right] + D\frac{\delta \ell}{\delta D}^{\mathsf{T}} \right)_{ij} = 0, \quad \text{for all } i \in N(j), \tag{3.16}$$

where $\mathbf{P} : \mathrm{Mat}(N) \to \Omega^1_d(\mathbb{M})$ is the projection obtained in Proposition 10. These equations are accompanied with the discrete continuity equation

$$\frac{d}{dt} D + D \bullet A = 0. \tag{3.17}$$

This result follows from a direct application of (3.15) by using the expression for δA and δD and isolating B which is an arbitrary curve in $\mathcal{S} \cap \mathcal{R}$. The equations then follow by an application of Proposition 10.

More precisely, using the definition of the pairings $\langle \, , \rangle$ and $\langle \, , \rangle_0$, we have

$$\delta \int_0^T \ell(A, D) dt = \int_0^T \left[\left\langle \frac{\delta \ell}{\delta A}, \partial_t B + [B, A] \right\rangle + \left\langle \frac{\delta \ell}{\delta D}, -\Omega^{-1} B^{\mathsf{T}} \Omega D \right\rangle_0 \right] dt \tag{3.18}$$

$$= -\int_0^T \left\langle \frac{d}{dt}\frac{\delta \ell}{\delta A} + \Omega^{-1}\left[A^{\mathsf{T}}, \Omega \frac{\delta \ell}{\delta A} \right] + D\frac{\delta \ell}{\delta D}^{\mathsf{T}}, B \right\rangle dt. \tag{3.19}$$

Since B is arbitrary in the space $\mathcal{S} \cap \mathcal{R}$, we obtain the equations (3.16) by application of Proposition 10. We provide in Table 3.1 a summary that enlightens the correspondence between the continuous and discrete objects.

Continuous diffeomorphisms	Discrete diffeomorphisms
$\mathrm{Diff}(M) \ni \varphi$	$\mathsf{D}(\mathbb{M}) \ni q$
Group action on functions	Group action on discrete functions
$f \mapsto f \circ \varphi$	$F \circ q = F \mapsto q^{-1}F$
Group action on densities	Group action on discrete densities
$h \mapsto h \bullet \varphi = (h \circ \varphi) J\varphi$	$D \mapsto D \bullet q = \Omega^{-1} q^{\mathsf{T}} \Omega D$
Eulerian velocity and depth	Eulerian discrete velocity and discrete depth
$\mathbf{u} = \dot{\varphi} \circ \varphi^{-1},\ \ h = (h_0 \circ \varphi^{-1}) J\varphi^{-1}$	$A = \dot{q} q^{-1},\ \ D = \Omega^{-1} q^{-\mathsf{T}} \Omega D_0$
Euler-Poincaré principle	Euler-Poincaré-d'Alembert principle
$\delta \int_0^T \ell(\mathbf{u}, h) dt = 0,$ $\delta\mathbf{u} = \partial_t \mathbf{v} + [\mathbf{v}, \mathbf{u}],\ \ \delta h = -\operatorname{div}(h\mathbf{v})$	$\delta \int_0^T \ell(A, D) dt = 0,$ $\delta A = \partial_t B + [B, A], \qquad \delta D = -\Omega^{-1} B^{\mathsf{T}} \Omega D$ constraint: $A, B \in \mathcal{S} \cap \mathcal{R}$

Table 3.1.: Continuous and discrete objects

For the RSW case, the discrete Lagrangian is

$$\ell(A, D) = \frac{1}{2} \sum_{i,j=1}^{N} D_i A_{ij}^{\flat} A_{ij} \Omega_{ii} + \sum_{i,j=1}^{N} D_i R_{ij}^{\flat} A_{ij} \Omega_{ii} - \frac{1}{2} \sum_{i=1}^{N} g (D_i + B_i)^2 \Omega_{ii}, \quad (3.20)$$

see (3.2), which requires the construction of a discrete "flat" operator $A \in \mathcal{S} \cap \mathcal{R} \mapsto A^{\flat} \in \Omega_d^1(\mathbb{M})$ associated to a given mesh, see [65].

The abstract developments made so far are valid for any kind of reasonable non-degenerate mesh (i.e. the mesh admits a non-degenerate circumcenter dual and is member of a shape-regular and quasi-uniform mesh family, cf. [8]). By choosing a fixed mesh, we will be able to express these abstract notions in concrete (implementable) equations.

3.3.3. Semi-discrete scheme on a 2D simplicial mesh

For our implementation, we use an icosahedral grid as pictured in Fig. 3.1. The construction of the grid is described in [44]. The icosahedron's edges are recursively

bisected, and the new vertices are projected onto the unit sphere. The triangles are used as the primal grid, and the circumcenter dual, consisting of pentagonal and hexagonal cells, as the dual grid. The vertex positions are optimized in the sense that the global maximum of the discrepancy between primal edge midpoints and intersection points of primal and dual edges is minimized [45]. Note that, for our variational integrator, any other optimization for which primal and dual edges intersect perpendicularly could be used too; only the order of convergence of the differential operators will be affected by the optimization.

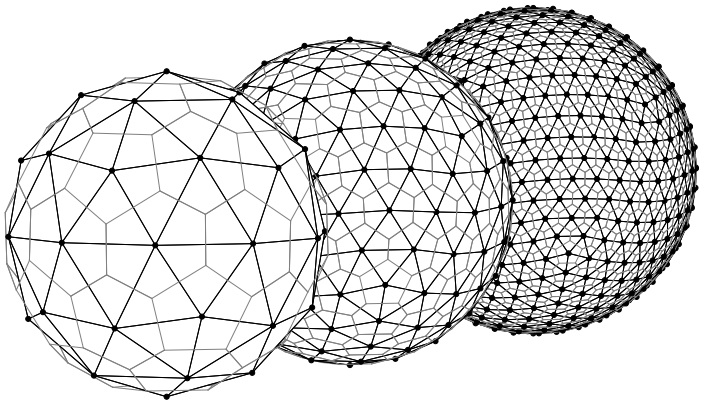


Figure 3.1.: Icosahedral meshes with refinement levels 2, 3, and 4 corresponding to 80, 320, and 1280 triangular cells.

In [84], the grid and its connectivity is described in more details. Fig. 3.2 shows a section of the simplicial mesh where we indicate our notation:

f_{ij} := length of a primal edge, triangle edge located between triangle i and triangle j;

h_{ij} := length of a dual edge that connects the circumcenters of triangle i and triangle j;

Ω_{ii} := area of a primal simplex (triangle) T_i;

$K_i^{\pm}$:= proportional area of the intersection of (triangle) T_i and (hexagon/pentagon) $\zeta_{\pm}$, see Equation (3.22).

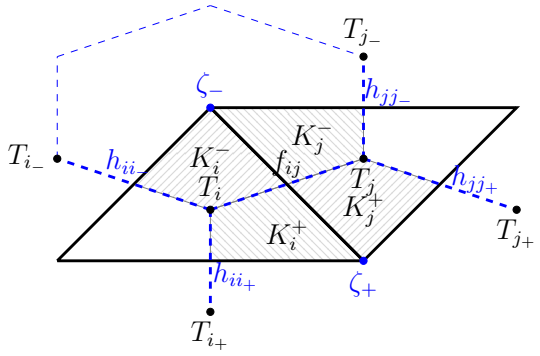


Figure 3.2.: Notation and indexing conventions for the 2D simplicial mesh.

Using the flat operator on a 2D simplicial mesh, see [65], we are able to represent the discrete one-forms $A^\flat \in \Omega_d^1(\mathbb{M})$ in terms of the discrete vector fields $A \in \mathcal{S} \cap \mathcal{R}$ by

$$\begin{aligned} A_{ij}^\flat &= 2\Omega_{ii}\frac{h_{ij}}{f_{ij}}A_{ij}, \quad \text{for } j \in N(i), \\ A_{ij}^\flat + A_{jk}^\flat + A_{ki}^\flat &= K_j^e\Big\langle \omega(A^\flat), \zeta_e \Big\rangle, \quad \text{for } i,k \in N(j),\ k \notin N(i), \end{aligned} \tag{3.21}$$

for the constant K_i^e and the vorticity $\omega(A^\flat)$, defined as, respectively

$$K_k^e := \frac{|\zeta_e \cap T_k|}{|\zeta_e|} \quad \text{and} \quad \Big\langle \omega(A^\flat), \zeta_e \Big\rangle := \sum_{h_{mn} \in \partial\zeta_e} A_{mn}^\flat \,. \tag{3.22}$$

Here, e denotes the node common to triangles T_i, T_j, T_k and $|\zeta_e \cap T_k|$ is the area of the intersection of T_k and ζ_e, where the latter denotes the dual cell to e, see Fig. 3.2. The discrete vorticity, see the second equation in (3.22), is calculated by taking the sum over the dual edges in the boundary $\partial\zeta_e$ counterclockwise around node e. The definition of $A^\flat$ in (3.21) leads to a skew-symmetric matrix, hence $A^\flat \in \Omega_d^1(\mathbb{M})$.

As shown in [8], for the discrete Lagrangian (3.20) on a simplicial grid, the

discrete variational equations (3.16) lead to the following semi-discrete equations

$$\begin{cases} \overline{D}_{ij}\dfrac{d}{dt}A^\flat_{ij} + \omega_+\left(K_i^+\overline{D}_{ji_+}A_{ii_+} + K_j^+\overline{D}_{ij_+}A_{jj_+}\right) - \omega_-\left(K_i^-\overline{D}_{ji_-}A_{ii_-} + K_j^-\overline{D}_{ij_-}A_{jj_-}\right) \\ \qquad +\overline{D}_{ij}\dfrac{1}{2}\left(A^\flat_{ii_-}A_{ii_-} + A^\flat_{ii_+}A_{ii_+} + A^\flat_{ij}A_{ij} - A^\flat_{ji}A_{ji} - A^\flat_{jj_-}A_{jj_-} - A^\flat_{jj_+}A_{jj_+}\right) \\ \qquad +\overline{D}_{ij}\left(\dfrac{\partial\epsilon}{\partial D_i} - \dfrac{\partial\epsilon}{\partial D_j}\right) = 0 \\ \dfrac{d}{dt}D_i = A_{ii_-}D_{i_-} + A_{ii_+}D_{i_+} + A_{ij}D_j - A_{ii}D_i, \end{cases} \tag{3.23}$$

in which $\overline{D}_{ij} := \frac{1}{2}(D_i + D_j)$ denotes the average of the cell values and $\omega_\pm := \sum_{h_{mn}\in\partial\zeta_\pm}(A^\flat_{mn} + R^\flat_{mn})$ is the discrete absolute vorticity at the nodes $\pm$ at endpoints of the edge between cells i and j, see Fig. 3.2. This is the discrete version of the RSW equations (3.6).

3.3.4. Semi-discrete RSW scheme in terms of the discrete velocity field

With a suitable choice of structure preserving time discretization, Equations (3.23) provide a set of fully discrete equations which can be implemented as they stand. However, as it is more familiar in the GFD community to work with velocity quantities, we proceed in rewriting the equations correspondingly.

From the original definitions (3.9) and the flat operator (3.21), we find the following relation between one-forms $A^\flat$, Lie algebra elements A and the normal velocity degrees of freedom V_{ij} on the triangle edges' midpoints:

$$\begin{aligned} A_{ij} &= -\frac{1}{2\Omega_{ii}}f_{ij}V_{ij}, \quad \text{for all } j\in N(i),\ j\neq i, \\ A_{ii} &= -A_{ij} - A_{ii_-} - A_{ii_+} = \sum_{k=j,i_-,i_+}\frac{1}{2\Omega_{ii}}f_{ik}V_{ik} =: \frac{1}{2}\operatorname{div}(V)_i. \end{aligned} \tag{3.24}$$

while $A^\flat_{ij} = -h_{ij}V_{ij}$ for $i\in N(j)$, see Fig. 3.2 for the index notations. Note that $\operatorname{div}(V)$ coincides with the natural finite volume divergence operator on a triangular

mesh (see e.g. [7]).

Momentum equation. The semi-discrete momentum equation in matrix–vector notation reads

$$\partial_t V_{ij} + \mathrm{Adv}(V, D)_{ij} = \mathrm{K}(V)_{ij} - \mathrm{G}(D)_{ij}, \tag{3.25}$$

where we define

$$\begin{aligned}
&\mathrm{Adv}(V, D)_{ij} := \\
&-\frac{1}{\overline{D}_{ij} h_{ij}} \Big(\frac{1}{|\zeta_-|} \sum_{h_{mn} \in \partial\zeta_-} h_{mn} (V_{mn} + \bar{R}_{mn}) \Big) \\
&\left(\frac{|\zeta_- \cap T_i|}{2\Omega_{ii}} \overline{D}_{ji_-} f_{ii_-} V_{ii_-} + \frac{|\zeta_- \cap T_j|}{2\Omega_{jj}} \overline{D}_{ij_-} f_{jj_-} V_{jj_-} \right) \\
&+ \frac{1}{\overline{D}_{ij} h_{ij}} \Big(\frac{1}{|\zeta_+|} \sum_{h_{mn} \in \partial\zeta_+} h_{mn} (V_{mn} + \bar{R}_{mn}) \Big) \\
&\left(\frac{|\zeta_+ \cap T_i|}{2\Omega_{ii}} \overline{D}_{ji_+} f_{ii_+} V_{ii_+} + \frac{|\zeta_+ \cap T_j|}{2\Omega_{jj}} \overline{D}_{ij_+} f_{jj_+} V_{jj_+} \right), \\
&\mathrm{K}(V)_{ij} := -\frac{1}{2h_{ij}} \Big(\frac{h_{jj_-} f_{jj_-} (V_{jj_-})^2}{2\Omega_{jj}} + \frac{h_{jj_+} f_{jj_+} (V_{jj_+})^2}{2\Omega_{jj}} + \frac{h_{ij} f_{ij} (V_{ji})^2}{2\Omega_{jj}} \\
&\qquad - \frac{h_{ii_-} f_{ii_-} (V_{ii_-})^2}{2\Omega_{ii}} - \frac{h_{ii_+} f_{ii_+} (V_{ii_+})^2}{2\Omega_{ii}} - \frac{h_{ij} f_{ij} (V_{ij})^2}{2\Omega_{ii}} \Big), \\
&\mathrm{G}(D)_{ij} := \frac{g}{h_{ij}} \Big(D_j + B_j - (D_i + B_i) \Big),
\end{aligned}$$

for values $\bar{R}_{mn}$ related to R_{mn} by $R_{ij} = -\frac{1}{2\Omega_{ii}} f_{ij} \bar{R}_{ij}$, analogously to the relation between V_{mn} and A_{mn}. We again use the definition $\overline{D}_{ij} = \frac{D_i + D_j}{2}$. We define as Coriolis parameter

$$f|_{\zeta_\pm} := \frac{1}{|\zeta_\pm|} \sum_{h_{mn} \in \partial\zeta_\pm} h_{mn} \bar{R}_{mn}. \tag{3.26}$$

Continuity equation. The semi-discrete continuity equations in matrix–vector form is given by

$$\partial_t D_i + \underbrace{\frac{1}{\Omega_{ii}} f_{ij} V_{ij} \overline{D}_{ij} + \frac{1}{\Omega_{ii}} f_{ii_-} V_{ii_-} \overline{D}_{ii_-} + \frac{1}{\Omega_{ii}} f_{ii_+} V_{ii_+} \overline{D}_{ii_+}}_{:=\mathrm{div}(V,D)_i} = 0. \tag{3.27}$$

Hence, the spatial variational discretization process leads to a standard finite volume representation of the divergence operator and hence of the continuity equation.

Time discretizations. Since the spatial discretization has been realized by variational principles in a structure-preserving way, a temporal variational discretization can be implemented by following the discrete (in time) Euler–Poincaré–d'Alembert approach, analogously to what has been done in [36] and [26], to which we refer for a detailed treatment. This variational approach is based on the introduction of a local approximant to the exponential map of the Lie group, see [18], chosen here as the Cayley transform. As explained in [26, 36], by dropping cubic terms, this results in a Crank–Nicolson-type time update for the momentum equation (3.16) (which, as such, is second order in time) and an update equation based on the Cayley transform for the advection equation (3.17). Following [8], we will use below the Crank–Nicolson-type time update directly on the momentum equation as reformulated in (3.25). This considerably simplifies the solution procedure without altering the behavior of the scheme.

Alternatively, we apply also a standard time integrator using a Crank–Nicolson-type time update for the continuity equations instead of the Cayley transform and compare both time stepping schemes with each other while keeping the spatial variational discretization unmodified.

1.) Fluid depth equation update by Cayley transform: This time integrator consists of two steps. We first compute the update equation for the fluid depth D,

which is based on the Cayley transform τ. This update equation is then given by $D^{t+1} = \tau(\Delta t A^t) D^t$ for the time t and a time step size Δt. In particular, τ can be represented as

$$\Big(I - \frac{1}{2}\Delta t A^t\Big) D^{t+1} = \Big(I + \frac{1}{2}\Delta t A^t\Big) D^t, \tag{3.28}$$

with I the identity matrix (cf. [26] for more details). Note that A can be expressed in terms of V using (3.24). In a second step, we solve the momentum equation, given by an implicit nonlinear equation (step 2), according to the fixed-point iteration:

1. Start loop over $k = 0$ with initial guess at t: $V^*_{k=0} = V^t$;

2. Calculate updated velocity V^*_{k+1} from the explicit equation:

$$\frac{V^*_{k+1} - V^t}{\Delta t} = -\frac{\mathrm{Adv}(V^*_k, D^{t+1}) + \mathrm{Adv}(V^t, D^t)}{2} + \frac{\mathrm{K}(V^*_k) + \mathrm{K}(V^t)}{2} - \mathrm{G}(D^{t+1});$$

3. Stop loop over k if $||V^*_{k+1} - V^*_k|| < \epsilon$ for a small positive ϵ, take $V^{t+1} = V^*_{k+1}$.

For more details, we refer the reader to [8]. Note that for this time integration scheme, we do not discretize the continuity equation (3.27) directly, but use the discretization of τ in (3.28).

Remark 12. The temporal discretization using the Cayley transform is only an approximation to a fully variational time integrator, see [26, 36] for more details. The energy of the resulting temporal scheme is not conserved at machine precision, but the energy fluctuates around a long term mean. We observe this behaviour when measuring the energy error for the numerical simulations.

2.) Fluid depth equation update by Crank–Nicolson: Here, we use a two-step time integration scheme to solve the system of fully discretized nonlinear momentum and continuity equations:

$$\frac{V_{ij}^{t+1} - V_{ij}^{t}}{\Delta t} = -\frac{\mathrm{Adv}(V^{t+1}, D^{t+1})_{ij} + \mathrm{Adv}(V^{t}, D^{t})_{ij}}{2} \tag{3.29}$$

$$+ \frac{\mathrm{K}(V^{t+1})_{ij} + \mathrm{K}(V^{t})_{ij}}{2} - \mathrm{G}(D^{t+1})_{ij}\,, \tag{3.30}$$

$$\frac{D_{i}^{t+1} - D_{i}^{t}}{\Delta t} = -\frac{\mathrm{div}(V^{t+1}, D^{t+1})_{i} + \mathrm{div}(V^{t}, D^{t})_{i}}{2}\,. \tag{3.31}$$

We solve this system of nonlinear equations by fixed-point iteration for all edges ij and cells i. To enhance readability, we skip the corresponding subindices in the following. The solution algorithm reads:

1. Start loop over $k = 0$ with initial guess at t: $V_{k=0}^{*} = V^{t}$ and $D_{k=0}^{*} = D^{t}$;

2. Calculate updated water depth (density) D_{k+1}^{*} from the explicit equation:

$$\frac{D_{k+1}^{*} - D^{t}}{\Delta t} = -\frac{\mathrm{div}(V_{k}^{*}, D_{k}^{*}) + \mathrm{div}(V^{t}, D^{t})}{2}$$

3. Calculate updated velocity V_{k+1}^{*} from the explicit equation:

$$\frac{V_{k+1}^{*} - V^{t}}{\Delta t} = -\frac{\mathrm{Adv}(V_{k}^{*}, D_{k+1}^{*}) + \mathrm{Adv}(V^{t}, D^{t})}{2} + \frac{\mathrm{K}(V_{k}^{*}) + \mathrm{K}(V^{t})}{2} - \mathrm{G}(D_{k+1}^{*});$$

4. Stop loop over k if $||V_{k+1}^{*} - V_{k}^{*}|| + ||D_{k+1}^{*} - D_{k}^{*}|| < \epsilon$ for a small positive ϵ.

Note that in case of convergence, i.e. $V_{k+1}^{*} \to V^{t+1}$ and $D_{k+1}^{*} \to D^{t+1}$, this algorithm solves equations (3.30) and (3.31).

3.4. Numerical analysis of the differential operators

We present a convergence study of the gradient, divergence and curl operators on the icosahedral meshes. These operators are used directly and indirectly in our scheme. The study will be done in x, y, z coordinates on $\mathbb{R}^3$, such that it is inde-

pendent of local coordinates of the hypersurface on which we solve the equations. We consider the Euclidean metric $\langle\,,\rangle$ on $\mathbb{R}^3$ and denote by ∇ the gradient relative to it.

Let $\mathbf{N}_\mathbf{x}$ be the outward unit normal vector on the sphere at $\mathbf{x} \in \mathcal{S}$ and $P_\mathbf{x} = I - \mathbf{N}_\mathbf{x}\mathbf{N}_\mathbf{x}^\top$ be the orthogonal projection onto the tangent space of $\mathcal{S}$ at $\mathbf{x}$, where I is the 3×3 identity matrix. In the following, we assume that $g\colon \mathcal{S} \to \mathbb{R}$ is a given real-valued function on $\mathcal{S}$, with $\tilde{g}\colon \mathbb{R}^3 \to \mathbb{R}$ being an arbitrary extension of g to $\mathbb{R}^3$, and $\mathbf{u}$ is a given vector field on $\mathcal{S}$, i.e. $\mathbf{u}\colon \mathcal{S} \to T\mathcal{S}$, with $\tilde{\mathbf{u}}\colon \mathbb{R}^3 \to \mathbb{R}^3$ being an arbitrary extension of $\mathbf{u}$ to $\mathbb{R}^3$. From the general expression of the covariant derivative induced on hypersurfaces, see e.g., [54], the gradient, divergence and curl operators on $\mathcal{S}$ relative to γ can be written as

$$
\begin{aligned}
\operatorname{grad} g(\mathbf{x}) &:= \nabla\tilde{g}(\mathbf{x}) - (\nabla\tilde{g}(\mathbf{x}) \cdot \mathbf{N}_\mathbf{x})\mathbf{N}_\mathbf{x} = (I - \mathbf{N}_\mathbf{x}\mathbf{N}_\mathbf{x}^\top)\nabla\tilde{g}(\mathbf{x}) = P_\mathbf{x}\nabla\tilde{g}(\mathbf{x}) \\
\operatorname{div} \mathbf{u}(\mathbf{x}) &:= \langle P_\mathbf{x}\nabla, \tilde{\mathbf{u}}(\mathbf{x})\rangle \\
\operatorname{curl} \mathbf{u}(\mathbf{x}) &:= \langle (P_\mathbf{x}\nabla) \times \tilde{\mathbf{u}}), \mathbf{N}_\mathbf{x}\rangle,
\end{aligned}
$$

see also [28, 29].

Note that in contrast to the numerical schemes derived in [28, 29] we use the above analytical expressions only to initialize the initial conditions for the numerical test cases presented in Section 3.5, and for comparing against the numerically obtained expressions for grad, div and curl.

As explained further below, in our variational scheme (3.25)–(3.27), the dis-

cretization of these operators are given as follows

$$(\mathrm{Grad}^{\mathrm{num}}\, g)_{ij} = \frac{g_j - g_i}{h_{ij}}, \tag{3.32}$$

$$(\overset{\mathrm{num}}{\mathrm{div}}\, \mathbf{u})_i = \frac{1}{\Omega_{ii}} \sum_{\ell \in \{j, i_-, i_+\}} f_{i\ell} V_{i\ell}, \tag{3.33}$$

$$(\overset{\mathrm{num}}{\mathrm{curl}}\, \mathbf{u})_{\zeta_e} = \frac{1}{|\zeta_e|} \sum_{mn \in \partial\zeta_e} h_{mn} V_{mn}, \tag{3.34}$$

where, from (3.9) and (3.24), we have $V_{i\ell} = \frac{1}{f_{i\ell}} \int_{D_{i\ell}} (\mathbf{u} \cdot \mathbf{n}_{i\ell}) dS$ and where $g_i = \frac{1}{\Omega_{ii}} \int_{T_i} g dx$. As before, $\mathbf{n}_{ij}$ is the unit normal pointing from T_i to T_j, ζ_e is the cell dual to a node e (a hexagon or pentagon), $|\zeta_e|$ is its area. Note that in (3.33) the sum is over the cells adjacent to T_i, and in (3.34) the sum is over the dual edges in the boundary $\partial\zeta_e$ counterclockwise around node e.

The gradient (3.32) appears in (3.25) and is denoted as $G(D)$. Therein, in the advection term, $\mathrm{Adv}(V, D)$, we find the curl operator (3.34) consisting of a counterclockwise sum over the edges of a dual cell. The divergence operator (3.33) with positive fluxes when pointing out of the triangles appears in the continuity equation (3.27).

For our convergence study, we use $g(\mathbf{x}) = \sin(x) + \sin(2y) + \sin(2z)$ for the gradient, $\mathbf{u}(\mathbf{x}) = (x - x^3, -x^2 y, -x^2 z)^\top$ for the divergence and $\mathbf{u}(\mathbf{x}) = (z, 0, -x)^\top$ for the curl.

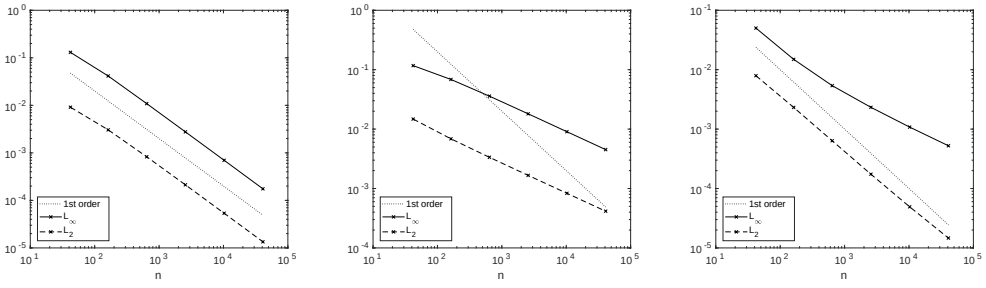


Figure 3.3.: Convergence table for the discrete gradient (left), divergence (center), and curl (right).

In Fig. 3.3, we see that the gradient and curl operator converge with first order.

The divergence operator converges too, but does not achieve first order when evaluated on the triangles, because the grid is optimized for the hexagons [45]. In Fig. 3.4, we clearly see the grid imprint of the original icosahedron in the pointwise errors for the divergence, showing that the error is highest along the edges of the icosahedron.

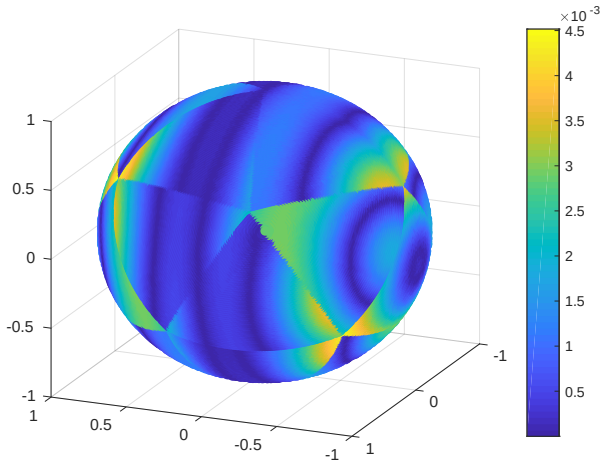


Figure 3.4.: The grid imprint of the absolute error for the evaluation of the divergence on triangles.

3.5. Numerical simulations

We consider four test cases; (1) the lake-at-rest solution to demonstrate that the model is well-balanced; (2) a global steady-state solution to study the convergence and energy and enstrophy loss of the model; (3) the flow over an isolated mountain; and (4) the Rossby–Haurwitz wave solution.

The following constants are kept fixed for all simulations:

$$R = 6.37122 \cdot 10^6 \quad [m], \qquad \Omega = 7.292 \cdot 10^{-5} \quad [s^{-1}], \qquad g = 9.80616 \quad [ms^{-2}].$$

The Coriolis parameter is $f = 2\Omega \sin\theta$. Unless indicated otherwise, the simula-

tions are carried out on a grid with $N = 40962$ Voronoi cells (corresponding to a resolution of about 120 km) for which we chose a time step of $\Delta t = 100$ s. To estimate the numerical errors, we use the following definitions for the relative L_∞-error and L_2-error,

$$\|u - u_0\|_\infty = \frac{\max_i |u(i) - u_0(i)|}{\max_i |u_0(i)|} \qquad \|u - u_0\|_2 = \frac{\sqrt{\sum_i (\Omega_{ii}(u(i) - u_0(i)))^2}}{\sqrt{\sum_i \Omega_{ii} u_0(i)^2}},$$

where Ω_{ii} is the area associated with cell i. Here, the function $u(i)$ is the numerical solution defined at $\mathbf{x}_i$ or the magnitude of the numerical solution at $\mathbf{x}_i$ (when used for calculating the error for the velocity) and $u_0(i)$ is the initial function at $\mathbf{x}_i$.

Since our scheme preserves mass and potential circulation up to machine precision, those error norms are not presented. In this section, we only present results obtained by using the Cayley transform time discretization, which we compare in Section 3.6 with the standard time integrator.

3.5.1. Case 1: Lake at rest

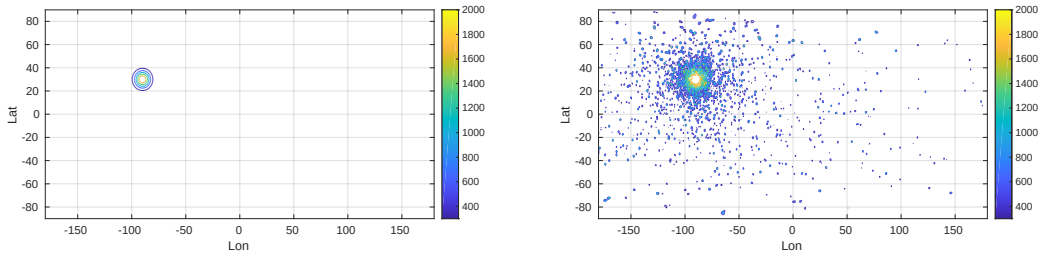
This test case verifies that the model is *well-balanced*, that is, the exact solution $\mathbf{u} = 0$, $h + B = constant$ of the RSW equations is preserved up to machine precision. Here, we choose the test case of a resting fluid over a conical shaped mountain. The initial velocity is $\mathbf{u} \equiv 0$ and the bottom topography is defined by

$$B(\lambda, \theta) = 2000 \exp\left(-(2.8 \cdot 9r/\pi)^2\right) \text{ with } r^2 = \min\left((\pi/9)^2, (\lambda - \lambda_c)^2 + (\theta - \theta_c)^2\right),$$

see also Fig. 3.5a. The total water depth is $D = 5960 - B$. In addition, we carry out a second simulation where we add some white noise to the bottom topography

profile, see Fig. 3.5b, to verify that the model remains well-balanced also for a noisy bottom topography.

The shallow-water equations are integrated over 15 days. The initial conditions are preserved to the order of machine precision, which verifies that the model is indeed well-balanced.



(a) Profile in [m] of the smooth bottom topography case.

(b) Profile in [m] of the noisy bottom topography case.

Figure 3.5.: Topography for smooth and noisy bottom topography.

3.5.2. Case 2: Global steady-state nonlinear zonal geostrophic flow

This test case, originally proposed in [90], is a geostrophically balanced flow over a flat bottom topography, i.e. $B \equiv 0$ $[m]$. This flow represents an exact solution to the rotating shallow-water equations. The initial conditions are:

$$V_{ij} = V(x_{ij}, y_{ij}, z_{ij}) = u_0(-y_{ij}, x_{ij}, 0)^{\mathsf{T}} \cdot \mathbf{n}_{ij}, \text{ where } u_0 = \frac{2\pi R}{12 \cdot 86400} \; [s^{-1}]$$

$$D_i = D(x_{T_i}, y_{T_i}, z_{T_i}) = h_0 - \frac{1}{g}\left(R\Omega u_0 + u_0^2/2\right) z_{T_i}^2, \text{ where } gh_0 = 2.94 \cdot 10^4 \; [m^2 s^{-2}].$$

Here $\mathbf{n}_{ij}$ denotes the normal vector of a triangle edge, and V_{ij} is the directional magnitude of the velocity normal to an edge f_{ij}, see Equation (3.24).

Although the nonlinear zonal geostrophic flow is a steady state solution of the RSW (in which any quantity is conserved because of no time dependence), it is

only a stationary solution of a numerical RSW scheme up to numerical errors. As such, it is also interesting to monitor the time series of the numerical values of the conserved quantities of the RSW, and to verify that the energy error converges at the expected first order.

Fig. 3.6 shows the initial conditions. In Fig. 3.7, we display the time evolution of the errors for energy, potential enstrophy, height and velocity. It can be seen that there is no trend in the evolution of the error. The energy is well-preserved at the order of 10^{-8} and the potential enstrophy at order of 10^{-7}. Fig. 3.8a contains the results of the spatial convergence study over 12 days with different resolutions. We integrate over 12 days, as one rotation of the fluid flow around the globe takes precisely 12 days. It can be seen that D and V do not achieve first order convergence, which is natural since the divergence operator likewise does not achieve first order convergence, see Fig. 3.4. Fig. 3.8b shows the expected first order convergence of the energy error with respect to the time step.

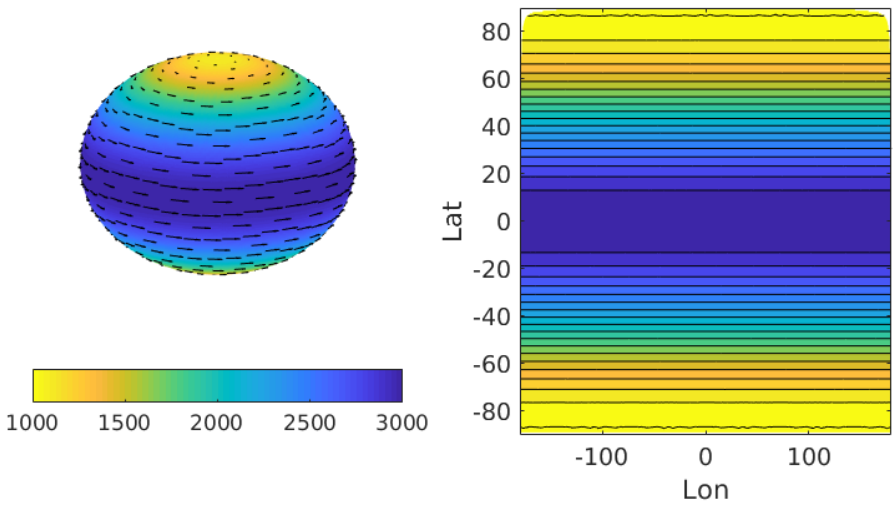


Figure 3.6.: Initial conditions for the nonlinear zonal geostrophic flow test case.

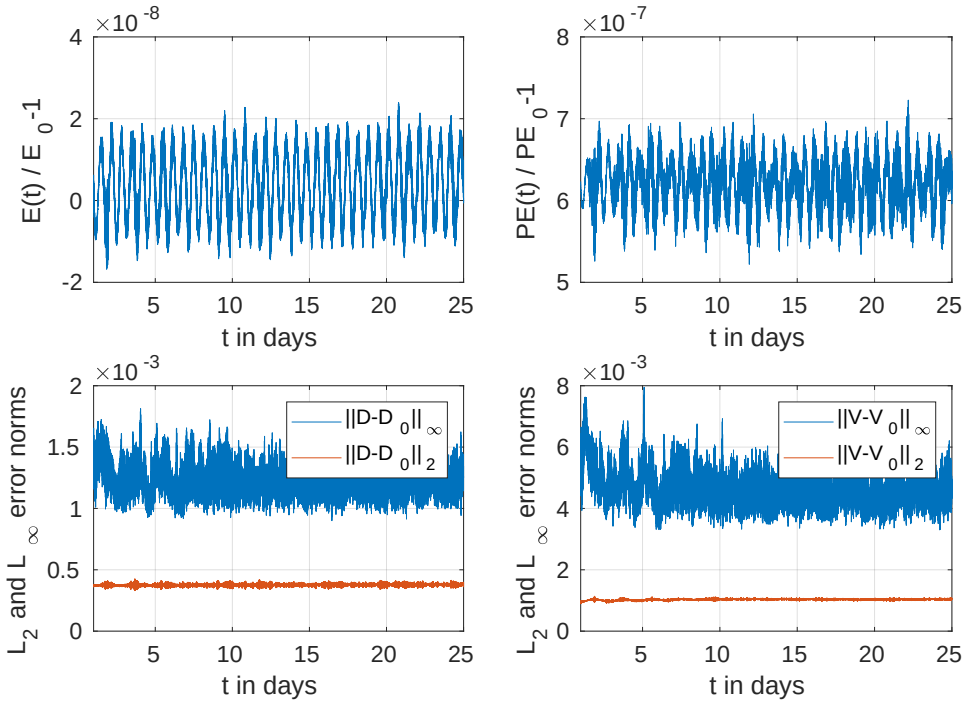
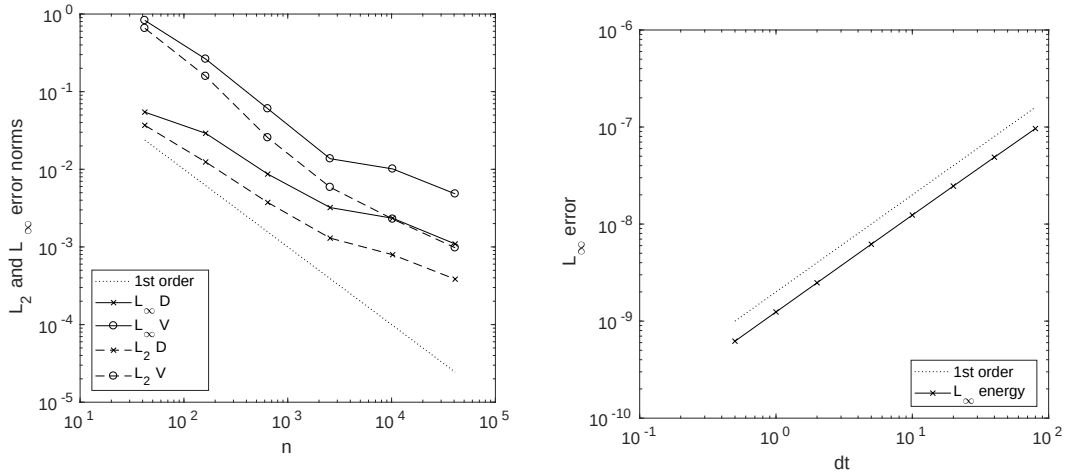


Figure 3.7.: Time series for the error norms for the conserved quantities for the nonlinear zonal geostrophic flow test case.



(a) Convergence plot of D and V for the nonlinear zonal geostrophic flow test case. Norms are taken at day 12.

(b) Convergence plot of energy error for $n = 10242$.

Figure 3.8.: Convergence plots for nonlinear zonal geostrophic flow test case

3.5.3. Case 3: Flow over an isolated mountain

Here, we consider the flow over a conically-shaped mountain which was also proposed in [90]. The initial conditions, see Fig. 3.9a, of this test case are the same as

for Case 2 considered in the previous subsection, except that now $h_0 = 5960$ $[m]$ and $u_0 = 20$ $[ms^{-1}]$. The following bottom topography is imposed,

$$B(\lambda, \theta) = 2000\,(1 - 9r/\pi) \quad \text{with} \quad r^2 = \min\Big((\pi/9)^2, (\lambda - \lambda_c)^2 + (\theta - \theta_c)^2\Big).$$

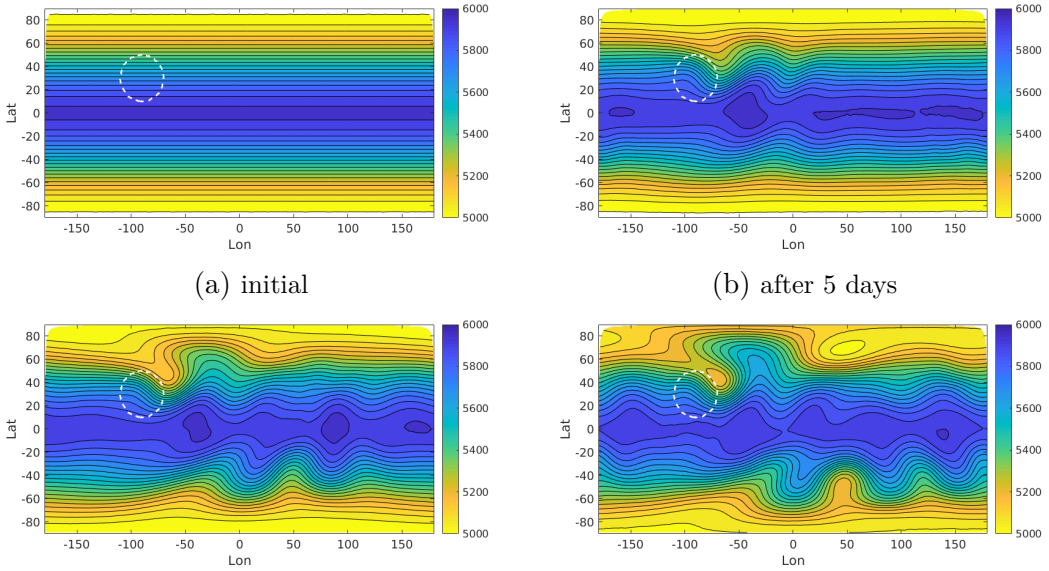
The mountain is centered at $\lambda_c = 3\pi/2$ and $\theta_c = \pi/6$. Note that there is no analytical solution for this problem.

Fig. 3.9 shows snapshots of the height field at times $t = 0$ (a), $t = 5$ days (b), $t = 10$ days (c) and $t = 15$ days (d), which are the times suggested in [90] to show the computed solutions. These results are visually similar to the results from different models, such as those given in [29, 85]. The time series for the errors in the energy and potential enstrophy are depicted in Figure 3.10.

Note that the energy is almost as well conserved as in Case 2, whereas potential enstrophy conservation is three orders of magnitudes less accurate (cf. Fig. 3.7 and Fig. 3.10). This is because the variational integrator preserves energy by construction, but we have no control over the conservation behavior of potential enstrophy (only PV is conserved). This, however, should not provide a problem in practice because usually enstrophy has to be dissipated anyways at small scales, cf. [60]. Nevertheless, our schemes preserves potential enstrophy relatively well given the nonlinearity of the test case.

3.5.4. Case 4: Rossby–Haurwitz waves

We consider a Rossby–Haurwitz wave with wavenumber $\kappa = 4$, which is proposed in [90]. Unlike the non-divergent barotropic vorticity equation, the shallow-water equations can only approximate this solution. For comparison, snapshots after 7 and 14 days are presented. For completeness we present the initial conditions here in latitude (θ) and longitude (λ),



(a) initial (b) after 5 days

(c) after 10 days (d) after 15 days

Figure 3.9.: Numerical solution for flow over an isolated mountain. Contour interval is 50m.

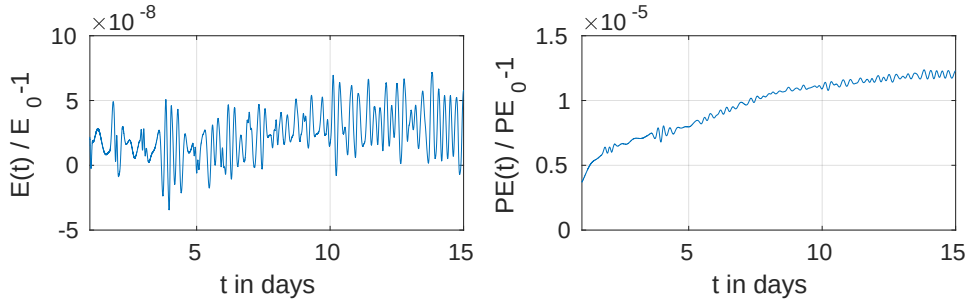


Figure 3.10.: Time series of the errors of energy and potential enstrophy for flow over an isolated mountain over the 15 days integration period.

$$K = 7.848 \cdot 10^{-6} \quad [s^{-1}], \qquad \kappa = 4 \quad \text{(wave number)}, \qquad h_0 = 8 \cdot 10^3 \quad [m].$$

The components of the velocity vector $\mathbf{u} = (u, v)$ are

$$u = RK\cos\theta + RK\cos^{\kappa-1}\theta(\kappa\sin^2\theta - \cos^2\theta)\cos\kappa\lambda,$$

$$v = -RK\kappa\cos^{\kappa-1}\theta\sin\theta\sin\kappa\lambda$$

and the water elevation is given by

$$D_i = h_0 + \tfrac{1}{g}(R^2 A(\theta_{T_i}) + R^2 B(\theta_{T_i}) \cos \kappa\lambda_{T_i} + R^2 C(\theta_{T_i}) \cos 2\kappa\lambda_{T_i}),$$

where

$$\begin{aligned}
A(\theta) &= \tfrac{K}{2}(2\Omega + K)\cos^2\theta + \tfrac{1}{4}K^2 \cos^{2\kappa}\theta\Big((\kappa+1)\cos^2\theta + (2\kappa^2 - \kappa - 2) - 2\kappa^2\cos^{-2}\theta\Big),\\
B(\theta) &= \frac{2(\Omega+K)K}{(\kappa+1)(\kappa+2)}\cos^{\kappa}\theta\Big((\kappa^2+2\kappa+2) - (\kappa+1)^2\cos^2\theta\Big),\\
C(\theta) &= \tfrac{1}{4}K^2\cos^{2\kappa}\theta\Big((\kappa+1)\cos^2\theta - (\kappa+2)\Big).
\end{aligned}$$

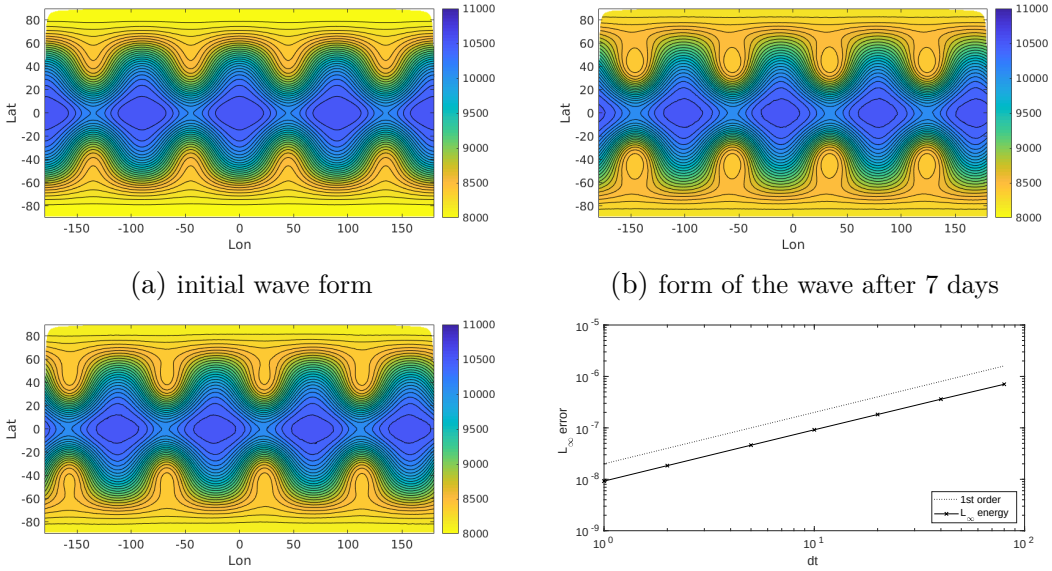
For our method, we need the directional magnitude of the velocity which is $V_{ij} = \mathbf{u}_{ij} \cdot \mathbf{n}_{ij}$ at edge ij.

In Fig. 3.11, it can be seen that the main features of the evolution of the Rossby–Haurwitz wave solution are reproduced correctly. The time series of the errors in the energy and potential enstrophy are depicted in Fig. 3.12.

3.6. Long term simulation

In the previous section, we have reproduced some of the standard test cases proposed in [90] for testing novel numerical schemes for the shallow-water equations. Note that these test cases require the integration of the shallow-water equations for relatively short time intervals, with the longest test case being integrated for $t = 15$ days. A main motivation for developing a geometric numerical integrator is that they should be suitable for longer integrations.

To test the ability of the variational discretization of shallow-water equations to carry out longer integration experiments, we revisit here Case 2, the nonlinear



(a) initial wave form

(b) form of the wave after 7 days

(c) form of the wave after 14 days

(d) convergence of energy error for $n = 10242$

Figure 3.11.: Numerical solution for Rossby-Haurwitz waves. Contour interval is 100m.

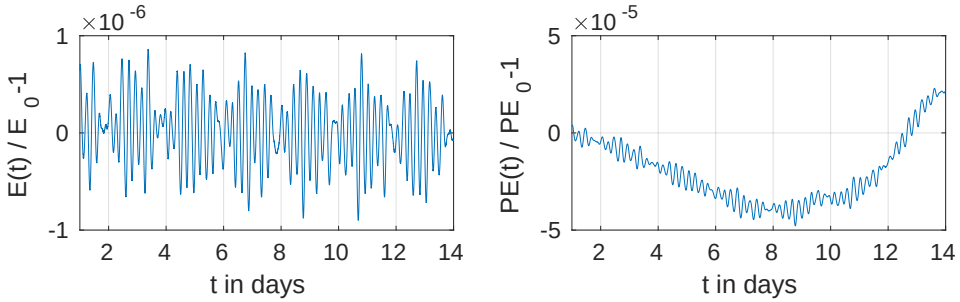


Figure 3.12.: Time series of the errors of energy and potential enstrophy for Rossby–Haurwirz waves over the 14 days integration period.

geostrophic flow on a rotating sphere and Case 3, the flow over a mountain. We test the spatial variational integrator with two different time discretizations, the (variational) Cayley transform and the (non-variational) standard method that applies a Crank–Nicolson time discretization of the continuity equation. Both test cases are integrated for a 50-day period. The time series of the total energy and potential enstrophy errors are depicted in Fig. 3.13 and 3.14. While both time integrators produce reasonable error time series for the short-term integration up to, say, 15 days, for long-term simulations only the variational scheme based on the Cayley transform shows hardly any error trends for both the nonlinear geostrophic

flow and the flow over the mountain test cases. Notably, while the standard time integrator performs well for Case 2, this non-variational method shows a clear trend to lose energy for Case 3.

Moreover, we note that for all test cases studied, the Cayley transform method shows first order convergence of the energy error with respect to the time step while for the standard method, though showing in general good energy conservation properties, such convergence behaviour cannot be guaranteed. For instance, while the energy error converges at first order for the Rossby–Haurwitz wave case, the standard time integrator yields only zero order convergence for the steady state case (not shown).

The slight positive drift in energy conservation that is visible in the top-left panel of Fig. 3.13 is probably related to the fact that the presented RSW scheme is not a fully variational integrator in time (see section Time discretizations). In particular, when using non-uniform grids (such as those used here, with both pentagonal and hexagonal cells in the dual mesh), this effect is enhanced. This drift gets smaller with increased resolution, and we observe very good long-term behaviour over a time period of a couple of months, a reasonable time period in atmospheric and ocean modelling. When considering longer integration times of up to one year, for instance, the drift is more enhanced but stays, nevertheless, very small, as illustrated in Figure 3.15 (right) for the geostrophic flow test case. Similarly to results shown in [8], these trends can be related to the irregularity of the mesh. In particular, we don't observe any trend in energy error when simulating a steady vortex solution on an f-plane approximation of the sphere with a uniform mesh (cf. Figure 3.15, left), but do see such a trend when using a non-uniform mesh (Figure 3.15, middle), until the vortex decays. The study of the influence of the grid structures and the construction of a fully variational time integrator to avoid these trends will be part of future work.

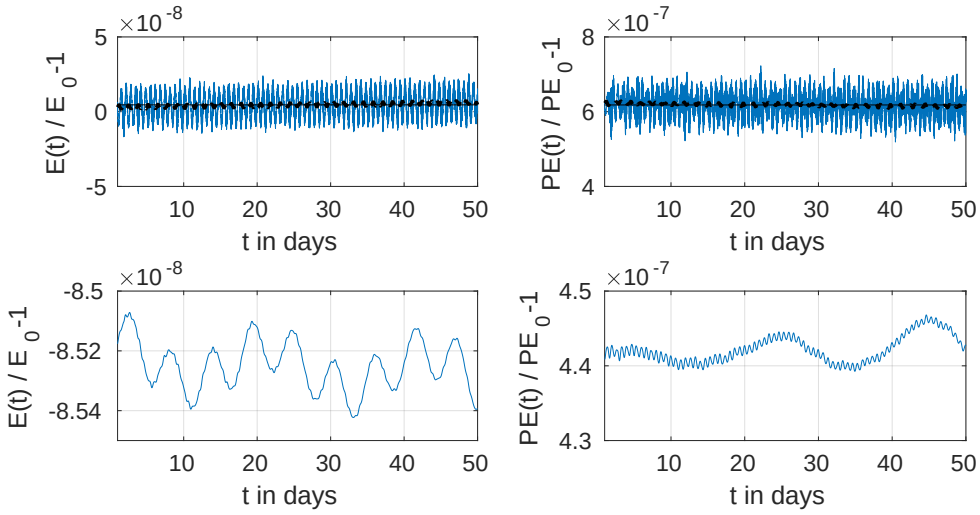


Figure 3.13.: Comparison between the Cayley transform (top panels) and standard time integrator (bottom panels) for the nonlinear geostrophic flow (Case 2) over 50 days. The dotted line indicates the moving mean and the solid straight line the overall mean.

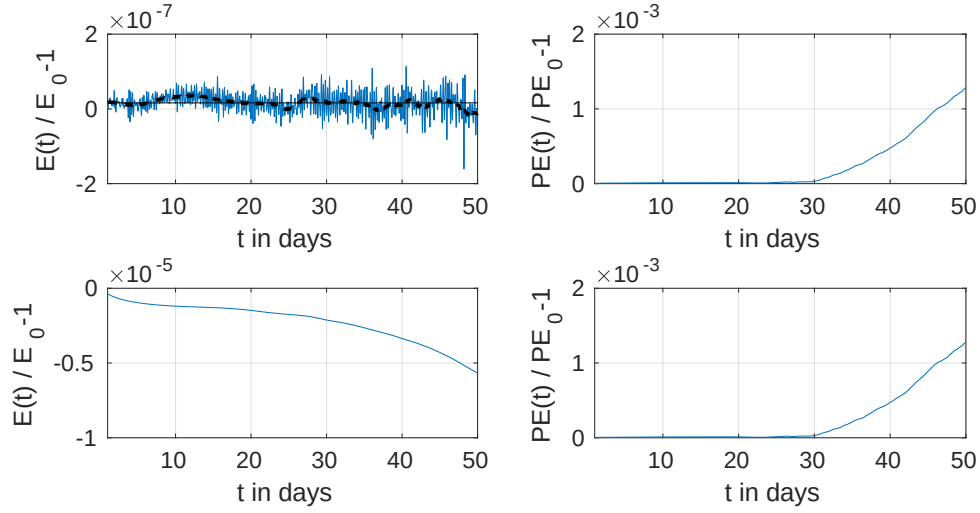


Figure 3.14.: Comparison between the Cayley transform (top panels) and standard time integrator (bottom panels) for the flow over the mountain (Case 3) over 50 days. The dotted line indicates the moving mean and the solid straight line the overall mean.

3.7. Conclusions and outlook

In this paper, we have constructed a space–time variational discretization for the shallow-water equations on a rotating sphere. This discretization is an extension of the variational integrator proposed in [9] for the rotating shallow-water equations

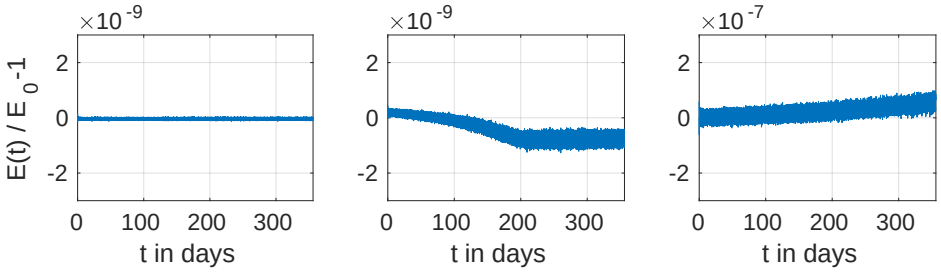


Figure 3.15.: Left and middle plots show the conservation of energy for an isolated vortex (for details see [8]) on a uniform and non-uniform mesh (64^2 triangles) on the f-plane over a period of 1 year. The right plot shows the conservation of energy for Case 2 on the sphere (10242 Voronoi cells) for 1 year.

on the plane. We have carried out some of the standard benchmark tests proposed in [90] and illustrated that the discretization converges at the order of about 0.5 to 1 (constrained by the convergence order of the divergence operator) to reference solutions of the shallow-water equations on the rotating sphere. As expected from the variational integrator, the magnitude of small error fluctuations of the energy around a conserved long-term mean reduce with first order with respect to the time step size. All numerical tests carried out demonstrate the excellent conservation properties of the variational integrator, regarding the conservation of energy, mass, and potential vorticity. Potential enstrophy is not an invariant of the discrete equations by construction; nevertheless it is well preserved.

We would like to stress that the variational integrator on the sphere proposed in this paper exactly conserves the lake-at-rest steady state solution over arbitrary bottom topography. In addition, mass is conserved up to machine precision. Both factors are of significant importance in tsunami propagation models, since they prevent the introduction of spurious waves by the discretization scheme. As such, the variational integrator can serve as dynamical core for a general purpose tsunami model, in particular as our framework permits application of different boundary conditions (e.g., free-slip), cf. [9]. A tsunami model requires a suitable inundation model to handle the time-dependent wet–dry interface. The development of such an inundation model for the presented variational integrator is

currently underway and will be presented elsewhere, as will be the addition of a framework for dynamic grid adaptation, which is another essential ingredient for a tsunami propagation and inundation model.

Acknowledgments

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, the NSERC Discovery Grant program and the InnovateNL CRC LeverageR&D program. WB has received funding from the European Union's Horizon 2020 research and innovative programme under the Marie Sklodowska-Curie grant agreement No 657016. AB is a recipient of an APART Fellowship of the Austrian Academy of Sciences. FGB was partially supported by the ANR project GEOMFLUID, ANR-14-CE23-0002-01. The authors thank Chris Eldred for providing us with the grid generator on the sphere.

3.8. Appendix

3.8.1. Differential forms, flat operator, and Lie derivatives

Let us consider a manifold M of dimension n, for instance the sphere $\mathcal{S} \subset \mathbb{R}^3$ with $n = 2$. For $x \in M$, the vector space T_xM consists of all the (tangent) vectors at the point x. A vector field $\mathbf{u}$ on M is a smooth map that associates to each point $x \in M$ a tangent vector $\mathbf{u}(x) \in T_xM$. In local coordinates $(x^1, ..., x^n)$ of M, a vector field is written $\mathbf{u}(x) = \sum_{i=1}^{n} \mathbf{u}^i(x) \frac{\partial}{\partial x^i}$.

A k-form ω is a skew-symmetric tensor field of rank k on M. That is, it is a smooth map that associates to each point $x \in M$ a multi-linear map $\omega(x)$ that takes k tangent vectors at x as input and returns a real number:

$$\omega(x) : T_xM \times .... \times T_xM \to \mathbb{R}, \quad (v_x^1, ..., v_x^k) \mapsto \omega(x)\Big(v_x^1, ..., v_x^k\Big).$$

A zero-form on M is thus a function on M. In local coordinates $(x_1, ..., x_n)$ of M, a k-form reads $\omega = \sum_{i_1<...<i_k} \omega_{i_1...i_k} dx^{i_1} \wedge ... \wedge dx^{i_k}$. For instance one-forms and two-forms on $M = \mathbb{R}^2$ read $\omega = \omega_1 dx^1 + \omega_2 dx^2$ and $\omega = \omega_{12} dx^1 \wedge dx^2$ in coordinates (x^1, x^2). The exterior differential is a mapping $\mathbf{d}$ from k-forms to $(k+1)$-forms that extends the notion of the differential of a function to differential forms. For instance, in coordinates, the differential of a k-form $\omega = \sum_{i_1<...<i_n} \omega_{i_1...i_k} dx^{i_1} \wedge ... \wedge dx^{i_1}$ is the $(k+1)$-form $\mathbf{d}\omega = \sum_{j=1}^{n} \sum_{i_1<...<i_n} \frac{\partial \omega_{i_1...i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge ... \wedge dx^{i_1}$. Given a vector field $\mathbf{u}$ and a k-form ω, the inner product of $\mathbf{u}$ with ω is the $(k-1)$-form $\mathbf{i}_\mathbf{u}\omega$ defined by $\mathbf{i}_\mathbf{u}\omega(x)\Big(v_x^1, ..., v_x^{k-1}\Big) = \omega(x)\Big(\mathbf{u}(x), v_x^1, ..., v_x^{k-1}\Big)$.

A Riemannian metric on M is a symmetric and positive-definite tensor field γ of rank 2 on M. That is, it is a smooth map that associates to each point $x \in M$ a symmetric bilinear map $\gamma(x)$ that takes 2 tangent vectors as input and returns a real number:

$$\gamma : T_xM \times T_xM \to \mathbb{R}, \quad (v_x^1, v_x^2) \mapsto \gamma(x)\Big(v_x^1, v_x^2\Big),$$

such that $\gamma(v_x, v_x) > 0$, for all $v_x \in T_xM$ with $v_x \neq 0$. In coordinates, we write $\gamma = \sum_{i,j=1}^{n} \gamma_{ij} dx^i dx^j$. The flat operator associated to γ is the linear map that sends a vector field $\mathbf{u}$ to the 1-form $\mathbf{u}^\flat$ defined by $\mathbf{u}^\flat(x)(v_x) = \gamma(x)(\mathbf{u}(x), v_x)$, for all $v_x \in T_xM$. In local coordinates, if $\mathbf{u} = \sum_{i=1}^{n} \mathbf{u}^i \frac{\partial}{\partial x^i}$, then $\mathbf{u}^\flat = \sum_{i,j=1}^{n} \gamma_{ij} \mathbf{u}^j dx^i$. The volume form associated to the Riemannian metric γ is the n-form $d\sigma$, given locally by $d\sigma = \sqrt{\det \gamma_{ij}} dx^1 \wedge ... dx^n$.

Given a diffeomorphism φ of M, the pull-back of a k-form ω is the k-form $\varphi^*\omega$ defined by

$$\varphi^*\omega(x)\Big(v_x^1, ..., v_x^k\Big) := \omega(\varphi(x))\Big(T_x\varphi(v_x^1), ..., T_x\varphi(v_x^k)\Big),$$

where $T_x\varphi : T_xM \to T_{\varphi(x)}M$ is the tangent map (the derivative) of φ. The Lie

derivative of a k-form ω with respect to a vector field $\mathbf{u}$ is the k-form $\mathcal{L}_{\mathbf{u}}\omega$ defined by $\mathcal{L}_{\mathbf{u}}\omega = \frac{d}{dt}\big|_{t=0} \varphi_t^*\omega$, where φ_t is the flow of the vector field $\mathbf{u}$, i.e. $\frac{d}{dt}\varphi_t(x) = \mathbf{u}(\varphi_t(x))$, with $\varphi_{t=0}(x) = x$. The divergence of a vector field relative to the Riemannian metric γ is the function $\operatorname{div}\mathbf{u}$ defined by $\mathcal{L}_{\mathbf{u}}d\sigma = (\operatorname{div}\mathbf{u})d\sigma$.

For the treatment of fluid mechanics, we also need the notion of Lie derivative of a k-form density, denoted $\mathsf{L}_{\mathbf{u}}$ to distinguish it from the Lie derivative $\mathcal{L}_{\mathbf{u}}$ of a k-form. Let us assume that a Riemannian metric γ is fixed (for the sphere $\mathcal{S}$ we chose the Riemannian metric induced from the Euclidean metric on $\mathbb{R}^3$). In this case, k-form densities can be identified with k-forms. The action of a diffeomorphism φ on a k-form density ω is $\omega \bullet \varphi := (\varphi^*\omega)J\varphi$, where $J\varphi$ is the Jacobian of the diffeomorphism with respect to the Riemannian metric γ, defined by $\varphi^*d\sigma = J\varphi d\sigma$. The Lie derivative is defined, similarly as before, by $\mathsf{L}_{\mathbf{u}}\omega = \frac{d}{dt}\big|_{t=0} \omega \bullet \varphi_t = \frac{d}{dt}\big|_{t=0} (\varphi_t^*\omega)J\varphi_t$, where φ_t is the flow of the vector field $\mathbf{u}$. In §3.2, this Lie derivative was applied to the fluid momentum density $\alpha = \frac{\delta \ell}{\delta \mathbf{u}}$ (a one-form density) and to the fluid depth h (a zero-form density), and reads

$$\mathsf{L}_{\mathbf{u}}\alpha = \mathbf{i}_{\mathbf{u}}\mathbf{d}\alpha + \mathbf{d}\mathbf{i}_{\mathbf{u}}\alpha + \alpha \operatorname{div}\mathbf{u} \qquad \text{and} \qquad \mathsf{L}_{\mathbf{u}}h = \operatorname{div}(h\mathbf{u}).$$

In local coordinates, we have $(\mathsf{L}_{\mathbf{u}}\alpha)_i = \sum_{k=1}^n \partial_k(\alpha_i \mathbf{u}^k) + \sum_{k=1}^n \alpha_k \partial_i \mathbf{u}^k$ and $\mathsf{L}_{\mathbf{u}}h = \sum_{k=1}^n \partial_k(h\mathbf{u}^k)$.

3.8.2. Lie group, Lie algebra, and actions

Recall that a Lie group G is a manifold and a group, such that the group operations are smooth maps with respect to the manifold structure. Basic examples are the general linear group $\mathsf{GL}(n) = \{A \in \mathrm{Mat}(n) \mid \det A \neq 0\}$ or the special orthogonal group $\mathsf{SO}(n) = \{R \in \mathrm{GL}(n) \mid R^{\mathsf{T}}R = I_n, \ \det R > 0\}$, where I_n is the $n \times n$ identity matrix. The tangent space to a Lie group at the identity, denoted $\mathfrak{g} = T_eG$, is called the Lie algebra of the Lie group and is naturally endowed with a Lie bracket

$[,]$. For matrix Lie groups, the Lie bracket is the usual commutator of matrices $[A, B] = AB - BA$.

Important Lie groups for this paper are the group $\mathrm{Diff}(M)$ of all smooth diffeomorphisms of the manifold M and its discrete version $\mathsf{D}(\mathbb{M})$ defined in (3.7). The Lie algebra of $\mathrm{Diff}(M)$ is the space of vector fields $\mathbf{u}$ on M, endowed with (minus) the Lie bracket of vector fields $[\mathbf{u}, \mathbf{v}] = \mathbf{u} \cdot \nabla \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u}$. The Lie algebra of the Lie group $\mathsf{D}(\mathbb{M}) = \{q \in \mathrm{GL}(N)^+ \mid q \cdot \mathbf{1} = \mathbf{1}\}$ of discrete diffeomorphisms is obtained by taking the ε-derivative of the relation $q_\epsilon \cdot \mathbf{1} = \mathbf{1}$, for a curve $q_\varepsilon \in \mathsf{D}(\mathbb{M})$ with $q_{\varepsilon=0} = I_N$. We obtain the Lie algebra $\mathfrak{d}(\mathbb{M}) = \{A \in \mathrm{Mat}(N) \mid A \cdot \mathbf{1} = 0\}$ as given in (3.8).

A (right) action of a Lie group G on a manifold M is a map $(g, x) \in G \times M \to x \cdot g \in M$ such that $x \cdot e = x$ and $x \cdot (gh) = (x \cdot g) \cdot h$, for all $x \in M$ and $g, h \in G$, where e is the neutral element in G. The relevant action for the present paper is the action of $\mathrm{Diff}(M)$ on k-forms densities $(\varphi, \omega) \mapsto \omega \bullet \varphi = (\varphi^* \omega) J\varphi$. In particular, we have the action on one-form densities $\alpha \bullet \varphi = (\varphi^* \alpha) J\varphi$ and on zero-form densities $h \bullet \varphi = (h \circ \varphi) J\varphi$, where α represents the momentum fluid density and h represents the fluid depth.

3.8.3. Euler–Poincaré variational principle

We quickly review the Euler–Poincaré theory by applying it to simpler examples, before considering the rotating shallow water equations. Euler–Poincaré variational principle applies to dynamical systems whose configuration manifold is a Lie group G and whose dynamics is given by the Euler–Lagrange equations associated to a Lagrangian which is invariant under a subgroup of G. We refer to [48] for a complete treatment.

The rigid body. One of the simplest example of dynamical system on a Lie group is a rigid body moving about its center of mass. The configuration Lie group is $\mathsf{SO}(3)$, with $R \in \mathsf{SO}(3)$ describing the orientation of the body. The dynamics of the body is described by the Euler–Lagrange equations for the Lagrangian $L : T\mathsf{SO}(3) \to \mathbb{R}$ given by the kinetic energy of the body. Euler–Lagrange equations arise from the Hamilton principle

$$\delta \int_0^T L(R, \dot{R}) \mathrm{d}t = 0, \tag{3.35}$$

with respect to variations δR with $\delta R(0) = \delta R(T) = 0$. This Lagrangian is invariant under the left action of $SO(3)$, i.e., it satisfies $L(AR, A\dot{R}) = L(R, \dot{R})$, for all $A \in \mathsf{SO}(3)$. This invariance allows one to define the reduced Lagrangian $\ell : \mathfrak{so}(3) \to \mathbb{R}$, with $\mathfrak{so}(3)$ the Lie algebra of $\mathsf{SO}(3)$, by $L(R, \dot{R}) = \ell(R^{-1}R)$. The skew-symmetric 3×3 matrix $\Omega = R^{-1}\dot{R}$ is the angular velocity of the rigid body in the body frame. It is natural to ask if it is possible to write the Hamilton principle (3.35) directly in terms of the body angular velocity Ω without referring to the initial variables R and $\dot{R}$. This is indeed the case if one considers the constrained variations of $\Omega = R^{-1}\dot{R}$ induced by the free variations of R. A straightforward computation shows that

$$\delta\Omega = \dot{A} + [\Omega, A], \tag{3.36}$$

where $A = R^{-1}\delta R$ is an arbitrary curve of skew-symmetric 3×3 matrices vanishing at $t = 0, T$. The Hamilton principle (3.35) is thus equivalent to the variational principle

$$\delta \int_0^T \ell(\Omega) dt = 0, \tag{3.37}$$

with respect to variations $\delta\Omega$ of the form (3.36). This is the Euler–Poincaré principle, as it applies to the rigid body. We do not present explicitly the Lagrangian and the equations for the rigid body as we will not use them. We will however use

the variational treatment of the rigid body, especially the passage from (3.35) to (3.37), as a guide for the understanding of the more involved treatment of fluids below.

The incompressible fluid. For the incompressible fluid on a Riemannian manifold M, the configuration Lie group is the group $\mathrm{Diff}_{\mathrm{vol}}(M) = \{\varphi \in \mathrm{Diff}(M) \mid J\varphi = 1\}$ of diffeomorphisms of M that preserve the volume. The equations of motions are given by the Euler–Lagrange equations for the Lagrangian $L : T\,\mathrm{Diff}_{\mathrm{vol}}(M) \to \mathbb{R}$ given by the kinetic energy of the fluid, namely

$$L(\varphi, \dot{\varphi}) = \int_M \frac{1}{2}\gamma(\dot{\varphi}, \dot{\varphi}) d\sigma,$$

where we recall that $d\sigma = \sqrt{\det \gamma_{ij}} dx^1 \wedge ... dx^n$ is the Riemannian volume form associated to γ. The Hamilton principle reads

$$\delta \int_0^T L(\varphi, \dot{\varphi}) \mathrm{d}t = 0, \tag{3.38}$$

with respect to variations $\delta\varphi$ with $\delta\varphi(0) = \delta\varphi(T) = 0$. In the same way the Lagrangian of the rigid body was left-invariant, the Lagrangian of the incompressible fluid is right-invariant: $L(\varphi \circ \psi, \dot{\varphi} \circ \psi) = L(\varphi, \dot{\varphi})$ for all $\psi \in \mathrm{Diff}_{\mathrm{vol}}(M)$. This invariance allows one to define the reduced Lagrangian $\ell : \mathfrak{X}_{\mathrm{vol}}(M) \to \mathbb{R}$, by $L(\varphi, \dot{\varphi}) = \ell(\dot{\varphi} \circ \varphi^{-1})$, where $\mathfrak{X}_{\mathrm{vol}}(M)$ is the Lie algebra of $\mathrm{Diff}_{\mathrm{vol}}(M)$, given by divergence free vector fields on M. The reduced Lagrangian is given by

$$\ell(\mathbf{u}) = \frac{1}{2}\int_M \gamma(\mathbf{u}, \mathbf{u}) d\sigma. \tag{3.39}$$

As before, it is natural to ask if it is possible to write the Hamilton principle (3.38) directly in terms of the Eulerian velocity $\mathbf{u}$ without referring Lagrangian fluid trajectory and velocity φ and $\dot{\varphi}$. This is indeed the case if one considers

the constrained variations of $\mathbf{u} = \dot{\varphi} \circ \varphi^{-1}$ induced by the free variations of φ. A straightforward computation shows that

$$\delta \mathbf{u} = \partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}], \tag{3.40}$$

where $\mathbf{v} = \delta\varphi \circ \varphi^{-1}$ is an arbitrary curve of divergence free vector fields vanishing at $t = 0, T$. The Hamilton principle (3.38) is thus equivalent to the variational principle

$$\delta \int_0^T \ell(\mathbf{u}) dt = 0,$$

with respect to constrained variations $\delta\mathbf{u}$ of the form (3.40). This is the Euler–Poincaré principle, as it applies to the incompressible fluid. A direct application of this principle yields

$$\delta \int_0^T \ell(\mathbf{u}) dt = \int_0^T \int_M \frac{\delta \ell}{\delta \mathbf{u}} \cdot \delta \mathbf{u} d\sigma dt = \int_0^T \int_M \frac{\delta \ell}{\delta \mathbf{u}} \cdot \Big(\partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}] \Big) d\sigma dt \tag{3.41}$$

$$= - \int_0^T \int_M \Big(\partial_t \frac{\delta \ell}{\delta \mathbf{u}} + \mathcal{L}_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} \Big) \cdot \mathbf{v} d\sigma dt, \tag{3.42}$$

where in the first equality the functional derivative $\frac{\delta \ell}{\delta \mathbf{u}}$ is defined as the one-form such that

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \ell(\mathbf{u} + \varepsilon \mathbf{v}) = \int_M \frac{\delta \ell}{\delta \mathbf{u}} \cdot \mathbf{v} \, d\sigma,$$

for arbitrary divergence free vector field $\mathbf{v}$. Since (3.41) has to be zero for all divergence free vector fields $\mathbf{v}$, by the Hodge decomposition there exists a function q such that

$$\partial_t \frac{\delta \ell}{\delta \mathbf{u}} + \mathcal{L}_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} = -\mathbf{d} q.$$

For the Lagrangian (3.39) of the incompressible fluid, we have $\frac{\delta \ell}{\delta \mathbf{u}} = \mathbf{u}^\flat$, hence using $\mathcal{L}_{\mathbf{u}} \mathbf{u}^\flat = \mathbf{i}_{\mathbf{u}} \mathbf{d} \mathbf{u}^\flat + \mathbf{d}(\mathbf{i}_{\mathbf{u}} \mathbf{u}^\flat)$, we get the Euler equations in the form $\partial_t \mathbf{u}^\flat + \mathbf{i}_{\mathbf{u}} \mathbf{d} \mathbf{u}^\flat = -\mathbf{d} p$, with $p = q + \mathbf{i}_{\mathbf{u}} \mathbf{u}^\flat$.

The rotating shallow water equations. The configuration Lie group for the rotating shallow water equations on the sphere is the group $\mathrm{Diff}(\mathcal{S})$ of all diffeomorphisms of $\mathcal{S}$. While the equations are still given by the Euler–Lagrange equations for L, obtained from the Hamilton principle $\delta \int_0^T L(\varphi, \dot{\varphi})dt = 0$, a major difference with the previous two cases is the symmetry group of L. It is no longer given by the whole configuration Lie group but by a subgroup, namely, the subgroup $\mathrm{Diff}(\mathcal{S})_{h_0} \subsetneq \mathrm{Diff}(\mathcal{S})$ of diffeomorphisms that preserves the initial fluid depth h_0. As a consequence, the reduced Lagrangian not only depends on the Eulerian velocity $\mathbf{u}$ as earlier, but also on the current fluid depth h. More precisely, we have

$$L(\varphi, \dot{\varphi}) = \ell(\mathbf{u}, h), \quad \text{for} \quad \mathbf{u} = \dot{\varphi} \circ \varphi^{-1} \quad \text{and} \quad h = (h_0 \circ \varphi^{-1}) J\varphi^{-1}. \tag{3.43}$$

From these relations, the free variations $\delta\varphi$ of the fluid trajectory φ induce the following constrained variations of $\mathbf{u}$ and h:

$$\delta\mathbf{u} = \partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}] \quad \text{and} \quad \delta h = -\operatorname{div}(h\mathbf{v}),$$

where $\mathbf{v} = \delta\varphi \circ \varphi^{-1}$ is an arbitrary curve of divergence free vector fields vanishing at $t = 0, T$. This is the variational principle given in (3.3). We now show the derivation of (3.4). We have

$$\begin{aligned} \delta \int_0^T \ell(\mathbf{u}, h)dt &= \int_0^T \int_M \Big(\frac{\delta \ell}{\delta \mathbf{u}} \cdot \delta\mathbf{u} + \frac{\delta \ell}{\delta h} \delta h \Big) d\sigma dt \\ &= \int_0^T \int_M \Big(\frac{\delta \ell}{\delta \mathbf{u}} \cdot \big(\partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}] \big) - \frac{\delta \ell}{\delta h} \operatorname{div}(h\mathbf{v}) \Big) d\sigma dt \\ &= \int_0^T \int_M \big(-\partial_t \frac{\delta \ell}{\delta \mathbf{u}} - \mathsf{L}_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} + h\mathbf{d} \frac{\delta \ell}{\delta h} \big) \cdot \mathbf{v} d\sigma dt, \end{aligned} \tag{3.44}$$

where we used integration by parts with $\mathbf{v}(0) = \mathbf{v}(T) = 0$. Since $\mathbf{v}$ is arbitrary, (3.4) follows.

Replacing the functional derivatives (3.5) for the RSW Lagrangian in (3.4), we

get

$$\partial_t(h(\mathbf{u}^\flat + \mathbf{R}^\flat)) + \mathsf{L}_{\mathbf{u}}\Big(h(\mathbf{u}^\flat + \mathbf{R}^\flat)\Big) = h\mathbf{d}\bigg(\frac{1}{2}\gamma(\mathbf{u},\mathbf{u}) + \gamma(\mathbf{R},\mathbf{u}) - g(h+B)\bigg). \quad (3.45)$$

Using the formulas for the Lie derivative recalled earlier, we have $\mathsf{L}_{\mathbf{u}}\Big(h(\mathbf{u}^\flat + \mathbf{R}^\flat)\Big) = \mathrm{div}(h\mathbf{u})(\mathbf{u}^\flat + \mathbf{R}^\flat) + \mathbf{i}_{\mathbf{u}}\mathbf{d}(\mathbf{u}^\flat + \mathbf{R}^\flat) + \mathbf{d}\Big(\mathbf{i}_{\mathbf{u}}(\mathbf{u}^\flat + \mathbf{R}^\flat)\Big)$. Inserting this relation into (3.45) and simplifying further by using the advection equation $\partial_t h + \mathrm{div}(h\mathbf{u}) = 0$, which follows from the second equation in (3.43), we get the RSW equations in the form given in (3.6).

4. Rotating shallow water flow under location uncertainty with a structure-preserving discretization

submitted to Journal of Advances in Modeling Earth Systems

Rüdiger Brecht[†], Long Li[§], Werner Bauer[‡], Etienne Mémin[§]

[†] Department of Mathematics and Statistics, Memorial University of Newfoundland,
St. John's (NL) A1C 5S7, Canada

[§] Inria/IRMAR, Campus universitaire de Beaulieu, Rennes, France

[‡] Imperial College London, Department of Mathematics, 180 Queen's Gate, London SW7 2AZ, United Kingdom.

Abstract

We introduce a new representation of the rotating shallow water equations based on a stochastic transport principle. The derivation relies on a decomposition of the fluid flow into a large-scale component and a noise term that models the unresolved small-scale flow. The total energy of such a random model is demonstrated to be preserved along time for any realization. To preserve this structure, we combine an energy (in space) preserving discretization of the underlying deterministic model with approximations of the stochastic terms that are based on standard finite volume/difference operators. This way, our method can directly be used in existing dynamical cores of global numerical weather prediction and climate models. For an inviscid test case on the f-plane we use a homogenous noise and illustrate that the spatial part of the stochastic scheme preserves the total energy of the system. Moreover, using an inhomogenous noise, we show for a barotropically unstable jet on the sphere that the proposed random model better captures the structure of a large-scale flow than a comparable deterministic model.

Plain summary

The motion of geophysical fluids on the globe needs to be modelled to get some insights of tomorrow's climate. These forecasts must be precise enough while remaining computationally affordable. An ideal system should also deliver, across time, an accurate measurement of the uncertainties introduced through physical or numerical approximations. To address these issues, we use the rotating shallow water equations, which provide a simplified version of the dynamics, and a stochastic representation of the unresolved small-scale processes. The former is approximated with a structure preserving numerical model enabling the conservation of physical quantities such as mass and energy and the latter is modelled

by the location uncertainty framework that relies on stochastic transport and has the great advantage to be energy conserving. Our method can directly be used in existing dynamical cores of global numerical weather prediction and climate models. Numerical results illustrate the energy conservation of the numerical model. Simulating a barotropically unstable jet on the sphere, we demonstrate that the random model better captures the structure of a large-scale flow than a comparable deterministic model. The random dynamical system is also shown to be associated with good uncertainty representation.

4.1. Introduction

Numerical simulations of the Earth's atmosphere and ocean plays an important role in developing our understanding of weather forecasting. A major focus lies in determining the large scale flow correctly, which is strongly related to the parameterizations of sub-grid processes [33]. The non-linear and non-local nature of the dynamical system make the large-scale flow structures interact with the smaller components. The computational expense for solving the Kolmogorov scales [67] of a geophysical flows is fare beyond reach today and likely in the future. Thus, the effect of unresolved scales has to be modeled or parametrized.

For several years, there is a growing interest in geophysical sciences to incorporate a stochastic representation [31, 39, 41, 57] of the small-scale processes. In this study, we propose to stick to a specific stochastic model, the so-called *Location Uncertainty* (LU) derived by [62], which emerges from a decomposition of the Lagrangian velocity into a time-smooth drift and a highly oscillating uncertainty term. Such random model allows us to develop by stochastic calculus a new stochastic transport operator [69] for the extensive scalars. In particular, this transport operator involves a multiplicative random forcing, a heterogeneous diffusion and a corrected advection resulting from the inhomogeneity of the random

flow. This stochastic transport principle has been used as a fundamental tool to derive stochastic representations of large-scale geophysical dynamics [10, 22, 69]. In the present work, we use this mathematical principle together with some physical conservation laws to derive a stochastic version of the rotating shallow water (RSW) system. One strong property of this random model is that it preserves the total energy of the resolved flow in time for each realization.

Recently, the LU model performed very well in [10, 11, 70, 71] when studying oceanic quasi-geostrophic flows. It was found to be more accurate in predicting the extreme events, in diagnosing the frontogenesis and filamentogenesis, in structuring the large-scale flow and in reproducing the long-term statistics. Besides, [22] investigated the Lorentz-63 test case and demonstrated that the LU model was more effective in exploring the range of the strange attractors compared to classical models.

In this work, the performance of the LU model is assessed for the numerical simulation of the RSW system, which can be considered as the first step towards developing global random numerical weather prediction and climate models. This is the first time that the LU model is implemented for the dynamics evolving on the sphere.

We propose to combine the discrete variational integrator for RSW fluids as introduced in [8] and [20] with the numerical LU model in order to mimic the continuous conservation properties. Variational integrators are designed by first discretizing the given Lagrangian, and then by deriving a discrete system of associated Euler-Lagrange equations from the discretized Lagrangian (see [58]). The advantage of this approach is that the resulting discrete system inherits several important properties of the underlying continuous system, notably a discrete version of Noether's theorem that guarantees the preservation of conserved quantities associated to the symmetries of the discrete Lagrangian (see [42]). Variational integrators also exhibit superior long-term stability properties. Therefore, they

typically outperform traditional integrators if one is interested in long-time integration or the statistical properties of a given dynamical system. The benefit of the proposed method that relies on a combination of a variational integrator with a potentially differently approximated LU model is that it can directly be applied to existing dynamical cores of numerical weather prediction and climate models.

Apart from taking into account the unresolved processes, it is paramount in uncertainty quantification and ensemble forecasting to model the uncertainties along time [72]. For a long time, operational weather forecast centres had relied on random perturbations of initial conditions (PIC) to spread the ensemble forecasts. However, in the application of data assimilation to geophysical fluid dynamics, such PIC model is known to under-estimate the true uncertainty compared to the observations [32, 40]. Hence, an assimilation system is overconfident for such a random model. To overcome this issue, the covariance inflation method [3] is often adopted, in which the ensemble covariance is increased by a carefully tuned parameter. In the present work, we compare the reliability of the ensemble spread of such a PIC model with our RSW-LU system, under the same strength of uncertainty.

The remainder of this paper is structured as follows. Section 4.2 describes the basic principles of the LU model and the derivation of the rotating shallow water system under LU associated with the energy conservation property. Section 4.3 explains the parameterizations of the uncertainty and the numerical discretization of the stochastic dynamical system. Section 4.4 discusses the numerical results for an inviscid test case with homogeneous noise and a viscous test case with heterogeneous noise. Finally, in Section 4.5 we draw some conclusions and provide an outlook for future work.

4.2. Rotating shallow water equations under location uncertainty

In this section, we first review the LU model introduced by [62], then we derive the rotating shallow water equations under LU, denoted as RSW–LU, following the classical strategy as shown in [86]. In particular, we demonstrate one important characteristic of the RSW–LU, namely it preserves the total energy of the large-scale flow.

4.2.1. Location uncertainty principles

The LU model is based on a temporal-scale-separation assumption of the following stochastic flow:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t, t)\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\,\mathrm{d}\boldsymbol{B}_t, \tag{4.1}$$

where $\boldsymbol{X}$ is the Lagrangian displacement defined within the bounded domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3), $\boldsymbol{w}$ is the large-scale velocity that is both spatially and temporally correlated, and $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is the small-scale uncertainty (also called noise) term that is only correlated in space. The spatial structure of such noise is specified through a deterministic integral operator $\boldsymbol{\sigma} : (L^2(\Omega))^d \to (L^2(\Omega))^d$, acting on square integrable vector-valued functions $\boldsymbol{f} \in (L^2(\Omega))^d$, with a bounded kernel $\breve{\boldsymbol{\sigma}}$ such that

$$\boldsymbol{\sigma}[\boldsymbol{f}](\boldsymbol{x}, t) = \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\boldsymbol{f}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \quad \forall \boldsymbol{f} \in (L^2(\Omega))^d. \tag{4.2}$$

The randomness of such noise is driven by the cylindrical $\mathbf{I}_d$-Wiener process $\boldsymbol{B}_t$ [25]. The fact that the kernel is bounded, *i.e.* $\sup_{(\boldsymbol{x},\boldsymbol{y})\in\Omega} |\breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y})| < +\infty$, implies that the operator $\boldsymbol{\sigma}$ is Hilbert-Schmidt on $(L^2(\Omega))^d$. Therefore, the resulting small-scale flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is a centered (of null ensemble mean) Gaussian process

with the following *covariance tensor*, denoted as $\boldsymbol{Q}$, being well-defined:

$$\begin{aligned}\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y},t,s) &= \mathbb{E}\Big[\Big(\boldsymbol{\sigma}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{B}_t\Big)\Big(\boldsymbol{\sigma}(\boldsymbol{y},s)\,\mathrm{d}\boldsymbol{B}_s\Big)^T\Big] \\ &= \delta(t-s)\,\mathrm{d}t\int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z},t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y},\boldsymbol{z},s)\,\mathrm{d}\boldsymbol{z},\end{aligned} \tag{4.3}$$

where $\mathbb{E}$ stands for the expectation and δ is the Kronecker symbol. The strength of the noise is measured by its *variance*, denoted as $\boldsymbol{a}$, which is given by the diagonal components of the covariance per unit of time:

$$\boldsymbol{a}(\boldsymbol{x},t) \triangleq \boldsymbol{Q}(\boldsymbol{x},\boldsymbol{x},t,t)/\mathrm{d}t = \boldsymbol{\sigma}\boldsymbol{\sigma}^T(\boldsymbol{x},t). \tag{4.4}$$

We remark that such variance tensor $\boldsymbol{a}$ has the same unit as a diffusion tensor ($\mathrm{m}^2\cdot\mathrm{s}^{-1}$) and that the density of the turbulent kinetic energy (TKE) can be specified by $\frac{1}{2}\mathrm{tr}(\boldsymbol{a})/\mathrm{d}t$.

The previous representation (4.2) is a general way to define the noise in LU models. In particular, the fact that $\boldsymbol{\sigma}$ is Hilbert-Schmidt ensures that the covariance operator per unit of time, $\boldsymbol{Q}/\mathrm{d}t$, admits an orthogonal eigenfunction basis $\{\boldsymbol{\Phi}_n(\bullet,t)\}_{n\in\mathbb{N}}$ weighted by the eigenvalues $\Lambda_n \geq 0$ such that $\sum_{n\in\mathbb{N}}\Lambda_n < \infty$. Therefore, one may equivalently define the noise and its variance, based on the following spectral decomposition:

$$\boldsymbol{\sigma}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{B}_t = \sum_{n\in\mathbb{N}}\boldsymbol{\Phi}_n(\boldsymbol{x},t)\,\mathrm{d}\beta_t^n, \tag{4.5a}$$

$$\boldsymbol{a}(\boldsymbol{x},t) = \sum_{n\in\mathbb{N}}\boldsymbol{\Phi}_n(\boldsymbol{x},t)\boldsymbol{\Phi}_n^T(\boldsymbol{x},t), \tag{4.5b}$$

where β^n denotes n independent and identically distributed (i.i.d.) one-dimensional standard Brownian motions.

The core of LU models is based on a stochastic Reynolds transport theorem

(SRTT), introduced by [62], which describes the rate of change of a random scalar q transported by the stochastic flow (4.1) within a flow volume $\mathcal{V}$. In particular, for incompressible small-scale flows, $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$, the SRTT can be written as

$$\mathrm{d}_t \Big(\int_{\mathcal{V}(t)} q(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \Big) = \int_{\mathcal{V}(t)} \Big(\mathbb{D}_t q + q\, \boldsymbol{\nabla}\cdot (\boldsymbol{w} - \boldsymbol{w}_s) \Big)\, \mathrm{d}\boldsymbol{x}, \tag{4.6a}$$

$$\mathbb{D}_t q \triangleq \mathrm{d}_t q + (\boldsymbol{w} - \boldsymbol{w}_s)\cdot\boldsymbol{\nabla}\, q\, \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot\boldsymbol{\nabla}\, q - \frac{1}{2}\, \boldsymbol{\nabla}\cdot (\boldsymbol{a}\boldsymbol{\nabla} q)\, \mathrm{d}t, \tag{4.6b}$$

$$\mathrm{d}_t q \triangleq q_{t+\mathrm{d}t} - q_t, \qquad \boldsymbol{w}_s \triangleq \frac{1}{2}\, \boldsymbol{\nabla}\cdot \boldsymbol{a}, \tag{4.6c}$$

in which the stochastic transport operator $\mathbb{D}_t$ [69] and the Itô-Stokes drift (ISD) $\boldsymbol{w}_s$ [10] are included. The latter term arises from the effect of statistical inhomogeneity of the small-scale flow on the large-scale component, which can be considered as a generalization of the Stokes drift in ocean circulations. In the definition of the stochastic transport operator in (4.6b), the first term on the right-hand side (RHS), defined in (4.6c), stands for a forward time-increment of q at a fixed point $\boldsymbol{x}$, and the last two terms describe, respectively, a backscattering from the small-scales to the large-scales and an inhomogeneous diffusion at the small-scales. In particular, for an isochoric flow with $\boldsymbol{\nabla}\cdot(\boldsymbol{w} - \boldsymbol{w}_s) = 0$, one may immediately deduce from (4.6a) the following transport equation of an extensive scalar:

$$\mathbb{D}_t q = 0, \tag{4.7}$$

where the energy of such random scalar q is globally conserved, as shown in [69]:

$$\mathrm{d}_t \Big(\int_{\Omega} \frac{1}{2} q^2\, \mathrm{d}\boldsymbol{x} \Big) = \Big(\underbrace{\frac{1}{2} \int_{\Omega} q\, \boldsymbol{\nabla}\cdot (\boldsymbol{a}\boldsymbol{\nabla} q)\, \mathrm{d}\boldsymbol{x}}_{\text{Energy loss by diffusion}} + \underbrace{\frac{1}{2} \int_{\Omega} (\boldsymbol{\nabla} q)^T \boldsymbol{a}\boldsymbol{\nabla} q\, \mathrm{d}\boldsymbol{x}}_{\text{Energy intake by noise}} \Big)\, \mathrm{d}t = 0. \tag{4.8}$$

Indeed, this can be interpreted as a process where the energy brought by the noise is exactly counter-balanced by that dissipated from the diffusion term.

4.2.2. Derivation of RSW–LU

This section describes in detail the derivation of the RSW–LU system. We remark that a formulation of the shallow water equations under LU in a non-rotating frame is outlined by [62], whereas the new model that we present in this work is fully stochastic and includes rotation such that it is suited for simulations of geophysical flows.

The above SRTT (4.6a) and Newton's second principle allow us to derive the following (three-dimensional) stochastic equations of motions in a rotating frame [10, 69]:

Horizontal momentum equation :

$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times \Big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t\Big) = -\frac{1}{\rho}\boldsymbol{\nabla}_{H}\Big(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}\Big) + \nu\nabla^2\Big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t\Big), \tag{4.9a}$$

Vertical momentum equation :

$$\mathbb{D}_t w = -\frac{1}{\rho}\partial_z\Big(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}\Big) - g\,\mathrm{d}t + \nu\nabla^2\Big(w\,\mathrm{d}t + \sigma_z\mathrm{d}B_t\Big), \tag{4.9b}$$

Mass equation :

$$\mathbb{D}_t \rho = 0, \tag{4.9c}$$

Continuity equation :

$$\boldsymbol{\nabla}\cdot\Big(\boldsymbol{w} - \boldsymbol{w}_s\Big) = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0, \tag{4.9d}$$

where $\boldsymbol{u}$ and w are the horizontal and vertical components of the three-dimensional large-scale flow $\boldsymbol{w}$; $\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t$ and $\sigma_z\mathrm{d}B_t$ are the horizontal and vertical components of the small-scale flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$; $\boldsymbol{f} = (2\tilde{\Omega}\sin\Theta)\boldsymbol{k}$ is the Coriolis parameter varying in latitude Θ, with the Earth's angular rotation rate $\tilde{\Omega}$ and the vertical unit

vector $\boldsymbol{k} = [0, 0, 1]^T$; ρ is the fluid density; $\boldsymbol{\nabla}_H = [\partial_x, \partial_y]^T$ denotes the horizontal gradient; p and $\dot{p}_t^\sigma \triangleq \mathrm{d}p_t^\sigma/\mathrm{d}t$ (informal definition) are the time-smooth and time-uncorrelated components of the pressure field, respectively; g is the Earth's gravity value and ν is the kinematic viscosity. For the following derivation of the shallow water equations we drop the viscous terms.

In order to model the large-scale circulations in the atmosphere and ocean, the hydrostatic balance approximation is widely adopted [86]. Under a small aspect ratio, $\mathcal{H}/\mathcal{L} \ll 1$ with $\mathcal{L}$ and $\mathcal{H}$ the horizontal and vertical scales of the motion, the acceleration term $\mathbb{D}_t w$ on the left-hand side (LHS) of Equation (4.9b) has a lower order of magnitude than the RHS terms, hence the vertical momentum equation reduces to

$$\partial_z\big(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\big) = -g\,\mathrm{d}t. \tag{4.10a}$$

According to the Doob's theorem – unique decomposition of a semimartingale process [53], the previous equation is equivalent to

$$\partial_z p = -\rho g, \qquad \partial_z \mathrm{d}p_t^\sigma = 0. \tag{4.10b}$$

Integrating vertically these hydrostatic balances (4.10b) from 0 to z (see Figure 4.1) under a constant density ρ_0, we have

$$p(x, y, z, t) = p_0(x, y, t) - \rho_0 g z, \tag{4.10c}$$

$$\mathrm{d}p_t^\sigma(x, y, z, t) = \mathrm{d}p_t^\sigma(x, y, 0, t), \tag{4.10d}$$

where p_0 denotes the pressure at the bottom of the basin ($z = 0$). Following [86], we assume that the weight of the overlying fluid is negligible, *i.e.* $p(x, y, \eta, t) \approx 0$ with η the height of the free surface, leading to $p_0 = \rho_0 g \eta$. This allows us to

rewrite Equation (4.10c) such that for any $z \in [0, \eta]$ we have

$$p(x, y, z, t) = \rho_0 g\Big(\eta(x, y, t) - z\Big). \tag{4.10e}$$

Subsequently, the pressure gradient forces in the horizontal momentum equation (4.9a) reduce to

$$-\frac{1}{\rho_0}\boldsymbol{\nabla}_{H}\Big(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}\Big) = -g\boldsymbol{\nabla}_{H}\eta - \frac{1}{\rho_0}\boldsymbol{\nabla}_{H}\mathrm{d}p_t^{\sigma}, \tag{4.10f}$$

which do not depend on z according to Equations (4.10e) and (4.10d). Therefore, the acceleration terms on the LHS of Equation (4.9a) must not depend on z, hence the shallow water momentum equation can be written as

$$\mathbb{D}_t^{H}\boldsymbol{u} + \boldsymbol{f} \times \Big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t\Big) = -g\boldsymbol{\nabla}_{H}\eta\,\mathrm{d}t - \frac{1}{\rho_0}\boldsymbol{\nabla}_{H}\mathrm{d}p_t^{\sigma}, \tag{4.11a}$$

$$\mathbb{D}_t^{H}u \triangleq \mathrm{d}_t u + \Big((\boldsymbol{u} - \boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t\Big) \cdot \boldsymbol{\nabla}_{H}u - \frac{1}{2}\boldsymbol{\nabla}_{H} \cdot \Big(\boldsymbol{a}_{H}\boldsymbol{\nabla}_{H}u\Big)\,\mathrm{d}t, \tag{4.11b}$$

$$\boldsymbol{u}_s \triangleq \frac{1}{2}\boldsymbol{\nabla}_{H} \cdot \boldsymbol{a}_{H}, \qquad \boldsymbol{a} = \begin{pmatrix} \boldsymbol{a}_{H} & \boldsymbol{a}_{Hz} \\ \boldsymbol{a}_{Hz} & a_z \end{pmatrix}, \tag{4.11c}$$

where $\mathbb{D}_t^{H}$ is the horizontal stochastic transport operator, $\boldsymbol{u}_s$ is the two-dimensional ISD, $\boldsymbol{a}_{H}$, a_z and $\boldsymbol{a}_{Hz}$ are the horizontal, vertical and cross components of the three-dimensional variance tensor $\boldsymbol{a}$. Note that Equation (4.11a) is valid only when the cross component $\boldsymbol{a}_{Hz}$ is vertically independent, *i.e.* $\partial_z \boldsymbol{a}_{Hz} = 0$. For instance, one may consider that the horizontal small-scale flow $\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t$ is spatially uncorrelated with the vertical small-scale flow $\sigma_z \mathrm{d}B_t$, *i.e.* $\boldsymbol{a}_{Hz} = 0$.

In order to derive the shallow water mass equation, let us first integrate vertically

the continuity equation (4.9d) from the bottom topography η_b to the free surface η (see Figure 4.1):

$$(w - w_s)|_{z=\eta} - (w - w_s)|_{z=\eta_b} = -h\boldsymbol{\nabla}_{H} \cdot (\boldsymbol{u} - \boldsymbol{u}_s), \tag{4.12a}$$

$$\sigma \mathrm{d}B_t|_{z=\eta} - \sigma \mathrm{d}B_t|_{z=\eta_b} = -h\boldsymbol{\nabla}_{H} \cdot \boldsymbol{\sigma}_{H} \mathrm{d}\boldsymbol{B}_t, \tag{4.12b}$$

where $h = \eta - \eta_b$ denotes the thickness of the water column. On the other hand, a small vertical (Eulerian) displacement at the top and the bottom of the fluid leads to a variation of the position of a particular fluid element [86]:

$$\Big((w - w_s)\,\mathrm{d}t + \sigma \mathrm{d}B_t\Big)\Big|_{z=\eta} = \mathbb{D}_t^H \eta, \tag{4.12c}$$

$$\Big((w - w_s)\,\mathrm{d}t + \sigma \mathrm{d}B_t\Big)\Big|_{z=\eta_b} = \mathbb{D}_t^H \eta_b. \tag{4.12d}$$

Combining Equations (4.12), we deduce the following stochastic mass equation:

$$\mathbb{D}_t^H h + h\boldsymbol{\nabla}_{H} \cdot \Big((\boldsymbol{u} - \boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}_{H} \mathrm{d}\boldsymbol{B}_t\Big) = 0. \tag{4.13}$$

The above two equations (4.11a) and (4.13) constitute a general formulation of the RSW–LU system. Using different levels of noise strength in such a stochastic system allows us to model different physical regimes of the large-scale flow. To characterise these regimes, [70] introduced the following scaling number

$$\epsilon = \frac{\mathcal{T}_\sigma}{\mathcal{T}} \frac{TKE}{MKE}, \tag{4.14a}$$

where $\mathcal{T}$ and $\mathcal{T}_\sigma$ are the correlation time scales of the large-scale flow and the small-scale component, respectively. The mean kinetic energy scale (MKE) is

given by $\mathcal{U}^2$ with $\mathcal{U} = \mathcal{L}/\mathcal{T}$ the typical velocity scale, and the turbulent kinetic energy scale (TKE) is defined by $\mathcal{A}/\mathcal{T}_\sigma$ with $\mathcal{A}$ the magnitude of the variance tensor $\boldsymbol{a}$. As such, the dimensional noise associated with its dimensional variance can be specified by

$$\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon}\,\mathcal{L}\,(\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t)', \qquad \boldsymbol{a} = \epsilon\mathcal{U}\mathcal{L}\,\boldsymbol{a}', \tag{4.14b}$$

where $\bullet'$ denotes adimensional variables. From expressions (4.14), one may easily conclude that the greater the scaling number ϵ, the stronger the noise $\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t$ (with higher variance $\boldsymbol{a}$). Furthermore, as interpreted in [71], a strong noise ($\epsilon \gg 1$) modifies the classical geostrophic equilibrium of the large-scale flow by including some correction terms to the isobaric velocities. In the present work, only moderate noise ($\epsilon \sim 1$) is adopted for the RSW–LU system. Under such assumption, the small-scale flow becomes approximately geostrophic and incompressible, *i.e.* $\boldsymbol{f}\times\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t \approx -\frac{1}{\rho_0}\boldsymbol{\nabla}_{H}\mathrm{d}p_t^{\sigma}$ and $\boldsymbol{\nabla}_{H}\cdot\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t = 0$. As a result, the RSW–LU system simplifies to

$$\mathbb{D}_t^{H}\boldsymbol{u} + \boldsymbol{f}\times\boldsymbol{u}\,\mathrm{d}t = -g\boldsymbol{\nabla}_{H}\eta\,\mathrm{d}t, \tag{4.15a}$$

$$\mathbb{D}_t^{H}h + h\boldsymbol{\nabla}_{H}\cdot(\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t = 0, \tag{4.15b}$$

$$\boldsymbol{\nabla}_{H}\cdot\boldsymbol{\sigma}_{H}\mathrm{d}\boldsymbol{B}_t = 0. \tag{4.15c}$$

We remark that an additional incompressible constraint must be imposed on the horizontal ISD, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0$, so that the previous system preserves the total energy of the large-scale flow. This will be shown in the subsequent section. For the sake of readability, in the following we drop the symbol H for all horizontal variables.

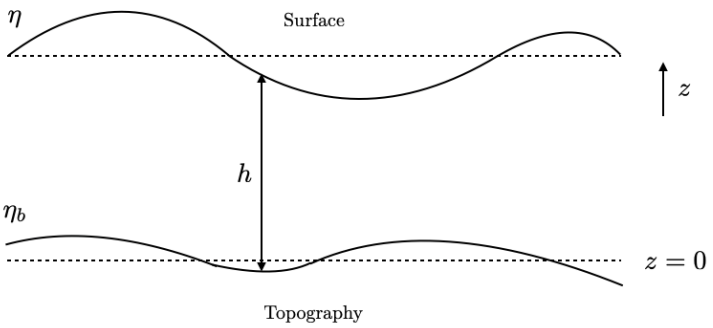


Figure 4.1.: Illustration of a single-layered shallow water system (inspired by [86]). h is the thickness of a water column, η is the height of the free surface and η_b is the height of the bottom topography. As a result, we have $h = \eta - \eta_b$.

4.2.3. Energy conservation of RSW–LU

This section demonstrates the energy conservation of the RSW–LU system (4.15). Let us recall that the density of the kinetic energy (KE) and of the potential energy (PE) of the large-scale flow in the shallow water system [86] is, respectively, given by

$$\mathrm{KE} = \int_0^h \frac{\rho_0}{2} |\boldsymbol{u}|^2 \, \mathrm{d}z = \frac{\rho_0}{2} h |\boldsymbol{u}|^2, \tag{4.16a}$$

$$\mathrm{PE} = \int_0^h \rho_0 g z \, \mathrm{d}z = \frac{\rho_0}{2} g h^2. \tag{4.16b}$$

The density of total energy is defined as the sum of them:

$$\mathrm{E} = \mathrm{KE} + \mathrm{PE} \tag{4.16c}$$

where $|\boldsymbol{u}|^2 = \boldsymbol{u} \cdot \boldsymbol{u}$ and we assume that $\rho_0 = 1$ and the bottom is flat, *i.e.* $\eta_b = 0$ for algebraic simplicity.

In order to explain the conservation of energy more concisely, we adopt an equivalent Stratonovich representation of the RSW–LU system (4.15), namely

$$\mathbb{D}_t \circ \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} \, \mathrm{d}t = -g \boldsymbol{\nabla} h \, \mathrm{d}t, \tag{4.17a}$$

$$\mathbb{D}_t \circ h + h\, \boldsymbol{\nabla}\boldsymbol{\cdot}\, (\boldsymbol{u} - \boldsymbol{u}_s)\, \mathrm{d}t = 0, \tag{4.17b}$$

$$\boldsymbol{f} \times \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = -\boldsymbol{\nabla} \mathrm{d}_t \circ p^{\sigma}, \qquad \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = 0, \tag{4.17c}$$

$$\mathbb{D}_t \circ u \triangleq \mathrm{d}_t \circ u + \Big((\boldsymbol{u} - \boldsymbol{u}_s)\, \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t \Big) \boldsymbol{\cdot} \boldsymbol{\nabla}\, u, \tag{4.17d}$$

where $\mathrm{d}_t \circ \triangleq u_{t+dt/2} - u_{t-dt/2}$ stands for a central time-increment based on the definition of Stratonovich integrals and $\mathbb{D}_t \circ$ denotes the stochastic transport operator under Stratonovich notations. We remark that the equivalence between the Itô form (4.11b) and the Stratonovich form (4.17d) are fully detailed in Appendix C of [10]. As shown by [53], Stratonovich integrals are defined such that the chain rule and the integration-by-part formula of ordinary calculus holds. In particular, for two random tracers f and g, we have

$$\mathrm{d}_t \circ (fg) = f \mathrm{d}_t \circ g + g \mathrm{d}_t \circ f. \tag{4.18a}$$

Therefore, from the definition of the Stratonovich transport operator (4.17d), we deduce the following product rule:

$$\mathbb{D}_t \circ (fg) = g \mathbb{D}_t \circ f + f \mathbb{D}_t \circ g. \tag{4.18b}$$

Applying this rule on the definition of PE (4.16b) together with the mass equation (4.15b),

$$\mathbb{D}_t \circ \mathrm{PE} = gh \mathbb{D}_t \circ h = -gh^2\, \boldsymbol{\nabla}\boldsymbol{\cdot}\, (\boldsymbol{u} - \boldsymbol{u}_s)\, \mathrm{d}t, \tag{4.19a}$$

or

$$\mathbb{D}_t \circ \mathrm{PE} + 2\mathrm{PE}\, \boldsymbol{\nabla}\boldsymbol{\cdot}\, (\boldsymbol{u} - \boldsymbol{u}_s)\, \mathrm{d}t = 0. \tag{4.19b}$$

Similarly, from both mass equation and momentum equation in (4.15), we derive the evolution of KE (4.16a):

$$\begin{aligned}\mathbb{D}_t \circ \mathrm{KE} &= h\boldsymbol{u} \cdot \mathbb{D}_t \boldsymbol{u} + \frac{1}{2}|\boldsymbol{u}|^2 \mathbb{D}_t \circ h \\ &= -\frac{1}{2}\boldsymbol{u} \cdot \boldsymbol{\nabla}\left(gh^2\right) \mathrm{d}t - \frac{1}{2} h|\boldsymbol{u}|^2 \, \boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \, \mathrm{d}t, \end{aligned} \tag{4.19c}$$

noting that $\boldsymbol{u} \cdot (\boldsymbol{f} \times \boldsymbol{u} \, \mathrm{d}t) = 0$ and recalling that $\eta_b = 0$, which yields

$$\mathbb{D}_t \circ \mathrm{KE} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \, \mathrm{PE} \, \mathrm{d}t + \mathrm{KE} \, \boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \, \mathrm{d}t = 0. \tag{4.19d}$$

Subsequently, we deduce the evolution of the density of total energy:

$$\mathbb{D}_t \circ \mathrm{E} + \boldsymbol{\nabla} \cdot (\boldsymbol{u} \, \mathrm{PE}) \, \mathrm{d}t - \mathrm{PE} \, \boldsymbol{\nabla} \cdot \boldsymbol{u}_s \, \mathrm{d}t + \mathrm{E} \, \boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \, \mathrm{d}t = 0. \tag{4.20a}$$

Expanding the Stratonovich transport operator (4.17d), the previous equation can be re-written as

$$\mathrm{d}_t \circ \mathrm{E} + \boldsymbol{\nabla} \cdot \left(\boldsymbol{F}_1 \, \mathrm{d}t + \boldsymbol{F}_2 \circ \mathrm{d}\boldsymbol{B}_t \right) = \mathrm{PE} \, \boldsymbol{\nabla} \cdot \boldsymbol{u}_s \, \mathrm{d}t, \tag{4.20b}$$

where $\boldsymbol{F}_1 \triangleq (\boldsymbol{u} - \boldsymbol{u}_s) \, \mathrm{E} + \boldsymbol{u} \, \mathrm{PE}$ and $\boldsymbol{F}_2 \circ \mathrm{d}\boldsymbol{B}_t \triangleq \mathrm{E} \, \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t$ are the total energy flux due to the corrected large-scale drift $\boldsymbol{u} - \boldsymbol{u}_s$ and the noise component, respectively. The additional term $\mathrm{PE} \, \boldsymbol{\nabla} \cdot \boldsymbol{u}_s$ stands for sources or sinks of the potential energy due to the compressibility of the ISD. In particular, if we assume that the ISD is incompressible, *i.e.* $\boldsymbol{\nabla} \cdot \boldsymbol{u}_s = 0$, the evolution of the energy density reduces to

$$\mathrm{d}_t \circ \mathrm{E} + \boldsymbol{\nabla} \cdot \left(\boldsymbol{F}_1 \, \mathrm{d}t + \boldsymbol{F}_2 \circ \mathrm{d}\boldsymbol{B}_t \right) = 0. \tag{4.21a}$$

If the fluid domain has zero boundary conditions (e.g. the normal velocities vanish on each wall or there are no boundaries at all as on the sphere), then one can show

that the total energy, $\overline{\mathrm{E}} \triangleq \int_\Omega \mathrm{E}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}$, is invariant in time:

$$\mathrm{d}_t \circ \overline{\mathrm{E}} = \int_\Omega \mathrm{d}_t \circ \mathrm{E}\,\mathrm{d}\boldsymbol{x} = -\int_{\partial\Omega} (\boldsymbol{F}_1\,\mathrm{d}t + \boldsymbol{F}_2 \circ \mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{n}\,\mathrm{d}l = 0, \tag{4.21b}$$

where $\partial\Omega$ and $\boldsymbol{n}$ denote the domain's boundaries and the unit normal vector, respectively.

In sum, in this work we propose the following RSW–LU system that preserves the global energy of the large-scale flow in time for any realization of a random noise:

Conservation of momentum :

$$\mathbb{D}_t\boldsymbol{u} + \boldsymbol{f}\times\boldsymbol{u}\,\mathrm{d}t = -g\boldsymbol{\nabla}\eta\,\mathrm{d}t, \tag{4.22a}$$

Conservation of mass :

$$\mathbb{D}_t h + h\,\boldsymbol{\nabla}\cdot\boldsymbol{u}\,\mathrm{d}t = 0, \tag{4.22b}$$

Random geostrophic constraint :

$$\boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}p_t^\sigma, \tag{4.22c}$$

Incompressible constraints :

$$\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0, \tag{4.22d}$$

Conservation of energy :

$$\mathrm{d}_t\int_\Omega \frac{\rho}{2}\big(h|\boldsymbol{u}|^2 + gh^2\big)\,\mathrm{d}\boldsymbol{x} = 0. \tag{4.22e}$$

Note that for a sufficiently small noise ($\boldsymbol{\sigma}\approx 0$), this system (4.22) reduces to the classical RSW system, in which the stochastic transport operator weighted by the unit of time, $\mathbb{D}_t/\mathrm{d}t$, reduces to the material derivative.

4.3. Discretization of RSW–LU and parametrization of noise

In order to perform a numerical simulation of the RSW–LU (4.22), the noise term $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and the variance tensor $\boldsymbol{a}$ have to be *a priori* parametrized. Then an adequate discretization in space-time have to be specified for solving the dynamical system. This section describes these two aspects.

4.3.1. Parameterizations of noise

In the following, we present two different kinds of spatial structure for the noise – homogeneous and heterogeneous. The first one is easy-to-implement, in particular when considering noise that respects the incompressible constraints (4.22d). We use such homogeneous noise to study the numerical energy behaviour, as shown in Section 4.4.1. On the other hand, because heterogeneous noise has more physical meaning, we will use the latter when studying realistic complex flows. As shown in [10], heterogeneous noise induces a structuration of the large-scale flow through the inhomogeneity of the small-scale flow. In Section 4.4.2, such heterogeneous noise is adopted for identifying the barotropic instability of a mid-latitude jet.

Homogeneous noise

From the general definitions (4.2) and (4.4), a homogeneous noise means that its correlation operator $\boldsymbol{\sigma}$ is a convolution operator and the variance tensor $\boldsymbol{a}$ reduces to a constant matrix (independent of any position in the fluid domain). Furthermore, to ensure that a two-dimensional noise is incompressible, [70] proposed an isotropic model defined through a random stream function

$$\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^{\perp}\Big(\breve{\varphi} \star \mathrm{d}B_t\Big)(\boldsymbol{x}), \tag{4.23}$$

where $\boldsymbol{\nabla}^{\perp} = [-\partial_y, \partial_x]^T$ denotes the perpendicular gradient and $\breve{\varphi} \star \mathrm{d}B_t$ stands for the random stream function with a convolution kernel $\breve{\varphi}$ (and the symbol $\star$ denotes a convolution). As shown in [70, 72], both isotropy and incompressibility of the noise (4.23) result in a (constant) diagonal variance tensor $a_0\mathbf{I}_2$ with the eddy-viscosity-like coefficient a_0 and the two-dimensional identity matrix $\mathbf{I}_2$. In fact, the divergence-free constraint of the ISD in Equation (4.22d) is naturally satisfied (since $\boldsymbol{\nabla}\cdot\boldsymbol{u_s} = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(a_0\mathbf{I}_2) = 0$). As discussed at the end of Section 4.2.2, for the RSW–LU system (4.22) under *geostrophic noise*, $\boldsymbol{f} \times \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t \approx -\boldsymbol{\nabla}_H \mathrm{d}p_t^{\sigma}$, one can identify, for a constant Coriolis parameter f_0, the random pressure $\mathrm{d}p_t^{\sigma}$ with the proposed random stream function by $\mathrm{d}p_t^{\sigma} = \frac{1}{f_0}\breve{\varphi} \star \mathrm{d}B_t$.

In practice, the convolution kernel $\breve{\varphi}$ is specified by three parameters: a fixed omni-directional spectrum slope s, a band-pass filter f_{BP} with support in the range of two wavenumbers κ_m and κ_M, and an eddy-viscosity-like coefficient a_0. In fact, the Fourier transform of the random stream function $\breve{\varphi} \star \mathrm{d}B_t$ can be defined as:

$$\widehat{\breve{\varphi} \star \mathrm{d}B_t}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}} f_{\mathrm{BP}}\left(\|\boldsymbol{k}\|\right) \|\boldsymbol{k}\|^{-\alpha} \widehat{\xi}_t(\boldsymbol{k}) \text{ with } \alpha = (3+s)/2, \tag{4.24}$$

where $\widehat{\bullet}$ denotes the Fourier transform coefficient, ξ_t is a space-time white noise, and A is a constant to ensure $\mathbb{E}\left\|\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right\|^2 = 2a_0\Delta t$ (see Equations (4.3) and (4.4)) with Δt the size of one time stepping. In the simulations, the maximal wavenumber k_M of the noise can usually be chosen as the effective resolution cutoff, the minimal wavenumber can be set to $k_m = k_M/2$, and the theoretical spectrum slope of a two-dimensional flow is given by $s = -3$. The noise strength parameter a_0 will be specified in Section 4.4.1.

Heterogeneous noise

The homogeneous noise defined in Section 4.3.1 is quite simple to construct and to interpret, however, it fails to represent physically important contributions of the

small-scale to the large scale flow, which is crucial in order to accurately model realistic scenarios in geophysical fluid dynamics. For this reason, two parameterizations of the heterogeneous noise are presented in the following.

These approaches result from the spectral decomposition (4.5) used to construct the eigenfunction basis of the spatial covariance. In practice, we work with a finite set of Empirical Orthogonal Functions (EOFs) of the small-scale Eulerian velocity rather than with the Lagrangian displacement. The first method for estimating the EOFs is an off-line procedure based on the Proper Orthogonal Decomposition (POD) technique of high-dimensional data in which the EOFs are assumed to be time-independent, whereas the second one is an on-line estimation from a coarse-grid simulation where the EOFs are time-dependent. As will be shown in Section 4.4.2, the former allows for incorporating data into the dynamical model and is more suitable for mid-term simulations, yet the latter is independent from observations and is more adequate for long-term simulations.

Off-line learning of EOFs Let us consider a set of velocity snapshots $\{\boldsymbol{u}_{\mathrm{o}}(\boldsymbol{x}, t_i)\}_{i=1,\dots,N_t}$, that have been *a priori* coarse-grained from high-dimensional data using a low-pass filter (such as the sharp spectral filter of [67] often used in large eddy simulations). Applying the snapshot POD procedure [78] for the fluctuations $\boldsymbol{u}'_{\mathrm{o}} = \boldsymbol{u}_{\mathrm{o}} - \overline{\boldsymbol{u}}_{\mathrm{o}}$ (where $\overline{\bullet}$ denotes a temporal average) enables us to build a set of EOFs $\{\boldsymbol{\phi}_i\}_{i=1,\dots,N_t}$. In addition, we suppose that the fluctuations of the large-scale flow live in a subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=1,\dots,m-1}$ (with $m < N_t$) and that the small-scale random drift $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\Delta t$ lives in the complemented subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=m,\dots,N_t}$ such that

$$\frac{1}{\Delta t}\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \sum_{i=m}^{N_t}\sqrt{\lambda_i}\boldsymbol{\phi}_i(\boldsymbol{x})\xi_i, \qquad \frac{1}{\Delta t}\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=m}^{N_t}\lambda_i\boldsymbol{\phi}_i(\boldsymbol{x})\boldsymbol{\phi}_i^T(\boldsymbol{x}), \tag{4.25}$$

where λ_i is the eigenvalue associated to the spatial mode $\boldsymbol{\phi}_i$ and ξ_i is a standard Gaussian variable. In practice, there exists an open question in (4.25), that is how to adequately choose the "splitting mode" $\boldsymbol{\phi}_m$. Recently, [11] proposed to fix it by comparing the time-averaged energy spectrum of the observations and the one from a coarse-grid deterministic simulation.

On-line learning of EOFs The previously described data-driven estimation of EOFs is a quite efficient procedure. However, such observation data, either from direct measurements or from high-dimensional simulations, are not always available. Therefore, [10, 72] proposed an alternative approach in which some local fluctuations, called *pseudo-observations* (PSO), are generated directly from a coarse-grid simulation. Then, the singular value decomposition (SVD) is applied on those PSO to estimate a set of EOFs such that the noise associated with its variance tensor will be built in the same way as in (4.25). Finally, the magnitude of the noise and variance should be scaled down to smaller scales based on a similarity analysis [49].

In the following, we describe in more details both the generation of PSO and the scaling technique. The approach proposed here defines N_o PSO (denoted as $\boldsymbol{u}'$) at each grid point. For a given time t and a current coarse velocity $\boldsymbol{u}$, we build the PSO by sliding a local window of size $N_w \times N_w$ over the spatial grid (with N_w the grid number in one direction of the local window). We denote the spatial scale of the window by $L = N_w l$, where l is the smallest scale of the simulation. At every grid point $\boldsymbol{x}_{i,j}$, we list the N_w^2 velocity values contained in the window centered at that point:

$$I(\boldsymbol{x}_{i,j}, t) \triangleq \left\{ \boldsymbol{u}(\boldsymbol{x}_{p,q}, t) \middle| |p-i| \leq \frac{N_w - 1}{2}, |q-j| \leq \frac{N_w - 1}{2} \right\}. \tag{4.26}$$

Note that appropriate boundary conditions (replication, periodicity, etc.) are

adopted when looking at a point on the border. Then, independently for each $n \in \{1, \ldots, N_o\}$ and for each point $\boldsymbol{x}_{i,j}$, we set the value of the PSO $\boldsymbol{u}'(\boldsymbol{x}_{i,j}, t, n)$ by randomly choosing a value in the set $I(\boldsymbol{x}_{i,j}, t)$. After this, we average over the realization index n. Then, from the SVD we obtain a set of EOFs $\{\boldsymbol{\phi}_i^{(L)}\}_{i=1,\ldots,N_o}$, and a spectral representation of the small-scale velocity:

$$\frac{1}{\Delta t}\boldsymbol{\sigma}^{(L)}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{B}_t = \sum_{i=1}^{N_o} \boldsymbol{\phi}_i^{(L)}(\boldsymbol{x}, t)\xi_i. \tag{4.27a}$$

Since the PSO $\boldsymbol{u}'$ have been generated at a spatial scale of the window $L = N_w l$, they must be scaled down to the "simulation scale" l. As such, the variance tensor $\boldsymbol{a}$ of the small-scale flow is rescaled according to a turbulence-power-law coefficient[49] such that

$$\boldsymbol{a}^{(l)} = \left(\frac{l}{L}\right)^{2/3} \boldsymbol{a}^{(L)}, \tag{4.27b}$$

where $\boldsymbol{a}^{(L)}$ and $\boldsymbol{a}^{(l)}$ are the variance tensors at the scales L and l respectively. Finally, the small-scale flow can be simulated at the "simulation scale" l as

$$\boldsymbol{\sigma}^{(l)}\mathrm{d}\boldsymbol{B}_t = \left(\frac{l}{L}\right)^{1/3} \boldsymbol{\sigma}^{(L)}\mathrm{d}\boldsymbol{B}_t. \tag{4.27c}$$

As will be shown in Section 4.4.2, such flow-dependent noise has a good performance in long-term simulation, yet the drawback is that the computational costs are significantly higher compared to the previous off-line procedure, as the SVD is computed at each time step.

4.3.2. Structure-preserving discretization of RSW–LU

In this subsection, we introduce an energy conserving (in space) approximation of the above derived stochastic system. Considering the definition of the stochastic transport operator $\mathbb{D}_t$ defined in (4.6b) with a time increment $\mathrm{d}_t q \triangleq q_{t+\mathrm{d}t} - q_t$ defined in (4.6c), the RSW–LU system in Eqn. (4.22a)–(4.22b) can be explicitly

written as

$$\mathrm{d}_t\boldsymbol{u} = \Big(-\boldsymbol{u}\cdot\boldsymbol{\nabla}\,\boldsymbol{u} - \boldsymbol{f}\times\boldsymbol{u} - g\boldsymbol{\nabla}\eta\Big)\,\mathrm{d}t + \Big(\frac{1}{2}\,\boldsymbol{\nabla}\cdot\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u})\,\mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{u}\Big), \tag{4.28a}$$

$$\mathrm{d}_t h = -\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{u}h)\,\mathrm{d}t + \Big(\frac{1}{2}\,\boldsymbol{\nabla}\cdot\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}h)\,\mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,h\Big). \tag{4.28b}$$

We suggest to develop an approximation of the stochastic RSW–LU model (4.28a)–(4.28b) by first discretizing the deterministic model underlying this system with a structure-preserving discretization method (that preserves energy in space) and, then, to approximate (with a potentially different discretization method) the stochastic terms. Here, we use for the former a variational discretization approach on a triangular C–grid while for the latter we apply a standard finite difference method. The *deterministic dynamical core* of our stochastic system results from simply setting $\boldsymbol{\sigma} \approx 0$ in the equations (4.28a)–(4.28b). To obtain the full discretized (in space and time) scheme for this stochastic system, we wrap the discrete stochastic terms around the deterministic core and combine this with an Euler–Marayama time scheme.

Introducing discretizations of the stochastic terms that do not necessarily share the same operators as the deterministic scheme has the advantage that our method can be directly applied to existing dynamical cores of global numerical weather prediction (NWP) and climate models.

Structure-preserving discretization of the deterministic RSW equations

As mentioned above, the deterministic model (or deterministic dynamical core) of the above stochastic system results from setting $\boldsymbol{\sigma} \approx 0$, which leads via (4.4) to $\boldsymbol{a} \approx 0$. Hence, Equations (4.28a)–(4.28b) reduce to the deterministic RSW

equations

$$\mathrm{d}_t \boldsymbol{u} = \Big(-(\boldsymbol{\nabla} \times \boldsymbol{u} + \boldsymbol{f}) \times \boldsymbol{u} - \boldsymbol{\nabla}(\frac{1}{2}\boldsymbol{u}^2) - g\boldsymbol{\nabla}\eta \Big)\,\mathrm{d}t, \qquad \mathrm{d}_t h = -\,\boldsymbol{\nabla}\boldsymbol{\cdot}\,(\boldsymbol{u}h)\,\mathrm{d}t, \tag{4.29}$$

where we used the vector calculus identity $\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla}\, \boldsymbol{u} = (\boldsymbol{\nabla} \times \boldsymbol{u}) \times \boldsymbol{u} + \frac{1}{2}\boldsymbol{u}^2$. Note that in the deterministic case $\mathrm{d}_t/\mathrm{d}t$ agrees (in the limit $\mathrm{d}t \to 0$) with the partial derivative $\partial/\partial t$.

Variational discretizations. In the following we present an energy conserving (in space) approximation of these equations using a variational discretization approach. While details about the derivation can be found in [8, 20], here we only give the final, fully discrete scheme.

To do so, we start with introducing the mesh and some notation. The variational discretization of (4.29) results in a scheme that corresponds to a C-grid staggering of the variables on a quasi uniform triangular grid with hexagonal/pentagonal dual mesh. Let N denote the number of triangles used to discretize the domain. As shown in Fig. 5.1, we use the following notation: T denotes the primal triangle, ζ the dual hexagon/pentagon, $e_{ij} = T_i \cap T_j$ the primal edge and $\tilde{e}_{ij} = \zeta_+ \cap \zeta_-$ the associated dual edge. Furthermore, we have $\mathbf{n}_{ij}$ and $\mathbf{t}_{ij}$ as the normalized normal and tangential vector relative to edge e_{ij} at its midpoint. Moreover, D_i is the discrete water depth at the circumcentre of T_i, η_{b_i} the discrete bottom topography at the circumcentre of T_i, and $V_{ij} = (\mathbf{u} \cdot \mathbf{n})_{ij}$ the normal velocity at the triangle edge midpoints in the direction from triangle T_i to T_j. We denote $\overline{D}_{ij} = \frac{1}{2}(D_i + D_j)$ as the water depth averaged to the edge midpoints.

The variational discretization method does not require to define explicitly approximations of the differential operators because they directly result from the discrete variational principle. It turns out that on the given mesh, these operators agree with the following definitions of standard finite difference and finite volume

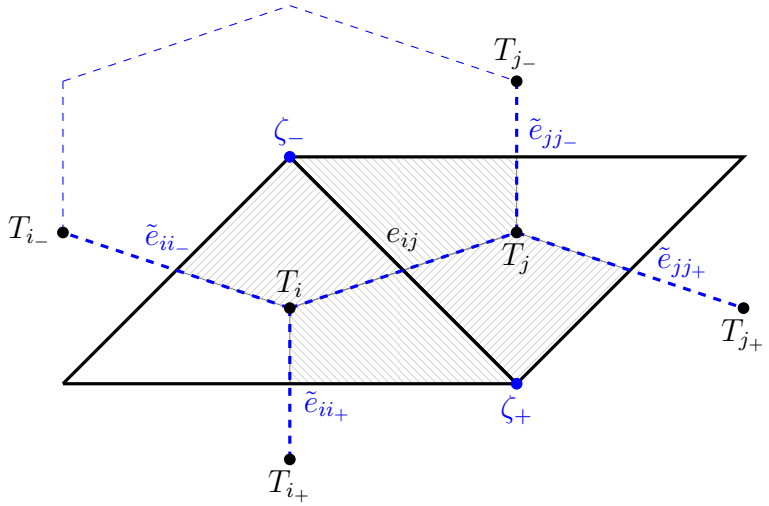


Figure 4.2.: Notation and indexing conventions for the 2D simplicial mesh.

operators:

$$\begin{aligned} (\text{Grad}_n\ F)_{ij} &\triangleq \frac{F_{T_j} - F_{T_i}}{|\tilde{e}_{ij}|}, & (\text{Div}\ V)_i &\triangleq \frac{1}{|T_i|} \sum_{k \in \{j, i_-, i_+\}} |e_{ik}| V_{ik}, \\ (\text{Grad}_t\ F)_{ij} &\triangleq \frac{F_{\zeta_-} - F_{\zeta_+}}{|e_{ij}|}, & (\text{Curl}\ V)_\zeta &\triangleq \frac{1}{|\zeta|} \sum_{\tilde{e}_{nm} \in \partial\zeta} |\tilde{e}_{nm}| V_{nm}, \end{aligned} \tag{4.30}$$

for the normal velocity V_{ij} and a scalar function F either sampled as F_{T_i} at the circumcentre of the triangle T_i or sampled as $F_{\zeta_\pm}$ at the centre of the dual cell $\zeta_\pm$. The operators Grad_n and Grad_t correspond to the gradient in the normal and tangential direction, respectively, and Div to the divergence of a vector field:

$$(\nabla F)_{ij} \approx (\text{Grad}_n\ F)\mathbf{n}_{ij} + (\text{Grad}_t\ F)\mathbf{t}_{ij}, \tag{4.31}$$

$$(\nabla \cdot \mathbf{u})_i \approx (\text{Div}\ V)_i, \tag{4.32}$$

$$(\nabla \times \mathbf{u})_\zeta \approx (\text{Curl}\ V)_\zeta. \tag{4.33}$$

The last Equation (4.33) defines the discrete vorticity and for later use, we also

discretize the potential vorticity as

$$\frac{\nabla \times \mathbf{u} + f}{h} \approx \frac{(\text{Curl } V)_\zeta + f_\zeta}{D_\zeta}, \qquad D_\zeta = \sum_{\tilde{e}_{ij} \in \partial\zeta} \frac{|\zeta \cap T_i|}{|\zeta|} D_i. \tag{4.34}$$

Semi-discrete RSW scheme. With the above notation, the deterministic semi-discrete RSW equations read:

$$\mathrm{d}_t V_{ij} = \mathcal{L}^V_{ij}(V, D)\,\Delta t, \quad \text{for all edges } e_{ij}, \tag{4.35a}$$

$$\mathrm{d}_t D_i = \mathcal{L}^D_i(V, D)\,\Delta t, \quad \text{for all cells } T_i, \tag{4.35b}$$

where $\mathcal{L}^V_{ij}$ and $\mathcal{L}^D_i$ denote the deterministic spatial operators, and Δt stands for the discrete time step. The RHS of the momentum equation (4.35a) is given by

$$\mathcal{L}^V_{ij}(V, D) \triangleq -\mathrm{Adv}(V, D)_{ij} - \mathrm{K}(V)_{ij} - \mathrm{G}(D)_{ij}, \tag{4.36}$$

where Adv denotes the discretization of the advection term $(\boldsymbol{\nabla} \times \boldsymbol{u} + \boldsymbol{f}) \times \boldsymbol{u}$ of (4.29), K the approximation of the gradient of the kinetic energy $\boldsymbol{\nabla}(\frac{1}{2}\boldsymbol{u}^2)$ and G of the gradient of the height field $g\boldsymbol{\nabla}\eta$. Explicitly, the advection term is given by

$$\begin{aligned}
&\mathrm{Adv}(V, D)_{ij} \triangleq \\
&-\frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\Big((\text{Curl } V)_{\zeta_-} + f_{\zeta_-}\Big)\left(\frac{|\zeta_- \cap T_i|}{2|T_i|}\overline{D}_{ji_-}|e_{ii_-}|V_{ii_-} + \frac{|\zeta_- \cap T_j|}{2|T_j|}\overline{D}_{ij_-}|e_{jj_-}|V_{jj_-}\right) \\
&+\frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\Big((\text{Curl } V)_{\zeta_+} + f_{\zeta_+}\Big)\left(\frac{|\zeta_+ \cap T_i|}{2|T_i|}\overline{D}_{ji_+}|e_{ii_+}|V_{ii_+} + \frac{|\zeta_+ \cap T_j|}{2|T_j|}\overline{D}_{ij_+}|e_{jj_+}|V_{jj_+}\right),
\end{aligned} \tag{4.37}$$

where $f_{\zeta_\pm}$ is the Coriolis term evaluated at the centre of $\zeta_\pm$. Moreover, the two gradient terms read:

$$\mathrm{K}(V)_{ij} \triangleq \frac{1}{2}(\mathrm{Grad}_n\ F)_{ij}, \qquad F_{T_i} = \sum_{k\in\{j,i_-,i_+\}} \frac{|\tilde{e}_{ik}|\ |e_{ik}|(V_{ik})^2}{2|T_k|}, \tag{4.38}$$

$$\mathrm{G}(D)_{ij} \triangleq g(\mathrm{Grad}_n\ (D+\eta_b))_{ij}. \tag{4.39}$$

The RHS of the continuity equation (4.35b) is given by

$$\mathcal{L}_i^D(V,D) \triangleq -\Big(\mathrm{Div}\ (\overline{D}V)\Big)_i, \tag{4.40}$$

which approximates the divergence term $-\boldsymbol{\nabla\cdot}(\boldsymbol{u}h)$.

Stabilization. In addition, as often used in the simulations of large-scale atmospheric and oceanic flows, in order to stabilize the numerical solution (which will be important for the stochastic model), we include a biharmonic eddy viscosity with uniform coefficient μ (of unit m^4/s) in the momentum equation:

$$\mathrm{d}_t V = \Big(-\mathrm{Adv}(V,D)_{ij} - \mathrm{K}(V)_{ij} - \mathrm{G}(D)_{ij} - \mu L(V)_{ij} \Big)\,\Delta t, \tag{4.41}$$

where:

$$L(V)_{ij} = \Big(\mathrm{Grad}_n(\mathrm{Div}\ V)_{ij} - \mathrm{Grad}_t(\mathrm{Curl}\ V)_{ij}\Big)^2. \tag{4.42}$$

Time scheme. For the time integrator we use a Crank-Nicolson-type scheme where we solve the system of fully discretized non-linear momentum and continuity equations by a fixed-point iterative method. The corresponding algorithm coincides for $\boldsymbol{\sigma} = 0$ and $\mu = 0$ with the one given in Section 4.3.2.

Spatial discretization of RSW–LU

The fully stochastic system has additional terms on the RHS of Equations (4.28a) and (4.28b). With these terms the discrete equations read:

$$\mathrm{d}_t V_{ij} = \mathcal{L}^V_{ij}(V, D)\,\Delta t + \Delta\mathcal{G}^V_{ij}, \tag{4.43a}$$

$$\mathrm{d}_t D_i = \mathcal{L}^D_i(V, D)\,\Delta t + \Delta\mathcal{G}^D_i, \tag{4.43b}$$

where the stochastic LU-terms are given by

$$\Delta\mathcal{G}^V_{ij} \triangleq \Big(-\frac{\Delta t}{2}\Big(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u})\Big)_{ij} + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{u})_{ij}\Big)\cdot\boldsymbol{n}_{ij}, \tag{4.43c}$$

$$\Delta\mathcal{G}^D_i \triangleq -\frac{\Delta t}{2}\Big(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}D)\Big)_i + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,D)_i. \tag{4.43d}$$

Note that the two terms within the large bracket in (4.43c) comprise two Cartesian components of a vector which is then projected onto the triangle edge's normal direction via $\boldsymbol{n}_{ij}$. The two terms in (4.43d) are scalar valued at the cell circumcenters i.

The parametrization of the noise described in Section 4.3.1 is formulated in Cartesian coordinates, because this allows using standard algorithms to calculate e.g. EOFs and POD. Likewise, we represent the stochastic LU-terms in Cartesian coordinates but to connect both deterministic and stochastic terms, we will calculate the occurring differentials with operators as provided by the deterministic dynamical core (see interface description below). Therefore, we write the second term in (4.43c) as

$$(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,F)_{ij} = \sum_{l=1}^{2}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^l_{ij}(\nabla F)^l_{ij}, \tag{4.44}$$

in which $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_{ij}$ denotes the discrete noise vector with two Cartesian components, constructed as described in Section 4.3.1 and evaluated at the edge midpoint ij. The scalar function F is a placeholder for the Cartesian components of the velocity field $\boldsymbol{u} = (u^1, u^2)$. Likewise, the first term in (4.43c) can be written component-wise as

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\cdot(\boldsymbol{a}F))_{ij} = \sum_{k,l=1}^{2} \left(\partial_{x_k} \left(\partial_{x_l}(a_{kl}F)\right)_{ij}\right)_{ij}, \tag{4.45}$$

where a_{kl} denotes the matrix elements of the variance tensor which will be evaluated, similarly to the discrete noise vector, at the edge midpoints. For a concrete realization of the differentials on the RHS of both stochastic terms, we will use the gradient operator (4.31) as introduced next.

To calculate the terms in (4.43d) we also use the representations (4.44) and (4.45) for a scalar function $F = D$ describing the water depth. However, as our proposed procedure will result in terms at the edge midpoint ij, we have to average them to the cell centers i.

Interface between dynamical core and LU terms. As mentioned above, the construction of the noise is done on a Cartesian mesh while the discretization of the deterministic dynamical core, corresponding to a triangular C-grid staggering, predicts the values for velocity normal to the triangle edges and for water depth at the triangle centers. We propose to exchange information between the noise generation module and the dynamical core via the midpoints of the triangle edges where on such C-grid staggered discretizations the velocity values naturally reside.

Starting with a given predicted velocity vector with edge values V_{ij}, we first have to reconstruct the full velocity vector field from these normal values. We use the reconstruction of the vector field in the interior of each triangle proposed by

[66]:

$$\boldsymbol{u}_i = \frac{1}{|T_i|} \sum_{k=j,i_-,i_+} |e_{ik}|(\boldsymbol{x}^{e_{ik}} - \boldsymbol{x}^{T_i})V_{ik}, \tag{4.46}$$

where $\boldsymbol{x}^{e_{ik}}$ are the coordinates of the edge midpoint and $\boldsymbol{x}^{T_i}$ are the coordinates of the triangle circumcentre. By averaging values from neighboring triangles, we obtain the corresponding values at the edge midpoints or vertices (see [7] for details).

This reconstructed velocity vector field will be used to generate the noise as described in Section 4.3.1. After the noise has been constructed on the Cartesian mesh, we evaluate the discrete noise vector $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_{ij}$ and the discrete variance tensor $(\boldsymbol{a})_{ij}$ at the triangle edge midpoints. This information will then be used to calculate the LU noise terms in (4.43c) and (4.43d).

To calculate the derivatives in these stochastic terms, we use the normal and tangential gradient operators, i.e. the gradient operator of (4.31). To use it, we have to average values, e.g. the term $(a_{kl}F)$, to cell centers and vertices and the resulting differential will be an expression located at the edge midpoint. In more detail, we can represent the partial derivative in Cartesian coordinates by

$$(\partial_{x_l}F)_{ij} = (\mathrm{Grad}_n\ F)n^l_{ij} + (\mathrm{Grad}_t\ F)t^l_{ij}, \qquad l = 1, 2. \tag{4.47}$$

Concretely, to discretize (4.45), we first compute $(\partial_{x_l}(a_{kl}F))_{ij}$ using Equation (4.47). The subindex ij indicates that the resulting term is associated to the edge midpoint. To apply the second derivative in (4.45), i.e. $\left(\partial_{x_k}\left(\partial_{x_l}(a_{kl}F)\right)_{ij}\right)_{ij}$, we proceed analogously, i.e. we first average the terms describing the first derivative to cells and vertices and then apply once more Equation (4.47). We proceed similarly to represent the term ∇F in (4.44).

As mentioned above, the terms in (4.43d) are calculated similarly to (4.43c) with the only difference that the former have to be averaged to the cell centers

after the proposed discretization procedure has been applied.

Temporal discretization of RSW–LU

The iterated Crank-Nicolson method presented in [20] is adopted for the temporal discretization. Keeping the iterative solver and adding the LU terms results in an Euler–Maruyama scheme, which decrease the order of convergence of the deterministic iterative solver (see [51] for details). To enhance readability, we denote V^t as the array over all edges e_{ij} of the velocity V_{ij} and D^t as the array over all cells T_i of the water depth D_i at time t. The governing algorithm reads:

Time-stepping algorithm

1. Start loop over $k = 0$ with initial guess at t : $V^*_{k=0} = V^t$ and $(D^*_{k=0})_i = D^t_i + \Delta\mathcal{G}^D_{ij}(D^t)$. Besides, we compute $\Delta\mathcal{G}^V_{ij}(V^t)$.

2. Update water depth D^*_{k+1} and velocity V^*_{k+1} using explicit equation:

$$\begin{aligned}
\frac{D^*_{k+1} - D^t}{\Delta t} &= -\frac{\text{Div }(\overline{D^*_k} V^*_k) + \text{Div }(\overline{D^t} V^t)}{2} \\
\frac{V^*_{k+1} - V^t}{\Delta t} &= -\frac{\text{Adv}(V^*_k, D^*_{k+1}) + \text{Adv}(V^t, D^t)}{2} - \frac{\text{K}(V^*_k) + \text{K}(V^t)}{2} - \text{G}(D^*_{k+1}) \\
&\quad - \mu L(V^*_k) + \Delta\mathcal{G}^V_{ij}(V^t)
\end{aligned}$$

   and set $k + 1 = k$.

3. Stop loop if $\|V^*_{k+1} - V^*_k\| + \|D^*_{k+1} - D^*_k\| <$ tolerance.

For all simulations in this manuscript, we used a tolerance of 10^{-6} for simulations on the plane and 10^{-10} for simulation on the sphere.

This algorithm will be used in the next section to evolve the fluid flow in time.

4.4. Numerical results

In this section, we first study the energy behaviour of the numerical RSW–LU scheme from above for an inviscid test flow. Then, we show that for a viscous test flow, the stochastic model captures more accurately the referent structure of the large-scale flow when compared to the deterministic model under the same coarse resolution. In addition, we demonstrate that the proposed RSW–LU system provides a more reliable ensemble forecast with larger spread, compared to a classical random model based on the perturbations of initial condition (PIC).

4.4.1. Inviscid test case – energy analysis

This first test case consists of two co-rotating vortices on the f-plane without viscosity (i.e. $\mu = 0$). To illustrate the energy conservation of the spatial discretization of the RSW–LU system (4.22), we use the homogeneous stationary noise defined in Section 4.3.1 since the two incompressible constraints $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0$ and $\boldsymbol{\nabla}\cdot\,\boldsymbol{\nabla}\cdot\,\boldsymbol{a} = 0$ in (4.22d) are naturally satisfied. Then, no extra steps are required to satisfy the incompressible constraints.

Initial conditions

The simulation is performed on a rectangular double periodic domain $\Omega = [0, L_x]\times[0, L_y]$ with $L_x = 5000\,\mathrm{km}$ and $L_y = 4330\,\mathrm{km}$, which is discretized into $N = 32768$ triangles. The large-scale flow is assumed to be under a geostrophic regime at the initial state, *i.e.* $f\boldsymbol{k}\times\boldsymbol{u} = -g\boldsymbol{\nabla}h$. We use an initial height field elevation (as e.g. in [8]) of the form

$$h\big(x, y, t = 0\big) = H_0 - H'\bigg(\exp\Big(-\frac{{x'_1}^2 + {y'_1}^2}{2}\Big) + \exp\Big(-\frac{{x'_2}^2 + {y'_2}^2}{2}\Big) - \frac{4\pi s_x s_y}{L_x L_y}\bigg), \tag{4.48a}$$

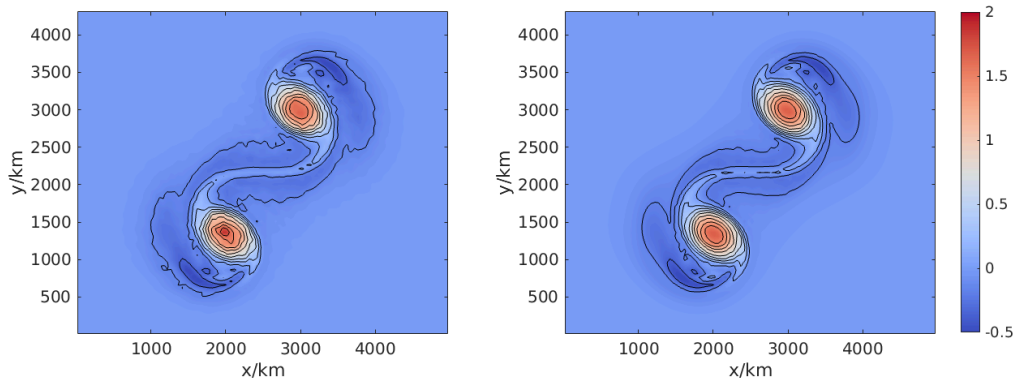


Figure 4.3.: Contour plots of the potential vorticity fields after 2 days for (left) one realization of a LU simulation with homogeneous noise and (right) a deterministic run. The contour interval is 0.4 days^{-1} km^{-1}.

where the background height H_0 is set to 10 km, the magnitude of the small perturbed height H' is set to 75 m and the periodic extensions are given by

$$x_i' = \frac{L_x}{\pi s_x} \sin\Big(\frac{\pi}{L_x}(x - x_{c_i})\Big), \qquad y_i' = \frac{L_y}{\pi s_y} \sin\Big(\frac{\pi}{L_y}(y - y_{c_i})\Big), \qquad i = 1, 2 \tag{4.48b}$$

with the centres of the vertices located at $(x_{c_1}, y_{c_1}) = \frac{2}{5}(L_x, L_y)$, $(x_{c_2}, y_{c_2}) = \frac{3}{5}(L_x, L_y)$ with parameters $(s_x, s_y) = \frac{3}{40}(L_x, L_y)$. To obtain the discrete initial water depth D_i, we sample the analytical function h at each cell centre. Subsequently, the discrete geostrophic velocities at each triangle edge ij at the initial state can be deduced via

$$V_{ij} = -\frac{g}{f}(\text{Grad}_t\ D)_{ij}, \tag{4.49}$$

where the Coriolis parameter f is set to 5.3108 days^{-1}. For the LU simulations, the magnitude of the homogeneous noise remains moderate with its constant variance a_0 set to be 169.1401 $\text{m}^2 \cdot \text{s}^{-1}$. This variance was chosen empirically, such that it is moderate.

Analysis of energy conservation

To analyze the energy conservation properties of our stochastic integrator, we use the above initial conditions to simulate the two co-rotating vortices for 2 days. In Figure 4.3, we show contour plots of the potential vorticity (as defined in (4.34)) fields of the deterministic and stochastic models. We observe that under the moderate noise with a_0 as chosen above, the large-scale structure of the stochastic system is similar to that of the deterministic run.

On the specific staggered grid as shown in Figure 5.1, the total energy of the shallow water equations (4.16) for both deterministic and stochastic case is approximated by

$$\mathrm{E}(t) \approx \sum_{i=1}^{N} \frac{1}{2} D_i(t)|T_i| \sum_{k=j,i_-,i_+} \frac{1}{2|T_i|} h_{ik} f_{ik} \Big(V_{ik}(t)\Big)^2 + \frac{1}{2} g \Big(D_i(t)\Big)^2 |T_i|. \tag{4.50}$$

As shown in [8], the proposed discrete variational integrator (see Section 4.3.2) together with an iterative Crank-Nicolson time stepping method exhibits a 1st order convergence rate of the energy error with smaller time step size. This will allows us immediately to simply include the stochastic terms to result in an Euler-Maruyama type time integrator for stochastic systems (cf. Section 4.3.2).

In the present work, we consider the energy behavior of the deterministic scheme (i.e. the variational integrator) as reference, which is denoted as $\mathrm{E}_{\mathrm{REF}}(t)$ in the following. For the stochastic RSW model, the Euler-Maruyama time scheme might lead to a different behavior with respect to energy conservation when compared to the deterministic model. In order to quantify numerically the energy conservation of the RSW–LU, we propose to measure the relative errors between the mean stochastic energy, denoted as $\overline{\mathrm{E}}_{\mathrm{LU}}(t)$, and the reference $\mathrm{E}_{\mathrm{REF}}(t)$ by $\overline{\mathrm{E}}_{\mathrm{LU}}(t)/\mathrm{E}_{\mathrm{REF}}(t) - 1$. This setup allows us to measure the influence of the stochastic terms on the energy conservation relative to the determinitic scheme. Figure 4.4a shows these relative errors for different time step sizes over a simulation time of 2 days. As we

can deduce from the curves, taking successively smaller time steps Δt results in smaller relative errors.

To determine more quantitatively the convergence rate of the stochastic scheme (relative to the reference) with respect to different time step sizes, we defined the following global (in space and time) error measure:

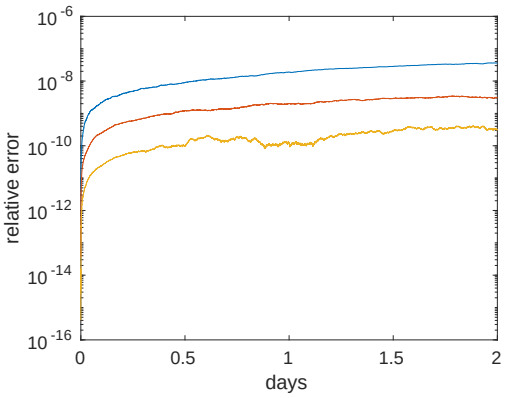
$$\varepsilon(\mathrm{E_{LU}}) \triangleq \frac{\|\mathrm{E_{LU}}(t) - \mathrm{E_{REF}}(t)\|_{L^2([0,T])}}{\|\mathrm{E_{REF}}(t)\|_{L^2([0,T])}}, \tag{4.51}$$

where $\|f(t)\|_{L^2([0,T])} = (\int_0^T |f(t)|^2 \mathrm{d}t)^{1/2}$ and T is set to 2 days. We determine for an ensemble with 10 members such global errors in order to illustrate the convergence rate of each ensemble member and the spread between those rates. This spread is illustrated as blue shaded area in Figure 4.4b. The area centre is determined by the mean of the errors, and the dispersion of this area is given by one standard derivation (*i.e.* 68% confident interval of the ensemble of $\varepsilon(\mathrm{E_{LU}})$). Besides, the minimal and maximal values of the errors of the ensemble are represented by the vertical bar-plots. The blue line of Figure 4.4b shows that the convergence rate (w.r.t. various Δt) of the ensemble mean energy is of 1st order. This is consistent with the weak convergence rate of order $\mathcal{O}(\Delta t)$ of the Euler-Maruyama scheme, cf. Section 4.3.2.

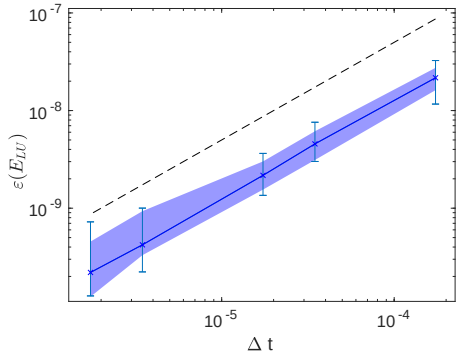
4.4.2. Viscous test case - ensemble prediction

Next, we want to show that our stochastic system better captures the structure of a large-scale flow than a comparable deterministic model. To this end, we use a viscous test case and heterogeneous noise.

The viscous test case we use is proposed by [34] and it consists of a barotropically unstable jet at the mid-latitude on the sphere. This strongly non-linear flow will be destabilized by a small perturbation of the initial field, which induces decaying turbulence after a few days. However, the development of the barotropic



(a) Evolution of the relative L_2 errors between the energy of the mean RSW–LU and the reference, using Δt (blue line), $\Delta t/10$ (red line) and $\Delta t/100$ (yellow line) respectively.



(b) Convergence of the energy path of the RSW–LU to that of the reference w.r.t. time step sizes. The blue line shows the global errors of the ensemble mean energy, the blue area describes the 68% confident interval of the ensemble errors and the dashed line stands for the 1st order convergence rate.

Figure 4.4.: Analysis of the numerical energy conservation of the RSW–LU.

instability in numerical simulations highly depends on accurately resolving the small-scale flow, which is particularly challenging for coarse-grid simulations. For the same reason, the performance of an ensemble forecast system in this test case is quite sensitive to the numerical resolution. In the following, we demonstrate that the RSW–LU simulation on a coarse mesh under heterogeneous noises, provides better prediction of the barotropic instability compared to the deterministic coarse simulation, and produces more reliable ensemble spread than the classical PIC simulation.

Initial conditions

The values of the principle parameters for the simulations are specified in Table 4.1. Under the geostrophic regime, the initial zonal velocity and height is respectively given by

$$u(\Theta, t=0) = \frac{U_0}{e_n} \exp\Big(\frac{1}{(\Theta-\Theta_0)(\Theta-\Theta_1)}\Big), \quad \text{for } \Theta_0 < \Theta < \Theta_1, \tag{4.52a}$$

$$h(\Theta, t=0) = H_0 - \frac{R}{g}\int_{\Theta} u(\theta, t=0)\Big(2\tilde{\Omega}\sin\theta + \frac{\tan\theta}{R}u(\theta, t=0)\Big)\,\mathrm{d}\theta, \tag{4.52b}$$

where $e_n = \exp\big(-4/(\Theta_1-\Theta_0)^2\big)$ is used to rescale the jet magnitude to the maximal value U_0 at the jet's mid-point $\Theta = \pi/4$. As introduced by [34], in order to initiate the barotropic instability, the following localized bump is included in the height field:

$$h'(\Upsilon, \Theta) = H' \cos\Theta \exp\Big(-(3\Upsilon)^2 - \Big(15(\frac{\pi}{4}-\Theta)\Big)^2\Big), \tag{4.52c}$$

where Υ denotes the longitude. Analogously to the previous inviscid test case, we then use these analytic functions (4.52) to sample the discrete velocity at the

edge mid-point and the height field at the cell centre on the staggered mesh (See Figure 5.1).

Parameters	Value	Description
(Θ_0, Θ_1)	$(2\pi, 5\pi)/14\,\mathrm{rad}$	Initial latitude limits
H_0	$10.158\,\mathrm{km}$	Background height
H'	$120\,\mathrm{m}$	Initial perturbation amplitude
R	$6.371 \times 10^3\,\mathrm{km}$	Mean radius of Earth
g	$9.806\,\mathrm{m}\cdot\mathrm{s}^{-2}$	Gravity of Earth
$\tilde{\Omega}$	$7.292 \times 10^{-5}\,\mathrm{s}^{-1}$	Angular rotation rate of Earth
U_0	$80\,\mathrm{m}\cdot\mathrm{s}^{-1}$	Maximum zonal velocity
μ_l	$3.975 \times 10^{14}\,\mathrm{m}^4\cdot\mathrm{s}^{-1}$	Fine-grid biharmonic viscosity
μ_L	$3.199 \times 10^{16}\,\mathrm{m}^4\cdot\mathrm{s}^{-1}$	Coarse-grid biharmonic viscosity
Δt_l	$12\,\mathrm{s}$	Fine-grid time step
Δt_L	$50\,\mathrm{s}$	Coarse-grid time step
N_l	327680	Number of triangles for fine grid
N_L	20480	Number of triangles for coarse grid

Table 4.1.: Parameter list for simulations of the barotropic instability.

For the LU simulations, we use the two heterogeneous noises described in Section 4.3.1, based on either the off-line learning of EOFs from the high-resolution simulation data, denoted as LU off-line, or on the on-line estimation of EOFs from the coarse-grid simulation, denoted as *LU on-line*. To allow for comparisons, the strength of these two noises are imposed to be the same.

The PIC stochastic model is obtained as follows: first, we perform ensemble simulations of the LU off-line and the LU on-line method over 1 day. Then, each ensemble realization is used as one initial random state for the *PIC off-line* and the *PIC on-line* simulations, respectively. For each stochastic model, an ensemble run with 20 realizations is done. Besides, a deterministic coarse-grid simulation, denoted as *LR*, is also performed. For all these coarse models, the biharmonic viscosity coefficient is fixed to be the same as given in Table 4.1.

Prediction of barotropic instability

In this section, we compare the predictions of the barotropic instability for different coarse models to that provided by the reference simulation. The latter is obtained from the coarse-graining procedure through a bilinear interpolation of the high-resolution snapshots. In Figure 4.5, we illustrate snapshots of the vorticity fields on the sphere for the reference, LU and deterministic models after a simulation time of 5 days. We can clearly see that at that day the LU ensemble mean better captures the large-scale structure of the reference flow than the deterministic simulation. To better distinguish the differences in the simulations, contour plots of the vorticity fields at day 4, 5 and 6, localized at the mid-latitude of the sphere, are given in Figure 4.6. From the evolution of the reference vorticity fields, we observe that the barotropic instability of the mid-latitude jet starts to develop at day 4. Subsequently, more and more small-scale features emerge and the flow becomes turbulent. Furthermore, both LU on-line and LU off-line simulations exhibit the stretched out wave at day 5 in the same way as the reference does, and that some big vortices start to separate from the wave at day 6. On the other hand, these characteristics are not correctly captured in both PIC off-line and LR simulations. We remark that the results of PIC on-line simulations are not include in Figure 4.6, since they behave quite similarly to the PIC off-line run.

To physically interpret the above results, it is useful to analyze the energy spectra of different models. From a basic knowledge of the two-dimensional turbulence theory [61], the potential enstrophy is transferred from the large scales to the small scales by the direct cascade, whereas the kinetic energy is transferred from the small scales to the large scales by the inverse cascade. However, introducing only a dissipation mechanism for coarse models often leads to an excessive decrease of the resolved kinetic energy [6, 50]. In our test case, this kind of issue is present in both PIC and the LR simulations, where the small-scale energy and

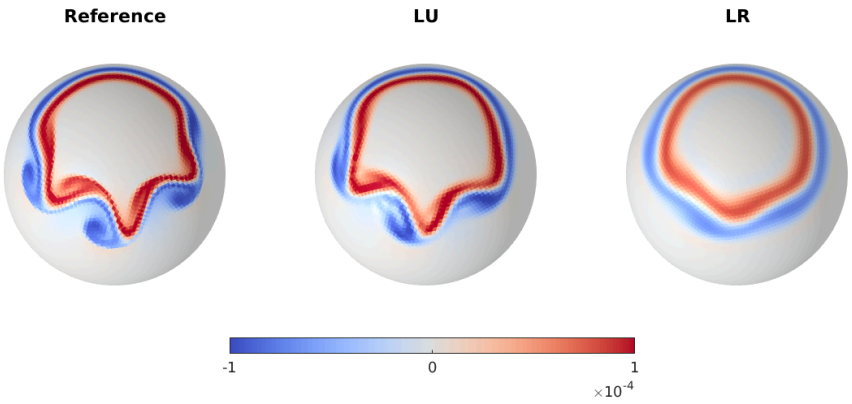


Figure 4.5.: Snapshots of the vorticity field on the sphere for different models (with 20480 triangles) after 5 days. From left to right: reference, ensemble mean of LU online and deterministic LR.

enstrophy are over-dissipated, as illustrated in Figure 4.7. On the other hand, introducing the non-linear convection by the noise, the LU dynamical systems bring higher turbulent energy and enstrophy to the small scales, which leads to better structuring of the large-scale flow. For instance, the ensemble mean of the energy and enstrophy spectra for both LU on-line and LU off-line simulations are much closer to that of the references at different days. Note that these spectra on the sphere are calculated using the method proposed by [1]: first, the energy and enstrophy is interpolated onto a Gaussian grid, then the spherical harmonics basis are used to compute the power spectral density.

Evaluation of ensemble forecasts

Once the ensembles have been produced by the random models, we measure the reliability of the ensemble forecast systems by some simple metrics. But before we do so, let us first demonstrate qualitatively the time evolution of each ensemble spread and compare it with the observation trajectory. To determine the latter, we evaluate the local vorticity field of the reference at different grid points in the region of the mid-latitude jet. These points serve as observation points. The

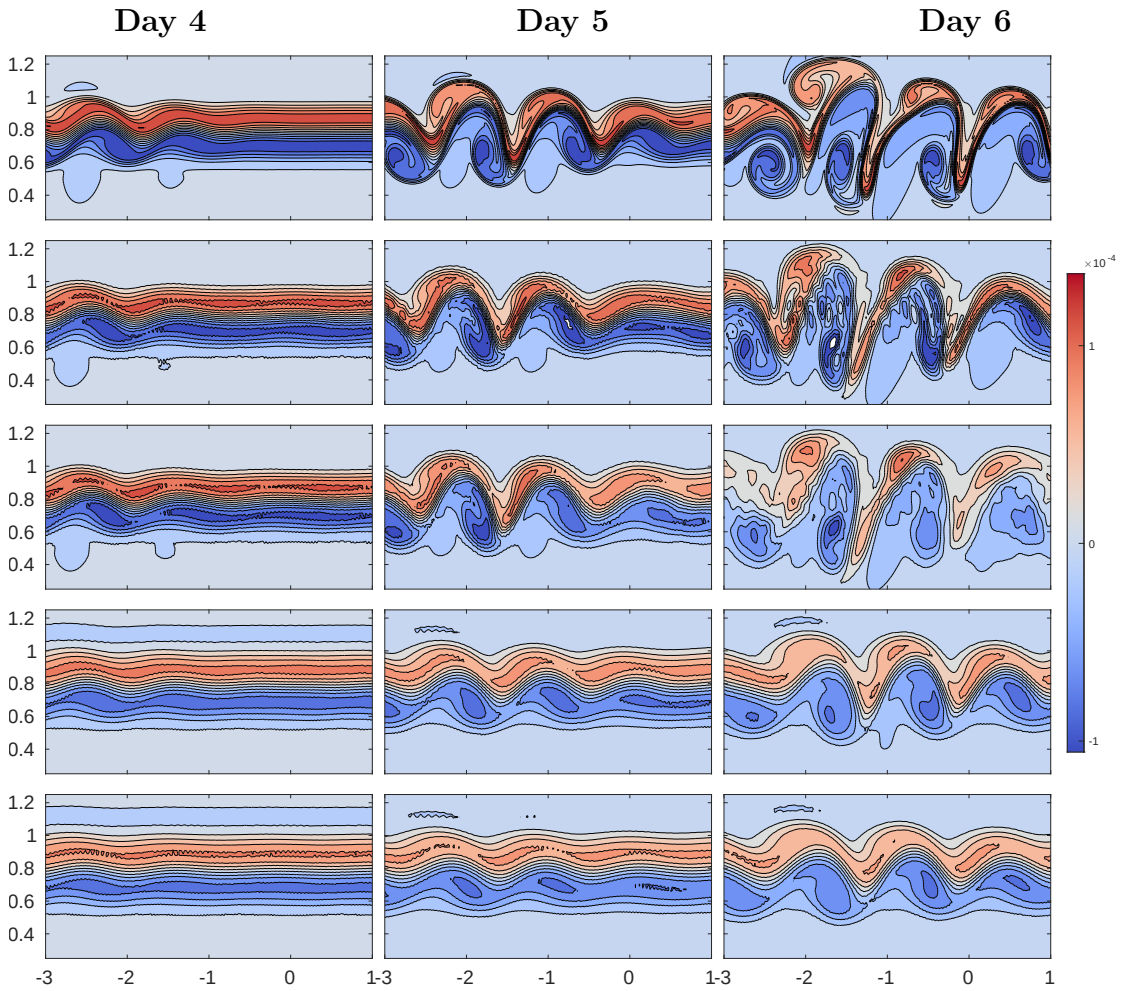


Figure 4.6.: Comparison of the vorticity contour plots along the mid-latitude jet for different models (with 20480 triangles) at day 4, 5 and 6 respectively. From top to bottom: reference, ensemble mean of LU on-line, ensemble mean of LU off-line, ensemble mean of PIC off-line and deterministic LR. The contour interval is fixed to $2 \times 10^{-5}\,\mathrm{s}^{-1}$, the x-axis is longitude (in rad) and the y-axis is latitude (in rad).

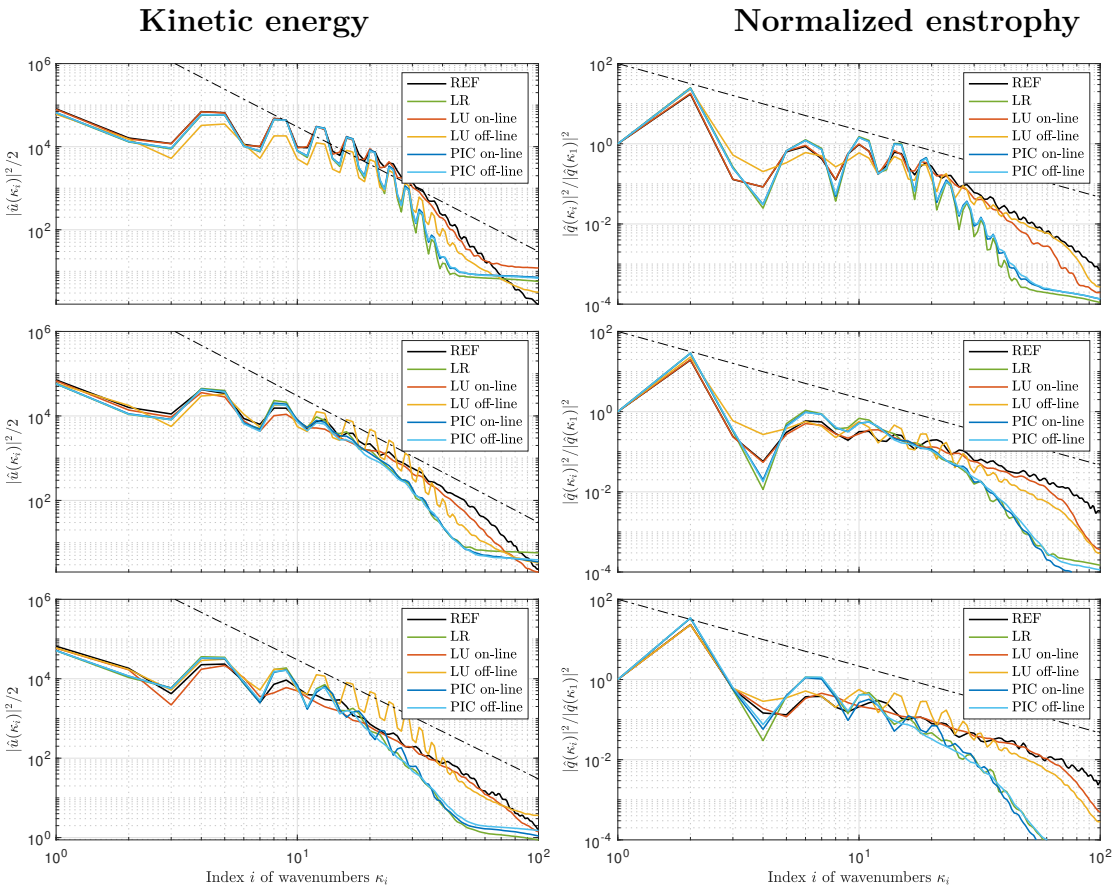


Figure 4.7.: Comparison of the ensemble mean of the kinetic energy (left column) spectrums and the potential enstrophy (right column) spectrums for different models (with 20480 triangles) at day 5 (1st row), 7 (2nd row) and 10 (3rd row) respectively. Note that the potential enstrophy is defined by the square of the potential vorticity and each potential enstrophy spectrum is normalized by its first value at the largest wavenumber. The dashed line is the k^{-3} (left column) and $k^{-5/3}$ (right column) power law.

evolution of the spread of the ensemble forecast systems is then build by the 95% confident interval of its ensemble trajectories at each selected point. As shown in Figure 4.8, for the six local points chosen along the longitude $\Upsilon = -1.53\,\mathrm{rad}$, the ensemble spreads of the LU off-line system are large enough to almost always include the observation trajectories, whereas the spreads of the PIC off-line system are quite small so that the observations are not always contained within the spread. For the latter, this will result in a wrong coupling of the measurement and the ensemble system, when performing data assimilation [32, 40].

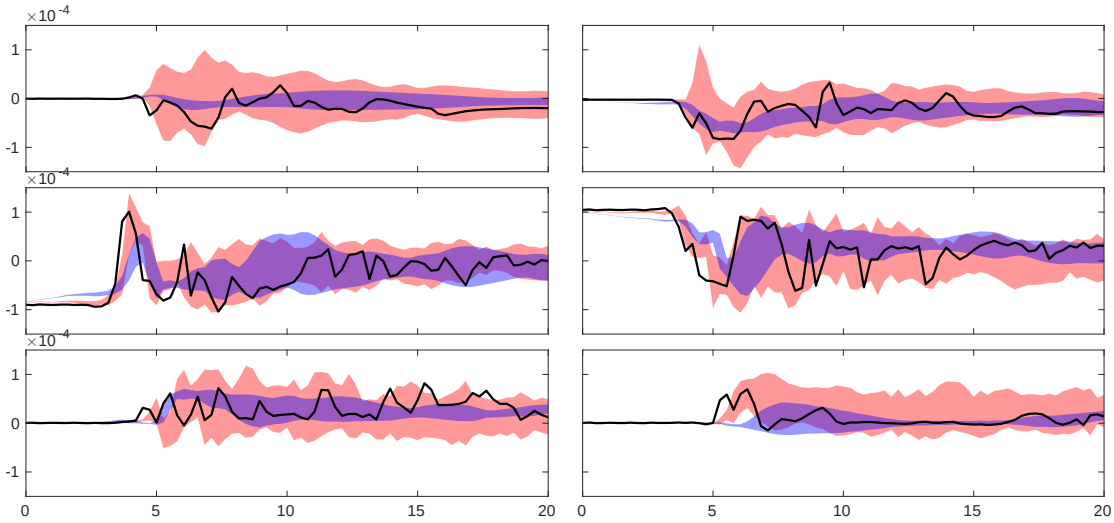


Figure 4.8.: Comparison of the ensemble spread evolution over 20 days of the vorticity field for the LU-offline (red area) runs and the PIC-offline (blue area) runs, at six different locations $\Theta = (0.4, 0.56, 0.72, 0.88, 1.04, 1.2)\,\mathrm{rad}$ along the longitude $\Upsilon = -1.53\,\mathrm{rad}$. The observation trajectories are shown by the black lines.

To quantify whether the ensemble spread of the forecast system represents the true uncertainty of the observations, the rank histogram [43, 80] is widely adopted as a diagnostic tool. This approach checks where the verifying observation usually falls w.r.t. the ensemble forecast states which are arranged in an increasing order at each grid point. In an ensemble with perfect spread, each member represents an equally likely scenario, so the observation is equally likely to fall between any two members. To construct the rank histogram in our test case, we proceed as

follows:

1. At every grid point $\boldsymbol{x}_i$, we rank the N_e vorticity values $\{q^{(j)}(\boldsymbol{x}_i)\}_{j=1,\ldots,N_e}$ of the ensemble from lowest to highest. This results in $N_e + 1$ possible bins which the observations can fall into, including the two extremes;

2. Identify which bin the observation vorticity $q^o(\boldsymbol{x}_i)$ falls into at each point $\boldsymbol{x}_i$;

3. Tally over all observations $\{q^o(\boldsymbol{x}_i)\}_{i=1,\ldots,N_o}$ to create a histogram of rank.

As shown in Figure 4.9, the histograms of both random models exhibit a U-shape for a few days in the beginning, while after a simulation time of about 10 days, the histograms of both LU on-line and LU off-line systems become mostly flat. A U-shape indicates that the ensemble spread is too small so that many observations are falling outside of the extremes of the ensemble while a dome-shape indicates the contrary. A flat histogram, in contrast, indicates that the ensemble members and observations are sampled from a common distribution. We observe that the LU off-line system performs slightly better than the LU on-line version. In contrast to these very good ensemble spreads, the histograms of both PIC on-line and PIC off-line systems remain in a U-shape during the entire simulation period which indicates that these systems do not accurately estimate the correct uncertainty around the observations.

It is important to notice that a flat rank histogram does not necessarily imply good forecasts, it only measures whether the observed probability distribution is well represented by the ensemble. To verify that a forecast is reliable, we need more criteria. One necessary criterion [89] for a reliable ensemble forecast is that the mean squared error (MSE) of the ensemble matches the mean intra-ensemble

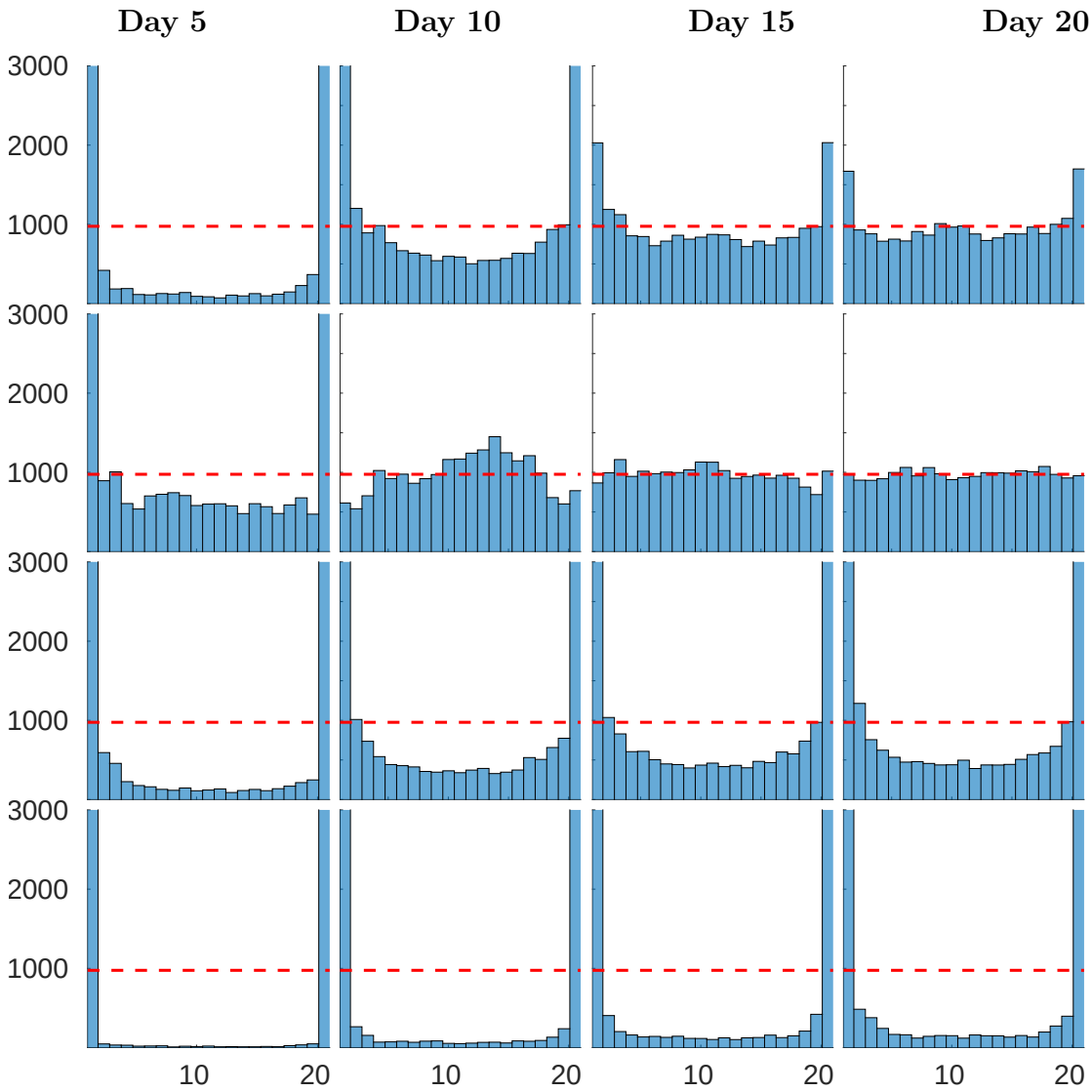


Figure 4.9.: Comparison of the rank histograms for the LU on-line (1st row) runs, the LU off-line (2nd row) runs, the PIC on-line (3rd row) runs and PIC off-line (last row) runs, at day 5, 10, 15 and 20 respectively.

variance (MEV), up to an ensemble size-dependent scaling factor, *i.e.*

$$\begin{aligned}\mathrm{MSE}\,(t) &= \frac{1}{N_o}\sum_{i=1}^{N_o}\Big(q^o - \widehat{\mathbb{E}}[q]\Big)^2(t,\boldsymbol{x}_i)\\ &\approx \Big(\frac{N_e+1}{N_e}\Big)\frac{1}{N_o}\sum_{i=1}^{N_o}\widehat{\mathrm{Var}}[q](t,\boldsymbol{x}_i) = \frac{N_e+1}{N_e}\,\mathrm{MEV}\,(t),\end{aligned} \tag{4.53}$$

where $\widehat{\mathbb{E}}[q] = \frac{1}{N_e}\sum_{j=1}^{N_e} q^{(j)}$ and $\widehat{\mathrm{Var}}[q] = \frac{1}{N_e-1}\sum_{j=1}^{N_e}\Big(q^{(j)} - \widehat{\mathbb{E}}[q]\Big)^2$ denote the empirical mean and the empirical variance, respectively.

In Figure 4.10, we compare the differences in time between the MSE and the MEV, normalized by the squared maximum of the initial vorticity, for the different random models from above. From these curves we can deduce that the LU off-line system exhibits the lowest errors during the entire simulation time of 20 days. In particular, during the first 10 days, these errors are significantly lower when compared to the other models, which can be explained by the fact that the LU off-line system incorporates data from the reference into the ensemble, which increases the reliability of the ensemble forecast. Although the errors between MSE and MEV of the LU on-line system is larger than the LU offline system from day 5 to day 10, they remain at low level from day 10 onwards, implying that the reliability of the former increases for longer simulation times. In contrast, both PIC off-line and PIC on-line systems show higher error values at most of the times and hence provide less reliable ensembles. We remark that other metrics, such as the continuous ranked probability score [72, 89], can also be used to measure a calibrated ensemble.

4.5. Conclusions

In this study, we introduced a stochastic version of the rotating shallow water equations under location uncertainty (RSW-LU). The derivation is based on a stochastic Reynolds transport theorem, where the fluid flow is decomposed into

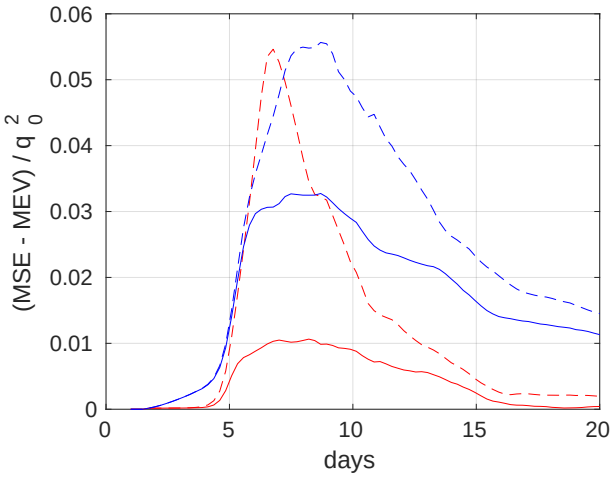


Figure 4.10.: Comparison of the differences between the mean square error (MSE) and the mean ensemble variance (MEV) of the ensemble vorticity fields for the LU on-line (red dashed line) runs, the LU off-line (red solid line) runs, the PIC on-line (blue dashed line) runs and the PIC off-line (blue solid line) runs. Note that these differences are normalized by $q_0 = \|q(\Upsilon, \Theta, t = 0)\|_\infty$.

a large-scale component and a noise term modelling the unresolved small-scale flow. A benefit of this approach is that the total energy is conserved along time for any realization. In order to preserve this structure, we combined an energy (in space) preserving discretization of the underlying deterministic equations of this RSW–LU system with approximations of the stochastic terms that are based on standard finite volume/difference operators.

We could show for an f-plane test case that this approach leads for homogeneous noise to a discretization of the RSW-LU system that preserves (spatially) the total energy. Moreover, using inhomogeneous noise that well captures the impact of small scales to the large-scale flow, we demonstrated that for a barotropically unstable jet on the sphere our proposed RSW–LU model better predicts the development of the instabilities than a comparable deterministic model, while the ensemble spread of the RSW–LU system is more likely to contain the observations compared to an ensemble of deterministic simulations with perturbed initial conditions (PIC). We also showed that the RSW–LU forecast systems follows a

common distribution of the observations and is more reliable than the PIC system.

Showing accurate ensemble spreads and reliable uncertainty quantification, we will next apply our developed RSW-LU system to data assimilation. We will also work towards discretizations of stochastic flow models in the framework of LU that preserve total energy both in space and time to which the present work provides a first step. Exploiting the modular approach of combining different discretizations for deterministic and stochastic terms, in future work we will explore the possibility to consistently extend existing atmospheric and ocean models with stochastic parametrizations.

Acknowledgments

The authors acknowledge the support of the Mitacs Globalink Research Award and of the ERC EU project 856408-STUOD. Besides, we would like to thank Alexander Bihlo and Scott MacLachlan for helpful discussions and thank Matthias Achtner for providing code to compute the energy spectrum on the sphere.

5. Selective decay for the rotating shallow-water equations with a structure-preserving discretization.

To be submitted to the journal Physics of Fluids

Rüdiger Brecht[†], Werner Bauer[‡], Alexander Bihlo[†], François Gay-Balmaz[§] and Scott MacLachlan[†]

[†] Department of Mathematics and Statistics, Memorial University of Newfoundland,
St. John's (NL) A1C 5S7, Canada

[‡] Imperial College London, Department of Mathematics, 180 Queen's Gate, London SW7 2AZ, United Kingdom.

[§] École Normale Supérieure/CNRS, Laboratoire de Météorologie Dynamique, Paris, France.

E-mail: rbrecht@mun.ca, w.bauer@imperial.ac.uk, abihlo@mun.ca, gaybalma@lmd.ens.fr, smaclachlan@mun.ca

Abstract

Numerical models of weather and climate critically depend on long-term stability of integrators for systems of hyperbolic conservation laws. While such stability is often obtained from dissipative terms (either physical or numerical), physical fidelity of such simulations also depends on properly preserving conserved quantities, such as energy, of the system. To address this apparent paradox, we develop a variational integrator for the shallow water equations that conserves energy, but dissipates potential enstrophy. Our approach follows the continuous selective decay framework [37], which enables dissipating an otherwise conserved quantity while conserving the total energy. We use this in combination with the variational discretization method [65] to obtain a *discrete selective decay* framework. This is applied to the shallow water equations, in both the plane and on the sphere, to dissipate the potential enstrophy. The resulting scheme significantly improves the quality of the approximate solutions, enabling long-term integrations to be carried out.

5.1. Introduction

Numerical weather and climate prediction require modelling of geophysical flows in the atmosphere and oceans on the globe. The atmosphere or oceans can be seen as thin layers of fluid above the surface of the Earth and, thus, the shallow water equations (SWE) are a useful simplified model of dynamic geophysical flows

around the Earth. These flows are approximately two dimensional, and we can get insight into their flow dynamics by studying the principles of two-dimensional turbulence. Important features of an incompressible turbulent flow are the cascades of enstrophy and energy, where the enstrophy transfers to higher wave numbers while the energy transfers to lower wave numbers, see e.g. [52, 56]. Numerical investigations of this phenomenon have led to the selective decay hypothesis [59], which states that the enstrophy accumulation at the grid-scale should be dissipated to improve stability, while energy should be conserved.

Following this perspective, energy conserving and enstrophy dissipating numerical schemes for the SWE and other equations have been developed, see, e.g., [4, 76]. However, many such methods directly manipulate the equations of motion to include the dissipation, which can have unpredictable consequences for the physical fidelity of the resulting numerical scheme. An alternative and more general method for developing energy-conserving dissipation schemes was introduced through the Lie-Poisson framework, see [37, 38]. While this approach appears to have great potential, it has not yet been applied to discrete models of geophysical fluid dynamics. In this paper, we aim to "connect the dots", leveraging the energy conservation from the Lie-Poisson framework via a structure-preserving discretization method, as structure-preserving integrators for differential equations generally guarantee long-term stability, consistency in statistical properties, and prevention of systematic drift in stationary or periodic solutions, see [42, 55, 88].

Here, we focus on variational integrators. These schemes are based on first discretizing the underlying variational principle and, then, deriving numerical schemes from the discrete Euler–Lagrange equations [58]. In [8], a variational discretization of the SWE was carried out. Then, in [20], the scheme was extended to the sphere, and it was observed that a stabilization of the scheme was needed to carry out long-term simulations. In this paper, we review the continuous selective decay theory to introduce a discretization of the selective decay that

mimics the continuous theory. We apply the new framework to obtain a discretization of the SWE that dissipates enstrophy and conserves energy. In particular, we extend the discrete SWE introduced in [8, 20] with the selective decay and carry out benchmarks on the plane and on the sphere.

This article is structured as follows. In Section 2, we review the continuous theory, to introduce the Casimir dissipation in Section 3. Then, Section 4 is devoted to a description of the discretization of the continuous Casimir dissipation. In Section 5, we verify the consistency of the discrete commutator and present results from numerical simulations. Conclusions are given in Section 6. Further, some detailed computations are presented in the appendix.

5.2. Euler–Poincaré equations

To obtain the selective decay in the numerical scheme we will use variational discretization, which mimics the continuous variational structure. On the continuous level, the equations of motion are obtained by defining a Lagrangian and computing the variational principle. This relies on the Euler–Poincaré reduction: the reformulation of Hamilton's principle from the Lagrangian to the Eulerian description. Thus, to understand the discretization procedure, we first review how we obtain the equations in the Euler–Poincaré framework.

The motion of a compressible fluid on a smooth manifold M is formally described by curves $\varphi : [0, T] \to \mathrm{Diff}(M)$ that are critical for the Hamilton principle

$$\delta \int_0^T L(\varphi, \dot{\varphi}) \, \mathrm{d}t = 0, \tag{5.1}$$

with respect to variations $\delta\varphi$ vanishing at $t \in [0, T]$. Here $\mathrm{Diff}(M)$ is the group of diffeomorphisms of the fluid domain M, and L is the Lagrangian of the fluid model expressed in terms of the Lagrangian fluid trajectory φ and Lagrangian fluid

velocity $\dot{\varphi}$. The variational principle (5.1) gives the equations in the Lagrangian description. For many computational approaches, it is more attractive to use a fixed Eulerian domain and, thus, Eulerian variables. Rewriting the principle in Eq. (5.1) in Eulerian variables yields the Euler–Poincaré variational principle which involves constrained variations, see [48] for a complete treatment. Here, we give a brief overview and refer to the appendix of [20] for a more detailed review for the case of the rotating shallow water equations on Riemannian manifolds.

We assume that M is endowed with a Riemannian metric and denote by $\mathrm{d}\sigma$ the associated Riemannian volume form. The examples treated in this paper will be a doubly periodic domain in $\mathbb{R}^2$ endowed with the Euclidean metric and a sphere endowed with its standard Riemannian metric; hence, we assume that M has no boundary. The Eulerian variables defined in terms of the Lagrangian fluid trajectory are the fluid velocity $\mathbf{u} = \dot{\varphi} \circ \varphi^{-1} \in \mathfrak{X}(M)$ (vector fields on M) and the fluid depth $h = (h_0 \circ \varphi^{-1}) J\varphi^{-1} \in \mathrm{Den}(M)$ (densities on M), where h_0 is the initial fluid depth and $J\varphi$ is the Jacobian of φ with respect to $\mathrm{d}\sigma$. The volume form allows the identification of the space of densities on M with the space of functions on M. From these relations, the Lagrangian $L(\varphi, \dot{\varphi})$ can be written in terms of $\mathbf{u}$ and h, which yields the reduced Lagrangian $\ell\colon \mathfrak{X}(M) \times \mathrm{Den}(M) \to \mathbb{R}$. A consequence of the definition of h is the mass continuity equation

$$\partial_t h + \operatorname{div}(h\mathbf{u}) = 0, \tag{5.2}$$

with div being the divergence operator on M defined by $\mathcal{L}_{\mathbf{u}}\mathrm{d}\sigma = (\operatorname{div}\mathbf{u})\mathrm{d}\sigma$. Then, (5.1) yields the Euler–Poincaré variational principle with respect to constrained variations,

$$\delta \int_0^T \ell(\mathbf{u}, h)\, \mathrm{d}t = 0 \quad \text{for} \quad \begin{cases} \delta\mathbf{u} = \partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}] \\ \delta h = -\operatorname{div}(h\mathbf{v}), \end{cases} \tag{5.3}$$

where $\mathbf{v}$ is an arbitrary vector field with $\mathbf{v}(0) = \mathbf{v}(T) = 0$ and $[\cdot, \cdot]$ is the Lie bracket

of vector fields. To compute the equations of motion in Eulerian variables, we need the functional derivatives $\frac{\delta \ell}{\delta \mathbf{u}} \in \Omega^1(M)$ (one-forms on M) and $\frac{\delta \ell}{\delta h} \in F(M)$ (scalar functions on M) which are defined by the duality pairings,

$$\left\langle \frac{\delta \ell}{\delta \mathbf{u}}, \delta \mathbf{u} \right\rangle_1 := \int_M \frac{\delta \ell}{\delta \mathbf{u}} \cdot \delta \mathbf{u} \, \mathrm{d}\sigma = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \ell(\mathbf{u} + \varepsilon \delta \mathbf{u}, h), \tag{5.4}$$

$$\left\langle \frac{\delta \ell}{\delta h}, \delta h \right\rangle_0 := \int_M \frac{\delta \ell}{\delta h} \delta h \, \mathrm{d}\sigma = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \ell(\mathbf{u}, h + \varepsilon \delta h), \tag{5.5}$$

for arbitrary $\delta \mathbf{u}$ and δh. Note that we denote the duality pairing between a one-form $\frac{\delta \ell}{\delta \mathbf{u}}$ and a vector field $\delta \mathbf{u}$ as $\langle \cdot, \cdot \rangle_1$ and that between a function $\frac{\delta \ell}{\delta h}$ and a density δh as $\langle \cdot, \cdot \rangle_0$. The variational principle (5.3) yields the Euler–Poincaré equations,

$$\partial_t \frac{\delta \ell}{\delta \mathbf{u}} + \mathcal{L}_{\mathbf{u}} \frac{\delta \ell}{\delta \mathbf{u}} = h \mathbf{d} \frac{\delta \ell}{\delta h}, \tag{5.6}$$

where $\mathcal{L}_{\mathbf{u}} \mathbf{m} = \mathbf{i}_{\mathbf{u}} \mathbf{d}\mathbf{m} + \mathbf{d}(\mathbf{i}_{\mathbf{u}} \mathbf{m}) + \mathbf{m} \operatorname{div} \mathbf{u}$ is the Lie derivative of the fluid momentum $\mathbf{m}$ (a one-form density on M) and $\mathbf{d}$ is the exterior derivative.

In Euclidean space, the Euler–Poincaré equations reduce to

$$\partial_t \frac{\delta \ell}{\delta \mathbf{u}} + \mathbf{u} \cdot \nabla \frac{\delta \ell}{\delta \mathbf{u}} + \nabla \mathbf{u}^\top \frac{\delta \ell}{\delta \mathbf{u}} + \frac{\delta \ell}{\delta \mathbf{u}} \operatorname{div} \mathbf{u} = h \nabla \frac{\delta \ell}{\delta h}.$$

5.2.1. Variational principle for the SWE

For the rotating shallow water equations on a two dimensional Riemannian manifold M, the Lagrangian is given by

$$\ell(\mathbf{u}, h) = \int_M \left[\frac{1}{2} h \mathbf{u}^\flat \cdot \mathbf{u} + h \mathbf{r}^\flat \cdot \mathbf{u} - \frac{1}{2} g (h + \eta_b)^2 \right] \mathrm{d}\sigma, \tag{5.7}$$

where η_b is the bottom topography, g is the gravitational acceleration and $\mathbf{r}$ is the vector potential of the angular velocity of the Earth. Here, $\flat \colon TM \to TM^*$ is the flat operator of the Riemannian metric, that associates a one-form $\mathbf{u}^\flat$ to $\mathbf{u}$. With

the variational derivatives $\frac{\delta \ell}{\delta \mathbf{u}} = h(\mathbf{u}^\flat + \mathbf{r}^\flat)$ and $\frac{\delta \ell}{\delta h} = \frac{1}{2}\mathbf{u}^\flat \cdot \mathbf{u} + \mathbf{r}^\flat \cdot \mathbf{u} - g(h + \eta_b)$, the Euler–Poincaré equation (5.6) gives the momentum equation of the SWE in the space of one-forms:

$$\partial_t \mathbf{u}^\flat + \mathbf{i}_\mathbf{u} \mathbf{d}(\mathbf{u}^\flat + \mathbf{r}^\flat) + \mathbf{d}\left(\frac{1}{2}\mathbf{u}^\flat \cdot \mathbf{u} + g(h + \eta_b)\right) = 0. \tag{5.8}$$

This general expression reduces in the Euclidean space $\mathbb{R}^2$ to:

$$\partial_t \mathbf{u} + (\nabla \times (\mathbf{u} + \mathbf{r})) \times \mathbf{u} + \nabla\left(\frac{1}{2}|\mathbf{u}|^2 + g(h + \eta_b)\right) = 0. \tag{5.9}$$

Biharmonic dissipation. To remove small scale noise and improve the stability of the scheme, a common approach is to apply a linear fourth-order diffusion to the velocity field, see, e.g., [29, 73, 75]. For instance, on $\mathbb{R}^2$, we consider:

$$\partial_t \mathbf{u} + (\nabla \times (\mathbf{u} + \mathbf{r})) \times \mathbf{u} + \nabla\left(\frac{1}{2}|\mathbf{u}|^2 + g(h + \eta_b)\right) = -\nu \Delta^2 \mathbf{u}, \tag{5.10}$$

where ν is the diffusion coefficient. We do not add any dissipation to the continuity equation, because it does not contain a turbulent mixing term. Also, adding such dissipation can break conservation of mass, see [75].

In the next section, we review a new dissipation scheme for our framework, that only acts on one conserved quantity while conserving the energy.

5.3. Selective decay with Casimir dissipation

Given the Lagrangian $\ell(\mathbf{u}, h)$ of the fluid in Eulerian variables, the associated Hamiltonian function $\mathfrak{h}(\mathbf{m}, h)$ is obtained by the Legendre transformation

$$\mathfrak{h}(\mathbf{m}, h) = \langle \mathbf{m}, \mathbf{u} \rangle_1 - \ell(\mathbf{u}, h),$$

with $\mathbf{u}$ defined in terms of $(\mathbf{m}, h)$ by the relation $\mathbf{m} = \frac{\delta \ell}{\delta \mathbf{u}} \in \Omega^1(M)$. We note the relations

$$\frac{\delta H}{\delta \mathbf{m}} = \mathbf{u} \quad \text{and} \quad \frac{\delta H}{\delta h} = -\frac{\delta \ell}{\delta h}. \tag{5.11}$$

The Eulerian Lie–Poisson formulation is given by

$$\frac{df}{dt} = \{f, \mathfrak{h}\}, \quad \forall\, f, \tag{5.12}$$

with Lie–Poisson bracket $\{\cdot, \cdot\}$ defined as

$$\{f, \mathfrak{h}\} = -\int_M \mathbf{m} \cdot \left[\frac{\delta f}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right] \mathrm{d}\sigma + \int_M \rho \Big(\mathbf{d}\frac{\delta f}{\delta h} \cdot \frac{\delta \mathfrak{h}}{\delta \mathbf{m}} - \mathbf{d}\frac{\delta \mathfrak{h}}{\delta h} \cdot \frac{\delta f}{\delta \mathbf{m}} \Big) \mathrm{d}\sigma;$$

see [48] for details. The Lie–Poisson equations (5.12) are equivalent to the system of equations (5.2) and (5.6), as can by directly verified using (5.11).

For the selective decay, we use the relationship between the Lie–Poisson bracket and the conservation laws. A function C is called a *Casimir* for the Lie–Poisson bracket if it satisfies $\{C, f\} = 0$ for all f. With this, we have the conservation law $\frac{dC}{dt} = 0$ along solutions of the Lie–Poisson system $\dot{f} = \{f, \mathfrak{h}\}$, for any Hamiltonian $\mathfrak{h}$. In the next section, the Lie–Poisson bracket is extended to dissipate a Casimir but still conserve energy.

5.3.1. Casimir dissipation

In this section, we recall the approach to selective decay developed in [37]. Let $\gamma \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathbb{R}$ be a positive and symmetric bilinear form (with associated norm $\|\cdot\|_\gamma$) and C a Casimir function. The Casimir dissipation is introduced in the Lie–Poisson formulation as follows

$$\frac{df}{dt} = \{f, \mathfrak{h}\} - \theta\gamma\left(\left[\frac{\delta f}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right], \left[\frac{\delta C}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right]\right), \tag{5.13}$$

for some $\theta > 0$. If $f = \mathfrak{h}$, then

$$\frac{d\mathfrak{h}}{dt} = 0 - \theta\gamma\left(0, \left[\frac{\delta C}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right]\right) = 0,$$

and we see that the energy remains conserved. For $f = C$, we have

$$\frac{dC}{dt} = 0 - \theta\gamma\left(\left[\frac{\delta C}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right], \left[\frac{\delta C}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right]\right) = -\theta\left\|\left[\frac{\delta C}{\delta \mathbf{m}}, \frac{\delta \mathfrak{h}}{\delta \mathbf{m}}\right]\right\|_{\gamma}^2,$$

thus, the Casimir decays in time.

The corresponding Lagrange–d'Alembert variational principle is given by (see [37, Eq. (3.7)])

$$\delta\int_0^T \ell(\mathbf{u}, h)\mathrm{d}t + \theta\int_0^T \gamma\left(\left[\frac{\delta C}{\delta \mathbf{m}}, \mathbf{u}\right], \left[\mathbf{u}, \mathbf{v}\right]\right)\mathrm{d}t = 0, \quad \text{for} \quad \begin{cases} \delta\mathbf{u} = \partial_t \mathbf{v} + [\mathbf{u}, \mathbf{v}] \\ \delta h = -\operatorname{div}(h\mathbf{v}). \end{cases} \tag{5.14}$$

Then, the Casimir dissipative Euler–Poincaré equations (see [37, Eq. (3.3)]) are

$$\partial_t \frac{\delta \ell}{\delta \mathbf{u}} + \mathcal{L}_{\mathbf{u}}\frac{\delta \ell}{\delta \mathbf{u}} = h\mathbf{d}\frac{\delta \ell}{\delta h} + \theta\mathcal{L}_{\mathbf{u}}\left(\left[\mathbf{u}, \frac{\delta C}{\delta \mathbf{m}}\right]^{\gamma}\right), \tag{5.15}$$

where, for a vector field $\mathbf{u} \in \mathfrak{X}(M)$, $\mathbf{u}^{\gamma}$ is the one-form on M defined by $\int_M (\mathbf{u}^{\gamma} \cdot \mathbf{v})\mathrm{d}\sigma = \gamma(\mathbf{u}, \mathbf{v})$, for all $\mathbf{v} \in \mathfrak{X}(M)$, and we recall that $\mathbf{m} = \frac{\delta \ell}{\delta \mathbf{u}}$. We assume that γ is such that the one-form $\mathbf{u}^{\gamma}$ is well-defined for all $\mathbf{u} \in \mathfrak{X}(M)$, see [37] for examples.

5.3.2. Enstrophy dissipation for SWE

Next, we will consider enstrophy dissipation for the SWE. For two-dimensional fluid flows dominated by geostrophic balance, enstrophy is known to cascade to small scales. Thus, in order to obtain physically relevant solutions, it is necessary to dissipate enstrophy at such scales, see [17, 60, 73]. For the SWE on 2D

Riemannian manifolds, the potential enstrophy Casimir is given by

$$C(\mathbf{m},h)=\frac{1}{2}\int_M h\,q(\mathbf{m},h)^2\mathrm{d}\sigma \quad \text{with} \quad q(\mathbf{m},h)\mathrm{d}\sigma=\frac{1}{h}\mathbf{d}\frac{\mathbf{m}}{h}. \tag{5.16}$$

The variational derivative of the enstrophy Casimir is found as $\frac{\delta C}{\delta \mathbf{m}}=-\frac{1}{h}(\star\mathbf{d}q)^\sharp$ with $\sharp : T^*M \to TM$ the Riemannian sharp operator, see Appendix 5.8.1. For a two dimensional planar domain, these formulas reduce to

$$q(\mathbf{m},h)=\frac{1}{h}\mathbf{z}\cdot\nabla\times\left(\frac{\mathbf{m}}{h}\right), \qquad \frac{\delta C}{\delta \mathbf{m}}=-\frac{1}{h}\mathbf{z}\times\nabla q. \tag{5.17}$$

With the Lagrangian (5.7), the Casimir dissipative Euler–Poincaré equations (5.15) are given by

$$h\partial_t\mathbf{u}^\flat+h\mathbf{i}_\mathbf{u}\mathbf{d}(\mathbf{u}^\flat+\mathbf{r}^\flat)=-h\mathbf{d}\left(\frac{1}{2}\mathbf{u}^\flat\cdot\mathbf{u}+g(h+\eta_b)\right)+\theta\mathcal{L}_\mathbf{u}\left(h\left[\mathbf{u},\frac{\delta C}{\delta\mathbf{m}}\right]^\flat\right), \tag{5.18}$$

where we choose γ to be the water depth weighted L^2 inner product, i.e., $\gamma(\mathbf{u},\mathbf{v})=\int_M h(\mathbf{u}^\flat\cdot\mathbf{v})\mathrm{d}\sigma$, and we note that $\mathbf{u}^\gamma=h\mathbf{u}^\flat$, with $\flat$ associated to the Riemannian metric on M.

5.4. Discrete selective decay

The discretization process translates each step of the continuous theory to the discrete level. We review the variational discretization process for fluid initially developed in [65], see also [8, 26, 35, 36] for extensions, and incorporate into it the Casimir selective decay.

We consider a two dimensional simplicial mesh $\mathbb{M}$ with n cells on the fluid domain, where triangles (T) are used as the primal grid, and the circumcenter dual (ζ) as the dual grid. On the grid (see Fig. 5.1) we adopt the following notation:

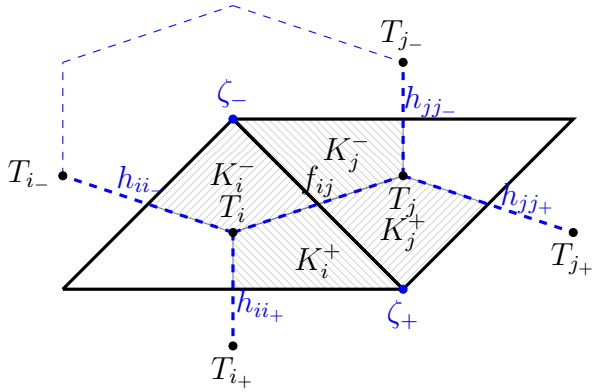


Figure 5.1.: Notation and indexing conventions for the 2D simplicial mesh.

- $e_{ij} = T_i \cap T_j$ as the primal edge,
- $\tilde{e}_{ij} = \zeta_+ \cap \zeta_-$ as the dual edge,
- Ω_{ii} as the area of triangle T_i,
- h_i as the discrete water depth on T_i,
- $(\eta_b)_i$ as the discrete bottom topography on T_i,
- V_{ij} is $\mathbf{u}_{e_{ij}} \cdot \mathbf{n}_{e_{ij}}$ at the edge midpoint.
- $\overline{h}_{ij} = \frac{1}{2}(h_i + h_j)$ as the water depth averaged to the edge midpoints.

Here, $\mathbf{n}_{e_{ij}}$ is the normal vector of edge e_{ij} pointing towards T_j.

5.4.1. Discrete setup

The Euler–Poincaré reduction from the Lie group of diffeomorphisms to its Lie algebra (as discussed in Section 5.2) is done analogously in the discrete setting by identifying the discrete analogues of $\mathrm{Diff}(M)$ and $\mathrm{Den}(M)$. For piecewise constant functions, the discrete diffeomorphism group is the matrix group

$$D(\mathbb{M}) = \{q \in GL(n)^+ \mid q \cdot \mathbf{1} = \mathbf{1}\}, \tag{5.19}$$

with $GL(n)^+$ the group of real $n \times n$ matrices with positive determinant and $\mathbf{1} = (1, ..., 1)^\mathsf{T}$. The condition $q \cdot \mathbf{1} = \mathbf{1}$ means that constants are preserved, which

is needed to obtain mass conservation. Then, the Lie algebra of $D(\mathbb{M})$ is

$$\mathfrak{d}(\mathbb{M}) = \{A \in \mathfrak{gl}(n) \mid A \cdot \mathbf{1} = 0\} \tag{5.20}$$

with the matrix commutator $[A, B] = AB - BA$ as the Lie bracket. This Lie algebra is a discrete version of the Lie algebra of $\mathrm{Diff}(M)$ given by vector fields on M. To reduce the computational complexity of the resulting scheme, nonholonomic constraints are imposed and, instead of $\mathfrak{d}(\mathbb{M})$, a subspace $\mathcal{R} \subset \mathfrak{d}(\mathbb{M})$ is considered [8], given by

$$\mathcal{R} = \Big\{A \in \mathfrak{d}(\mathbb{M}) \mid A^\top \Omega + \Omega A^\top \text{ is diagonal } \Big\} \cap \Big\{A \in \mathfrak{d}(\mathbb{M}) \mid A_{ij} = 0 \quad \forall j \notin N(i)\Big\},$$

with $N(i)$ being the set of cells sharing an edge with the cell T_i and with Ω being the $n \times n$ diagonal matrix with elements Ω_{ii}.

Remark 13. For $A, B \in \mathcal{R}$ we have $[A, B]_{ij} = 0$ for all $j \in N(i)$. Since elements of $\mathcal{R}$ are zero for non neighbouring cells, we get $[\mathcal{R}, \mathcal{R}] \cap \mathcal{R} = \{0\}$. In particular $[\mathcal{R}, \mathcal{R}] \neq \mathcal{R}$.

Next, we identify the dual space $\mathcal{R}^*$ with the space $\Omega^1_d(\mathbb{M})$ of discrete one-forms relative to the duality pairing on $\mathfrak{gl}(n)$:

$$\langle L, A\rangle_1 = \mathrm{Tr}(L^\top \Omega A). \tag{5.21}$$

To obtain an element in $\mathcal{R}^*$, we use the projection $P\colon \mathfrak{gl}(n) \to \Omega^1_d(\mathbb{M})$ defined by

$$P(L)_{ij} = \frac{1}{2}(L_{ij} - L_{ji} - L_{ii} + L_{jj}), \tag{5.22}$$

which satisfies $\langle L, A\rangle_1 = \langle P(L), A\rangle_1$, for all $A \in \mathcal{R}$, see [8]. Piecewise constant functions on M are represented by vectors $F \in \mathbb{R}^n$, with value F_i on cell i being the cell average of the continuous function on cell i. The space of discrete functions

is denoted by $\Omega_d^0(\mathbb{M})$, and the space of discrete densities $\mathrm{Den}_d(\mathbb{M}) \simeq \mathbb{R}^n$ is defined as the dual space to $\Omega_d^0(\mathbb{M})$ relative to the pairing:

$$\langle F, G\rangle_0 = F^\top \Omega G. \tag{5.23}$$

If a matrix $A \in \mathcal{R}$ approximates a vector field $\mathbf{u}$, then its entries satisfy

$$\begin{aligned} A_{ij} &\approx -\frac{1}{2\Omega_{ii}} \int_{T_i \cap T_j} (\mathbf{u}\cdot\mathbf{n}) \mathrm{d}S, \quad j \in N(i) \\ A_{ii} &\approx \frac{1}{2\Omega_{ii}} \int_{T_i} \operatorname{div} \mathbf{u}\, \mathrm{d}\sigma. \end{aligned} \tag{5.24}$$

In the next section, we will use this discrete setup to state the discrete variational principle and compute the numerical scheme.

5.4.2. Discrete variational equations for selective decay

Let $\ell\colon \mathfrak{d}(\mathbb{M}) \times \mathrm{Den}_d(\mathbb{M}) \to \mathbb{R}$ be a semi-discrete Lagrangian and $C\colon \mathfrak{d}(\mathbb{M}) \times \mathrm{Den}_d(\mathbb{M}) \to \mathbb{R}$ be a semi-discretized approximation of a Casimir. As above, let $\gamma\colon \mathfrak{d}(\mathbb{M}) \times \mathfrak{d}(\mathbb{M}) \to \mathbb{R}$ be a positive, symmetric bilinear form. Analogous to the continuous Casimir dissipative variational principle in Eq. (5.14), we consider the discrete dissipative variational principle given by

$$\delta \int_0^T \ell(A,h)\mathrm{d}t + \theta \int_0^T \gamma\left(\left[\frac{\delta C}{\delta M}, A\right], [A,B]\right) \mathrm{d}t = 0 \qquad \text{for} \qquad \begin{cases} \delta A = \partial_t B + [B,A] \\ \delta h = -\Omega^{-1} B^\top \Omega h \end{cases} \tag{5.25}$$

with $B \in \mathcal{R}$ and $B(0) = B(T) = 0$. The discrete functional derivatives $\frac{\delta \ell}{\delta A} \in \mathfrak{d}(\mathbb{M})^*$, $\frac{\delta \ell}{\delta h} \in \Omega^0(\mathbb{M})$ and $\frac{\delta C}{\delta M} \in \mathfrak{d}(\mathbb{M})$ are defined by

$$\begin{aligned} \left\langle \frac{\delta \ell}{\delta A}, \delta A \right\rangle_1 &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \ell(A + \varepsilon \delta A, h), \\ \left\langle \frac{\delta \ell}{\delta h}, \delta h \right\rangle_0 &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \ell(A, D + \varepsilon \delta h), \\ \left\langle \delta M, \frac{\delta C}{\delta M} \right\rangle_1 &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} C(M + \varepsilon \delta M, h), \end{aligned} \tag{5.26}$$

for all $\delta A \in \mathfrak{d}(\mathbb{M})$, $\delta h \in \mathrm{Den}_d(\mathbb{M})$, $\delta M \in \mathfrak{d}(\mathbb{M})^*$.

Theorem 14 (**Discrete dissipative variational equations**). *For a semi-discrete Lagrangian $\ell(A, D)$, the curves $A(t), h(t)$ are critical for the variational principle of Eq.* (5.25) *if and only if they satisfy*

$$P\left(\frac{d}{dt} \frac{\delta \ell}{\delta A} + \mathcal{L}_A \left(\frac{\delta \ell}{\delta A} \right) - \theta \mathcal{L}_A \left(h \left[\frac{\delta C}{\delta M}, A \right]^\flat \right) + h \frac{\delta \ell}{\delta h}^\top \right)_{ij} = 0, \tag{5.27}$$

where $\langle \mathcal{L}_A M, B \rangle_1 = \langle M, [A, B] \rangle_1$.

Proof: The variational principle (5.25) gives

$$0 = \delta \int_0^T \ell(A, h) \mathrm{d}t + \theta \int_0^T \gamma \left(\left[\frac{\delta C}{\delta M}, A \right], [A, B] \right) \mathrm{d}t.$$

Next, we use the definition of the flat operator and γ to be the water depth weighted inner product, giving

$$0 = \delta \int_0^T \ell(A, h) \mathrm{d}t + \theta \int_0^T \left\langle h \left[\frac{\delta C}{\delta M}, A \right]^\flat, [A, B] \right\rangle_1 \mathrm{d}t.$$

Finally, we use the expression of the variations in (5.25) and the definition of $\mathcal{L}$, see [8], which yields

$$
\begin{aligned}
0 &= -\int_0^T \left\langle \frac{d}{dt}\frac{\delta \ell}{\delta A} + \mathcal{L}_A\Big(\frac{\delta \ell}{\delta A}\Big) + h\frac{\delta \ell}{\delta h}^{\top}, B \right\rangle_1 \mathrm{d}t + \theta \int_0^T \left\langle h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat, [A,B] \right\rangle_1 \mathrm{d}t \\
&= -\int_0^T \left\langle \frac{d}{dt}\frac{\delta \ell}{\delta A} + \mathcal{L}_A\Big(\frac{\delta \ell}{\delta A}\Big) + h\frac{\delta \ell}{\delta h}^{\top}, B \right\rangle_1 \mathrm{d}t + \theta \int_0^T \left\langle \mathcal{L}_A\Big(h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat\Big), B \right\rangle_1 \mathrm{d}t \\
&= -\int_0^T \left\langle \frac{d}{dt}\frac{\delta \ell}{\delta A} + \mathcal{L}_A\Big(\frac{\delta \ell}{\delta A}\Big) - \theta\mathcal{L}_A\Big(h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat\Big) + h\frac{\delta \ell}{\delta h}^{\top}, B \right\rangle_1 \mathrm{d}t.
\end{aligned}
$$

The result then follows from $\int_0^T \langle L, B\rangle_1 \, \mathrm{d}t = 0, \ \forall B \in \mathcal{R} \iff P(L)_{ij} = 0$ (see [8, Proposition 2.3]).

□

Remark 15. We note that Theorem 14 becomes the *discrete variational equations* theorem in [8] for $\theta = 0$. This form of the discrete equations is valid on Cartesian and simplicial meshes in 2D and 3D. We focus below on two dimensional simplicial meshes.

The following proposition demonstrates that, for the resulting semi-discrete scheme, the energy is conserved.

Proposition 16. *Let $A(t)$ and $h(t)$ be the solution of* (5.27) *and $\dot{h} + \Omega^{-1}A^\top \Omega h = 0$. Then,*

$$
\frac{d}{dt}\left(\left\langle \frac{\delta \ell}{\delta A}, A \right\rangle_1 - \ell(A,h)\right) = 0
$$

Proof: We compute

$$\begin{aligned}
\frac{d}{dt}\left(\left\langle \frac{\delta \ell}{\delta A}, A\right\rangle_1 - \ell(A,h)\right) &= \left\langle \frac{d}{dt}\frac{\delta \ell}{\delta A}, A\right\rangle_1 + \left\langle \frac{\delta \ell}{\delta A}, \frac{d}{dt}A\right\rangle_1 - \left\langle \frac{\delta \ell}{\delta A}, \frac{d}{dt}A\right\rangle_1 - \left\langle \frac{\delta \ell}{\delta h}, \frac{d}{dt}h\right\rangle_0 \\
&= \left\langle P\frac{d}{dt}\frac{\delta \ell}{\delta A}, A\right\rangle_1 + \left\langle \frac{\delta \ell}{\delta h}, -\Omega^{-1}A^\top \Omega h\right\rangle_0 \\
&= \left\langle P\left(\frac{d}{dt}\frac{\delta \ell}{\delta A} + h\frac{\delta \ell}{\delta h}^\top\right), A\right\rangle_1 \\
&= -\left\langle P\left(\mathcal{L}_A \frac{\delta \ell}{\delta A} - \theta \mathcal{L}_A\Big(h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat\Big)\right), A\right\rangle_1 \\
&= -\left\langle \mathcal{L}_A \frac{\delta \ell}{\delta A} - \theta \mathcal{L}_A\Big(h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat\Big), A\right\rangle_1 = 0,
\end{aligned}$$

where the last equality follows from $\langle \mathcal{L}_A(X), A\rangle_1 = \langle X, [A,A]\rangle_1 = 0$. This holds independently of the chosen discretization of $\left[\frac{\delta C}{\delta M}, A\right]^\flat$.

□

5.4.3. Variational discretization of the Casimir dissipative SWE

Before presenting the discretization including the Casimir dissipation term, we briefly recall the variational discretization for the scheme without Casimir dissipation [8, 20]. We discretize the Lagrangian (5.7) with piecewise constant functions, giving

$$\ell(A,h) = \frac{1}{2}\sum_{i,j=1}^{n} h_i A_{ij}^\flat A_{ij}\Omega_{ii} + \sum_{i,j=1}^{n} h_i R_{ij}^\flat A_{ij}\Omega_{ii} - \frac{1}{2}\sum_{i=1}^{n} g(h_i + (\eta_b)_i)^2\Omega_{ii}. \quad (5.28)$$

To compute the variational derivatives we use the duality pairing (5.21) and the definition in Eq. (5.26). This gives

$$\frac{\delta \ell}{\delta A_{ij}} = h_i(A_{ij}^\flat + R_{ij}^\flat) \qquad \text{and} \qquad \frac{\delta \ell}{\delta h_i} = \frac{1}{2}\sum_j A_{ij}^\flat A_{ij} + \sum_j R_{ij}^\flat A_{ij} - g(h_i + (\eta_b)_i). \quad (5.29)$$

In [19, 20], it was noted that the approximations of the differential operators

result from the variational discretization method and agree with the following standard finite difference and finite volume operators:

$$
\begin{aligned}
(\mathrm{Grad}_n\ F)_{ij} &:= \frac{F_{T_j} - F_{T_i}}{|\tilde{e}_{ij}|}, & (\mathrm{div}\,\mathbf{u})_i \approx (\mathrm{Div}\ V)_i &:= \frac{1}{\Omega_{ii}} \sum_{k\in\{j,i_-,i_+\}} |e_{ik}| V_{ik}, \\
(\mathrm{Grad}_t\ F)_{ij} &:= \frac{F_{\zeta_-} - F_{\zeta_+}}{|e_{ij}|}, & (\nabla\times\mathbf{u})_\zeta \approx (\mathrm{Curl}\ V)_\zeta &:= \frac{1}{|\zeta|} \sum_{\tilde{e}_{nm}\in\partial\zeta} |\tilde{e}_{nm}| V_{nm},
\end{aligned}
\tag{5.30}
$$

for a scalar field F sampled either at the triangle or dual cell centres and a normal velocity V_{ij}. The gradient in the tangential and normal direction is noted with Grad_t and Grad_n, respectively. The normal velocity V_{ij} is related to the matrix elements $A \in \mathcal{R}$ in (5.24) as

$$
A_{ij} = -\frac{|e_{ij}|}{2\Omega_{ii}} V_{ij},\ j \in N(i) \quad \text{and} \quad A_{ii} = \frac{1}{2\Omega_{ii}} \sum_{k\in N(i)} |e_{ik}| V_{ik}.
$$

Remark 17. The gradient in the normal direction and the discrete divergence are adjoints with respect to the natural inner products on the triangles and their edges. Similarly, the tangential gradient and the discrete curl operator are adjoints with respect to the natural inner products on dual cells and their edges.

Remark 18. The continuous gradient, divergence, and curl operators are naturally written in Cartesian coordinates, but also can be defined (via parametrization) in a local neighbourhood on the sphere. The discrete counterparts are always locally defined and independent of the coordinate system. We will use the notation in Cartesian coordinates for this section for the continuum operators, to simplify notation.

Computing the projection (5.27) with $\theta = 0$, we obtain the momentum equation in [8]. This results in multiple terms. Here, we combined the terms involved in the advection term $(\nabla\times(\mathbf{u}+\mathbf{r}))\times\mathbf{u}$ and denoted them with Adv, terms involved in the kinetic energy term $\nabla(\frac{1}{2}\mathbf{u}^2)$ with K and the terms involved in the gradient

term $g\nabla h$ with G.

$$\partial_t V_{ij} = -\mathrm{Adv}(V,h)_{ij} + \mathrm{K}(V)_{ij} - \mathrm{G}(h)_{ij}, \tag{5.31}$$

where

$$\begin{aligned}
&\mathrm{Adv}(V,h)_{ij} := \\
&-\frac{1}{\overline{h}_{ij}|\tilde{e}_{ij}|}\Big((\mathrm{Curl}\ V)_{\zeta_-} + f_{\zeta_-}\Big)\left(\frac{|\zeta_- \cap T_i|}{2\Omega_{ii}}\overline{h}_{ji_-}|e_{ii_-}|V_{ii_-} + \frac{|\zeta_- \cap T_j|}{2\Omega_{jj}}\overline{h}_{ij_-}|e_{jj_-}|V_{jj_-}\right) \\
&+\frac{1}{\overline{h}_{ij}|\tilde{e}_{ij}|}\Big((\mathrm{Curl}\ V)_{\zeta_+} + f_{\zeta_+}\Big)\left(\frac{|\zeta_+ \cap T_i|}{2\Omega_{ii}}\overline{h}_{ji_+}|e_{ii_+}|V_{ii_+} + \frac{|\zeta_+ \cap T_j|}{2\Omega_{jj}}\overline{h}_{ij_+}|e_{jj_+}|V_{jj_+}\right), \\
&\mathrm{K}(V)_{ij} := \frac{1}{2}(\mathrm{Grad}_n\ F)_{ij}, \qquad\qquad F_{T_i} = \sum_{k\in\{j,i_-,i_+\}} \frac{|\tilde{e}_{ik}|\ |e_{ik}|(V_{ik})^2}{2\Omega_{kk}}, \\
&\mathrm{G}(h)_{ij} := g(\mathrm{Grad}_n\ (h+\eta_b))_{ij}.
\end{aligned}$$

The Coriolis parameter is defined by

$$f_\zeta = \frac{1}{|\zeta|}\sum_{\tilde{e}_{nm}\in\partial\zeta} |\tilde{e}_{nm}| r_{nm}, \qquad \text{with } r_{ij} = \mathbf{r}_{e_{ij}} \cdot \mathbf{n}_{e_{ij}}.$$

Casimir dissipative scheme:

Including the extra term for $\theta > 0$, the Casimir dissipative momentum equation is

$$\partial_t V_{ij} = -\mathrm{Adv}(V,h)_{ij} + \mathrm{K}(V)_{ij} - \mathrm{G}(h)_{ij} + \theta \mathrm{L}(V,h,\frac{\delta C}{\delta M})_{ij}, \tag{5.32}$$

where

$$P\left(\mathcal{L}_A\Big(h\Big[\frac{\delta C}{\delta M}, A\Big]^\flat\Big)\right)_{ij} =: \mathrm{L}(V,h,\frac{\delta C}{\delta M})_{ij}.$$

To compute this term, we first need to discretize the commutator. Here, we cannot follow the discretization procedure of [65] for the commutator of vector fields $[A,B]$. This is due to the fact that $[A,B] \in [\mathcal{R},\mathcal{R}]$, for $A,B \in \mathcal{R}$ and $[\mathcal{R},\mathcal{R}] \neq \mathcal{R}$, see Remark 13; and the flat operator $\flat$ is only defined for matrices in

$\mathcal{R}$. To obtain a discrete vector $W \in \mathcal{R}$ approximating the commutator $\left[\frac{\delta C}{\delta \mathbf{m}}, \mathbf{u}\right]$ at the edge midpoint, we will use the standard operators from Eq. (5.30). Then, we can compute $P(\mathcal{L}_A \left(hW^\flat\right))_{ij}$ using [8, Lemma 3.1].

Discrete commutator. Let $U_{ij} = \mathbf{u}_{e_{ij}} \cdot \mathbf{n}_{e_{ij}}$ be the edge normal for a vector field $\mathbf{u}$ at edge e_{ij} and V_{ij} for a vector field $\mathbf{v}$ respectively. The Lie bracket for vector fields $\mathbf{u}, \mathbf{v}$ is given below and can be rewritten using a standard vector calculus identity, giving

$$[\mathbf{u}, \mathbf{v}] = \mathbf{u} \cdot \nabla \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u} = \mathbf{u} \operatorname{div} \mathbf{v} - \mathbf{v} \operatorname{div} \mathbf{u} - \nabla \times (\mathbf{u} \times \mathbf{v}).$$

We discretize $\mathbf{u} \operatorname{div} \mathbf{v}$ and $\mathbf{v} \operatorname{div} \mathbf{u}$ using the discrete divergence on each triangle (Eq. (5.30)), averaging over the two adjacent triangles to obtain an edge value,

$$\begin{aligned}\Big(\mathbf{u} \operatorname{div} \mathbf{v}\Big)_{ij} &= U_{ij}\Big(\frac{\operatorname{div}(V)_i + \operatorname{div}(V)_j}{2}\Big)\\ \Big(\mathbf{v} \operatorname{div} \mathbf{u}\Big)_{ij} &= V_{ij}\Big(\frac{\operatorname{div}(U)_i + \operatorname{div}(U)_j}{2}\Big).\end{aligned}$$

Then, to obtain a discrete version of $\nabla \times (\mathbf{u} \times \mathbf{v})$, we use the following procedure:

- Reconstruct the full vector field $\mathbf{u}_\zeta$ and $\mathbf{v}_\zeta$ at the dual cell centres from the normal values U_{ij} and V_{ij}. We use the reconstruction in the interior of each triangle proposed by [66] and map it to the dual cell:

$$\begin{aligned}\mathbf{u}_\zeta &= \sum_{i \in N(\zeta)} \frac{|\zeta \cap T_i|}{|\zeta|}\mathbf{u}_i, \qquad \text{where} \qquad \mathbf{u}_i = \frac{1}{\Omega_{ii}} \sum_{k \in \{j, i_-, i_+\}} |e_{ik}|(\mathbf{x}_{e_{ik}} - \mathbf{x}_{T_i})U_{ij},\\ \mathbf{v}_\zeta &= \sum_{i \in N(\zeta)} \frac{|\zeta \cap T_i|}{|\zeta|}\mathbf{v}_i, \qquad \text{where} \qquad \mathbf{v}_i = \frac{1}{\Omega_{ii}} \sum_{k \in \{j, i_-, i_+\}} |e_{ik}|(\mathbf{x}_{e_{ik}} - \mathbf{x}_{T_i})V_{ij}.\end{aligned}$$

- Compute the cross product $c_\zeta = (\mathbf{u}_\zeta \times \mathbf{v}_\zeta) \cdot \mathbf{k}_\zeta$, where $\mathbf{k}$ is the unit vector that points in the local vertical direction.

- Obtain $\Big(\nabla\times(\mathbf{u}\times\mathbf{v})\Big)_{ij} = \mathrm{Grad}_t\, c_\zeta$. Since c_ζ is located at the dual cell centres and the resulting value after taking the curl should be an edge normal value (tangential for the dual grid cells), we use Grad_t, which is the adjoint curl (see Remark 17).

We obtain W approximating $\left[\frac{\delta C}{\delta \mathbf{m}}, \mathbf{u}\right]$, as $W_{ij} = \frac{|e_{ij}|}{2\Omega_{ii}}\widetilde{W}_{ij}$ and $W_{ii} = (\mathrm{Div}\ \widetilde{W})_i$, where

$$\begin{aligned}\widetilde{W}_{ij} = \left(\frac{\delta C}{\delta M}\right)_{ij}\Big(\frac{\mathrm{div}(V)_i + \mathrm{div}(V)_j}{2}\Big) - V_{ij}\Big(\frac{\mathrm{div}\left(\frac{\delta C}{\delta M}\right)_i + \mathrm{div}\left(\frac{\delta C}{\delta M}\right)_j}{2}\Big)\\ - \mathrm{Grad}_t\left((\frac{\delta C}{\delta \mathbf{m}_\zeta}\times\mathbf{u}_\zeta)\cdot\mathbf{k}_\zeta\right)_{ij}.\end{aligned}$$

In [8, Lemma 3.1], the discrete projection of the Lie derivative $\mathcal{L}^d_A$ is given by

$$\begin{aligned}P\Big(\mathcal{L}^d_A(hW^\flat)\Big)_{ij} &= (\underset{\zeta_-}{\mathrm{curl}}\,\widetilde{W})\left(\frac{|\zeta_-\cap T_i|}{2\Omega_{ii}}\overline{h}_{ji_-}|e_{ii_-}|V_{ii_-} + \frac{|\zeta_-\cap T_j|}{2\Omega_{jj}}\overline{h}_{ij_-}|e_{jj_-}|V_{jj_-}\right)\\ &+ (\underset{\zeta_+}{\mathrm{curl}}\,\widetilde{W})\left(\frac{|\zeta_+\cap T_i|}{2\Omega_{ii}}\overline{h}_{ji_+}|e_{ii_+}|V_{ii_+} + \frac{|\zeta_+\cap T_j|}{2\Omega_{jj}}\overline{h}_{ij_+}|e_{jj_+}|V_{jj_+}\right)\\ &+ \overline{h}_{ij}\left(\sum_{k\in N(i)}\frac{|e_{ik}||\tilde{e}_{ik}|}{\Omega_{ii}}V_{ik}\widetilde{W}_{ik} - \sum_{k\in N(j)}\frac{|e_{jk}||\tilde{e}_{jk}|}{\Omega_{jj}}V_{jk}\widetilde{W}_{jk}\right)\\ &+ \frac{\mathrm{div}(V\overline{h})_i + \mathrm{div}(V\overline{h})_j}{2}(2|\tilde{e}_{ij}|\widetilde{W}_{ij}),\end{aligned} \tag{5.33}$$

which we can evaluate using the discrete operators above once we know $\widetilde{W}$.

Discrete enstrophy variational derivative Analogously to Section 5.3.2, we compute the discrete variational derivative of the approximation of the enstrophy Casimir and substitute it into Eq. (5.32). The discrete enstrophy Casimir is

$$\mathcal{C}(M,h) = \frac{1}{2}\sum_\zeta h_\zeta\Big(q(M,h)_\zeta\Big)^2|\zeta|,\quad q(M,h)_\zeta = \frac{(\mathrm{Curl}\ V)+f}{h_\zeta},\quad h_\zeta = \sum_{T_i\cap\zeta\neq\emptyset}\frac{|T_i\cap\zeta|}{|\zeta|}h_i,$$

where $M = \frac{\delta \ell}{\delta A}$ and f is the Coriolis parameter.

Then, computing the variational derivative (see Appendix 5.8.2 for details) we obtain

$$\frac{\delta \mathcal{C}}{\delta M}_{ij} = \frac{q_{\zeta_+} - q_{\zeta_-}}{\Omega_{ii} h_{ij}} = -\frac{|e_{ij}|}{2\Omega_{ii}} \Big(2 \frac{q_{\zeta_-} - q_{\zeta_+}}{|e_{ij}|} \frac{1}{h_{ij}} \Big) = -\frac{|e_{ij}|}{2\Omega_{ii}} \frac{2 \ \mathrm{Grad}_t \ q}{h_{ij}}.$$

Substituting this into Eq. (5.32) results in the discretized potential enstrophy dissipating SWE.

Remark 19. The approximation of the enstrophy Casimir is not a Casimir of the discrete system. Therefore, we cannot directly prove that enstrophy is dissipated for the semi-discrete scheme. However, the numerical results demonstrate that the numerical scheme indeed dissipates enstrophy.

5.4.4. Temporal discretization

A temporal variational discretization can be obtained by following the discrete (in time) Euler–Poincaré–d'Alembert approach, see [26, 36]. This approach is based on the Cayley transform, a local approximation to the exponential map of the Lie group. In particular, the resulting scheme uses the Cayley transform in the update for the continuity equation and a Crank–Nicolson-type update for the momentum equation at each timestep. For the selective decay, the dissipation term is added to the Crank–Nicolson-type update.

Based on the Cayley transformation, the continuity update equation is then given by $h^{t+1} = \tau(\Delta t A^t) h^t$ for the time t and a time step size Δt, where the action of τ can be represented by solving

$$\Big(I - \frac{1}{2} \Delta t A^t \Big) h^{t+1} = \Big(I + \frac{1}{2} \Delta t A^t \Big) h^t, \tag{5.34}$$

with I being the identity matrix . Then, we use the following fixed-point iteration

to approximately solve the discrete momentum equation:

1. Start loop over k with initial guess as solution at time t: $V^*_{k=0} = V^t$;

2. Calculate updated velocity V^*_{k+1} from the explicit equation:

$$\frac{V^*_{k+1} - V^t}{\Delta t} = -\frac{\mathrm{Adv}(V^*_k, h^{t+1}) + \mathrm{Adv}(V^t, h^t)}{2} + \frac{\mathrm{K}(V^*_k) + \mathrm{K}(V^t)}{2} - \mathrm{G}(h^{t+1}) - \theta\frac{\mathrm{L}(V^*_k, h^{t+1}, \frac{\delta C}{\delta M}^t) + \mathrm{L}(V^t, h^t, \frac{\delta C}{\delta M}^t)}{2};$$

3. Stop loop over k if $||V^*_{k+1} - V^*_k|| < \epsilon$ for a small positive ϵ, take $V^{t+1} = V^*_{k+1}$.

For more details, we refer the reader to [8, 20].

Remark 20. This temporal discretization is only an approximation to a fully variational time integrator, see [26, 36] for more details. In particular, it has been observed that the temporal integrator does not conserve energy at the level of machine precision but, rather, the energy error fluctuates around a long term mean. Thus, while the energy is conserved by the semi-discrete Casimir dissipative equations, independent of the discretization of the commutator (see Proposition 16), this does not guarantee that the energy will be conserved after temporal discretization. We observe this in the numerical results that follow, but ascribe the small energy growth seen there to errors from this temporal discretization.

5.4.5. Biharmonic dissipation

To compare against a standard dissipation such as a biharmonic eddy viscosity, we discretize the dissipation term in Eq. (5.10), using the vector calculus identity for the vector Laplacian

$$\Delta \mathbf{u} = \nabla\ \mathrm{div}\, \mathbf{u} - \nabla \times (\nabla \times \mathbf{u}).$$

Then, using the discrete operators (5.30), we obtain

$$\begin{aligned}\mathrm{lap}(V)_{ij} &= \mathrm{Grad}_n(\mathrm{Div}\ V)_{ij} - \mathrm{Grad}_t(\mathrm{Curl}\ V)_{ij}\\ \nu L(V)_{ij} &= \nu\ \mathrm{lap}(\mathrm{lap}(V))_{ij}.\end{aligned}$$

We obtain the discrete version of Eq. (5.10)

$$\partial_t V_{ij} = -\mathrm{Adv}(V,h)_{ij} + \mathrm{K}(V)_{ij} - \mathrm{G}(h)_{ij} - \nu \mathrm{L}(V)_{ij}. \tag{5.35}$$

The temporal discretization is the same as above, but with $-\nu \mathrm{L}(V)_{ij}$ instead of $\theta\, \mathrm{L}(V^t, h^t, \frac{\delta C}{\delta M}^t)$.

5.5. Numerical results

The numerical simulations on the plane are performed on a doubly periodic rectangular domain $M = [0, L_x] \times [0, L_y]$ with $L_x = 5000\ km$ and $L_y = 4330\ km$. We consider an f-plane approximation with constant Coriolis parameter f set to 5.3108 $days^{-1}$ and $g = 9.81\ m/s$. Unless otherwise noted, the simulations are performed using a resolution of $N = 32768$ triangles. For the simulations on the sphere, we use an icosahedral grid and set the Earth's radius $R = 6.37122 \times 10^6\ m$, the Coriolis parameter to be $f = 2\Omega \sin(\Theta)$, where $\Omega = 7.292 \times 10^{-5}\ s^{-1}$, and $g = 9.81\ m/s$. Here, Θ is the latitude and Λ the longitude. The simulations are performed using a resolution of $N = 81920$ triangles.

We define the discrete total energy H, namely the Hamiltonian, and the discrete potential enstrophy $\mathcal{C}$:

$$H = \sum_{T_i} \frac{g}{2}(h_i + (\eta_b)_i)^2 \Omega_{ii} + \frac{1}{2}\Omega_{ii} h_i \sum_{k=j,i_-,i_+} \frac{|e_{ik}|\ |\tilde{e}_{ik}| V_{ik}^2}{2} \tag{5.36}$$

$$\mathcal{C} = \frac{1}{2}\sum_{\zeta} \frac{\mathrm{Curl}(V(t)) + f}{h_\zeta(t)} |\zeta|. \tag{5.37}$$

For each simulation, we first choose the dissipation coefficient ν, for the biharmonic dissipation, and then empirically choose θ for the Casimiar dissipation term so that the dissipation of potential enstrophy is very similar.

5.5.1. Numerical analysis of the discrete commutator

We first present a convergence study for the discrete commutator on both the plane and sphere. We define

$$\mathbf{u} = \begin{pmatrix} \sin(\frac{2\pi x}{L_x}) \\ 0 \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} \cos(\frac{2\pi x}{L_x}) \\ 0 \end{pmatrix} \text{ such that } [\mathbf{u}, \mathbf{v}] = \begin{pmatrix} \frac{2\pi}{L_x} \\ 0 \end{pmatrix}$$

for the test case on the plane, and

$$\mathbf{u} = \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} 0 \\ -z \\ y \end{pmatrix} \text{ such that } [\mathbf{u}, \mathbf{v}] = \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix}$$

for the test case on the sphere. We approximate these vector fields with piecewise constant functions and follow the algorithm in 5.4.3 to discretize $[\mathbf{u}, \mathbf{v}]$. Then, we compute the error between the approximation of the discrete commutator to the analytic field projected on the edge normal direction. To estimate the numerical errors, we use the following definitions for the relative L_2 and L_∞ error on edge values:

$$L_2 = \frac{\sqrt{\sum_{ij} |\mathfrak{e}_{ij}| \Big(u_n(e_{ij}) - u_r(e_{ij})\Big)^2}}{\sqrt{\sum_{ij} |\mathfrak{e}_{ij}| u_r(e_{ij})^2}}, \qquad L_\infty = \frac{\max_{ij} |u_n(e_{ij}) u_r(e_{ij})|}{\max_{ij} |u_r(e_{ij})|},$$

where $u_n(e_{ij})$ is the numerical solution defined at edge e_{ij} and $u_r(e_{ij})$ is the analytical solution evaluated at the edge midpoint e_{ij}. Moreover $|\mathfrak{e}_{ij}| = \frac{1}{2}|e_{ij}|\ |\tilde{e}_{ij}|$ is the area associated to an edge.

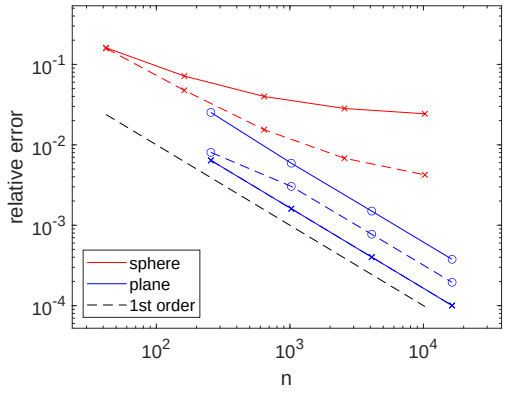


Figure 5.2.: Convergence of the commutator on the plane and sphere. Solid lines represent the L_∞ error and dashed lines the L_2 error. On the plane, we present results using both a regular grid, with data denoted by x, and an irregular grid (with a central refinement region), with data denoted by o.

We observe that the resulting approximation on both a regular and irregular grid on the plane is first-order accurate, see Fig. 5.2. On the sphere, the approximation is less than first-order accurate, however, because the icosahedral grid is optimised for the dual grid (pentagons/hexagons), see [44] for further details, but we evaluate the discrete divergence and reconstruction of the vector fields on the triangles. This low-order convergence of the divergence was also observed in [20].

5.5.2. Vortex interaction

This test case consists of two counter-rotating vortices in the plane and is dominated by non-linear processes. The two vortices are initially placed too far apart to merge. Thus, a key point in this simulation is that adding the Casimir dissipation does not change the evolution of the vortices.

Initial conditions. The initial height function for this example [8] is given by

$$h\big(x,y,t=0\big) = H_0 - H'\bigg(\exp\Big(-\frac{{x_1'}^2+{y_1'}^2}{2}\Big)+\exp\Big(-\frac{{x_2'}^2+{y_2'}^2}{2}\Big)-\frac{4\pi s_x s_y}{L_x L_y}\bigg), \tag{5.38a}$$

where $H_0 = 750\,m$, $H' = 75\,m$, and the periodic extensions are given by

$$x_i' = \frac{L_x}{\pi s_x}\sin\Big(\frac{\pi}{L_x}(x - x_{c_i})\Big), \qquad y_i' = \frac{L_y}{\pi s_y}\sin\Big(\frac{\pi}{L_y}(y - y_{c_i})\Big), \qquad i = 1, 2 \tag{5.38b}$$

with the centres located at $(x_{c_1}, y_{c_1}) = (2/5)\,(L_x, L_y)$, $(x_{c_2}, y_{c_2}) = (3/5)\,(L_x, L_y)$ and $(s_x, s_y) = \frac{3}{40}(3/40)\,(L_x, L_y)$. The discrete initial water depth on each triangle, h_i, is obtained by sampling the analytical water depth at the cell centre. Then, the initial condition for the velocity is given by the discrete geostrophic velocity,

$$V_{ij} = -\frac{g}{f}\mathrm{Grad}_t(h)_{ij}.$$

In these simulations, we use dissipation parameters $\nu = 1.2724 \times 10^5$ and $\theta = 2$.

We first integrate the initial conditions for 2 days for different values of the time step, to analyse the convergence of the energy. In Fig. 5.3, we observe that the energy converges with first-order accuracy. Then, to analyse the effects of the Casimir dissipation, we integrate the initial conditions for 10 days and compare the relative potential vorticity field against a simulation with no dissipation and one with biharmonic dissipation, see Fig. 5.4. All simulations behave similarly, with the cores of the two vortices being mutually repelled, due to non-linear effects. We note that the simulation with no dissipation becomes noisy loses physical fidelty on this timescale, while the two simulations with dissipation retain their accuracy.

The quantities of interest, total energy and potential enstrophy, are shown in Fig. 5.5. We observe that the enstrophy is dissipated at the same rate for the simulations with biharmonic and Casimir dissipation, as expected with this choice of dissipation parameters. While the energy is dissipated in the simulation with biharmonic dissipation, conservation of energy for the simulation using Casimir dissipative simulation is similar to that of the simulation with no dissipation. As noted in Remark 20 above, the temporal discretization used here is not completely

energy conserving, leading to the oscillations seen at the left of Fig. 5.5.

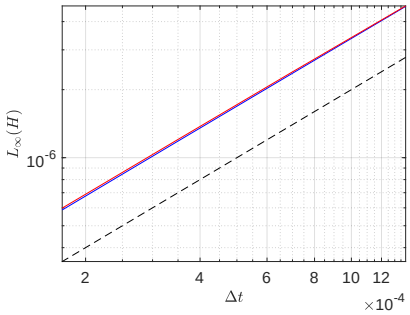


Figure 5.3.: Convergence of the energy with respect to the time step size on a regular (blue) and irregular (red) grid. The dashed black line indicates first order.

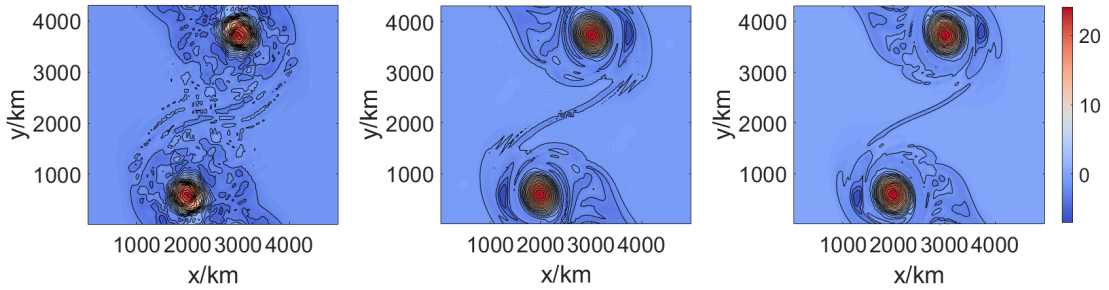


Figure 5.4.: Interacting vortices test case: Comparison of the relative potential vorticity for a simulation without dissipation (left), with Casimir dissipation (middle) and with biharmonic dissipation (right) after 10 days.

5.5.3. Shear flow

We next consider a shear flow test case in the quasi-geostrophic regime [8], with strongly dominant non-linear effects. The shear flow is initialized to an unstable equilibrium state so that, after a few days, the instability develops. This test case demonstrates that adding the Casimir dissipation does not change the development and growth of this instability.

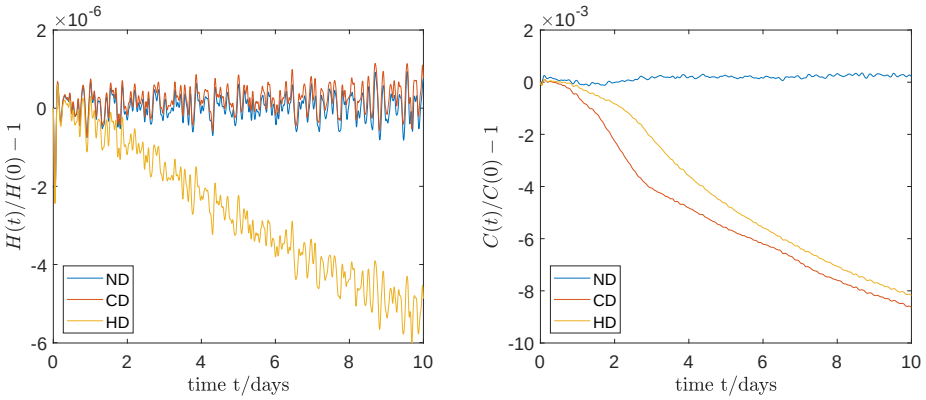


Figure 5.5.: Interacting vortices test case: Comparison of the relative errors in the energy (left) and potential enstrophy (right) for a simulation without dissipation (blue), with Casimir dissipation (red) and with a biharmonic dissipation (yellow).

Initial conditions. The initial height for this example is given by

$$h(x, y, t = 0) = H_0 - H' \frac{y''}{\sigma_y} e^{-\frac{y'^2}{2\sigma_y^2} + \frac{1}{2}} \left(1 - \kappa \sin\left(\frac{2\pi x'}{\lambda_x}\right)\right),$$

where

$$x' = \frac{x}{L_x}, \qquad y' = \frac{1}{\pi} \sin\left(\frac{\pi}{Ly}\left(y - \frac{L_y}{2}\right)\right), \qquad y'' = \frac{1}{2\pi} \sin\left(\frac{2\pi}{L_y}\left(y - \frac{L_y}{2}\right)\right),$$

with parameters $\lambda_x = \frac{1}{2}, \sigma_y = \frac{1}{12}, \kappa = 0.1, H_0 = 1.076\ km$ and $H' = 0.03\ km$. Again, the velocity field is initialized to be the discrete geostrophic velocity,

$$V_{ij} = -\frac{g}{f} \mathrm{Grad}_t(h)_{ij}.$$

Here, the dissipation parameters are chosen as $\nu = 3.7145 \times 10^5$ and $\theta = 2$.

We integrate the initial conditions for 10 days. The instability develops in the first three days, then the flow evolves into pairs of counter-rotating vortices. The filaments between the vortices become thinner until they can no longer be resolved by the spatial resolution of the mesh. This causes a noisy pattern in

the vorticity field at day ten for the simulation without any dissipation, see Fig. 5.6. In contrast, the simulations with Casimir and biharmonic dissipation are not polluted.

The quantities of interest for this simulation are shown in Fig. 5.7. Again, we observe the similar dissipation rate of the potential enstrophy for the Casimir and biharmonic dissipation, by construction. The simulation with no dissipation and the Casimir dissipative simulation have a similar conservation of energy. In contrast, the simulation with biharmonic dissipation has a loss of energy about 100 times greater.

Fig. 5.8 shows the kinetic energy and potential enstrophy spectra for simulations on refined spatial meshes with 524288 triangles. Expected scaling laws for these spectra are discussed in [23, 73]. Both dissipative simulations follow the expected k^{-1} power law for the enstrophy and k^{-3} power law for the kinetic energy over a significant region of the resolved wavenumbers. However, we note that using Casimir dissipation results in better resolution of the spectra over the small scales (higher wavenumbers) in comparison with the biharmonic dissipation. As expected, the biharmonic dissipation results in much faster dissipation over small scales in both the energy and enstrophy.

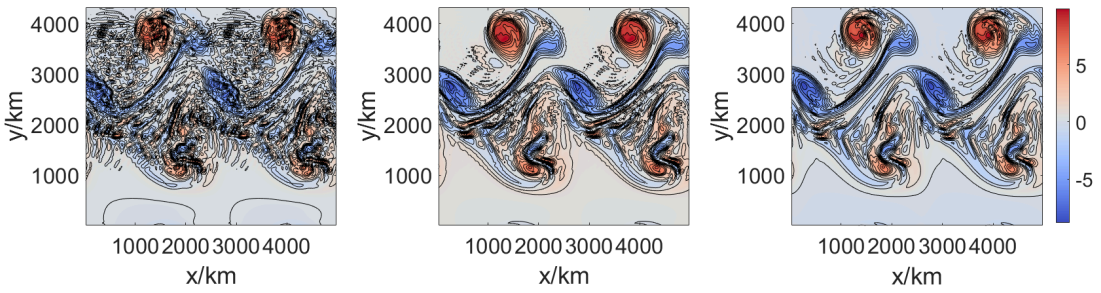


Figure 5.6.: Shear flow test case: Comparison of the relative potential vorticity for a simulation without dissipation (left), with Casimir dissipation (middle) and with biharmonic dissipation (right) after 10 days.

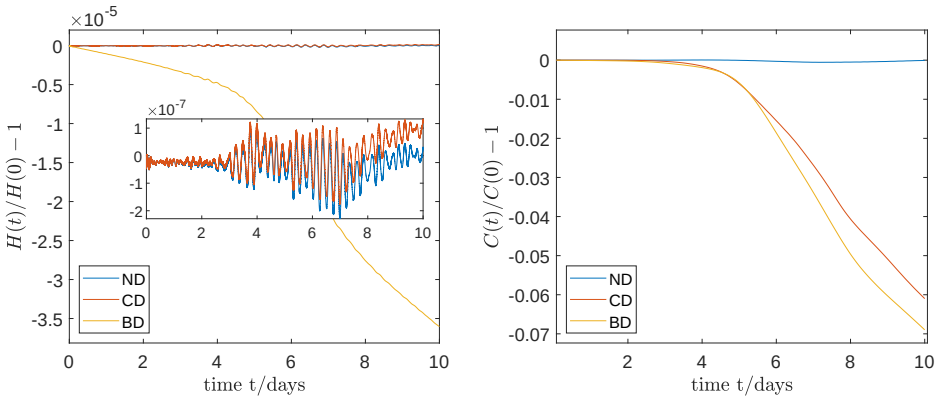


Figure 5.7.: Shear flow test case: Comparison of the relative errors in the energy (left) and potential enstrophy (right) for a simulation without dissipation (blue), with Casimir dissipation (red) and with biharmonic dissipation (yellow).

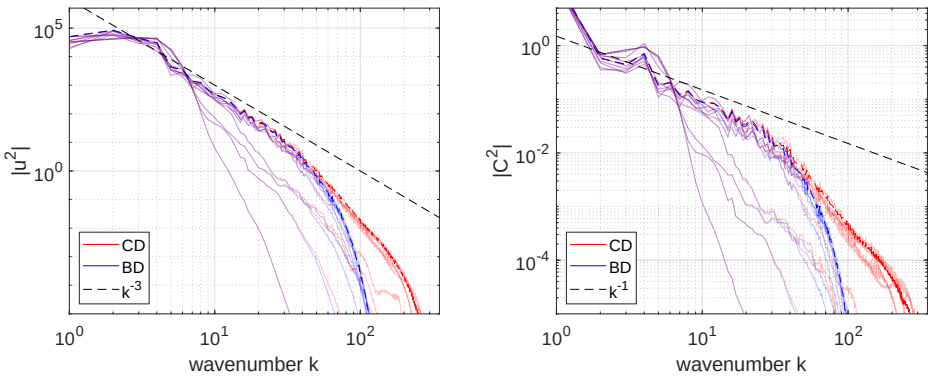


Figure 5.8.: Shear flow test case: Comparison of the kinetic energy (left) and potential enstrophy (right) spectrum for simulations with Casimir dissipation (red) and biharmonic dissipation (blue). The spectra are obtained from simulations with a resolution of $N = 524288$ triangles. The spectra are shown for days 1 to 10 of the simulation. The dashed red and blue lines show the averages of the spectra from days 6 to 10.

5.5.4. Flow over an isolated mountain

As a final example, we consider the flow over a conically-shaped mountain on the sphere, as proposed in [90]. The initially balanced flow runs over the mountain, which initiates turbulence. The flow stays turbulent for a long period of time.

Initial conditions. The discrete initial velocity and height fields are given in spherical coordinates as

$$V_{ij} = u_0(\cos(\Theta), 0)^\top \cdot \mathbf{n}_{ij} \qquad u_0 = 20\ m/s$$
$$h_i = h_0 - \frac{1}{g}(R\Omega u_0 + u_0^2/2)\cos(\Theta) \qquad h_0 = 5960\ m$$

The conically shaped bottom topography is given by

$$\eta_b(\Lambda, \Theta) = 2000(1 - 9r/\pi), \qquad r^2 = \min\Big((\pi/9)^2, (\Lambda - \Lambda_c)^2 + (\Theta - \Theta_c)^2\Big),$$

where $\Lambda_c = 3\pi/2$ and $\Theta_c = \pi/6$.

We integrate the initial conditions for 100 days and observe that the simulation without any stabilization becomes noisy and unphysical, see Fig. 5.9. In contrast, the stabilized schemes produce coherent structures in the vorticity field. In Fig. 5.10, we show the quantities of interest. The simulation without dissipation shows an increase in potential enstrophy, which is related to the noisy vorticity field. The dissipative schemes, as expected, dissipate potential enstrophy at the same rate. The error in the energy of the simulation with the Casimir dissipation stays on the same order as the simulation without any dissipation, while the simulation using biharmonic dissipation has an energy loss.

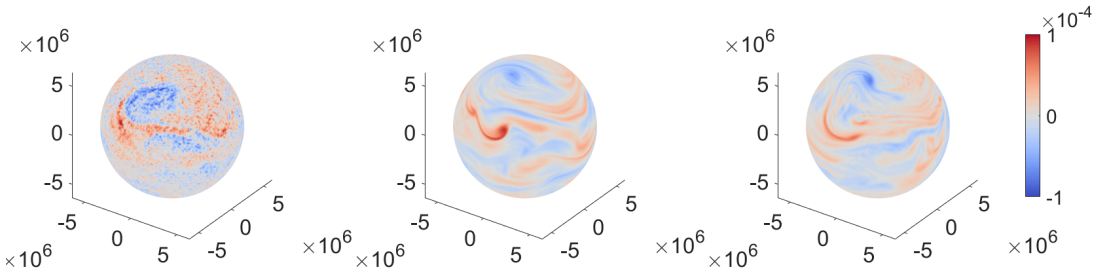


Figure 5.9.: Flow over a mountain test case: Comparison of the relative vorticity for a simulation without dissipation (left), with Casimir dissipation (middle), and with biharmonic dissipation (right) after 100 days.

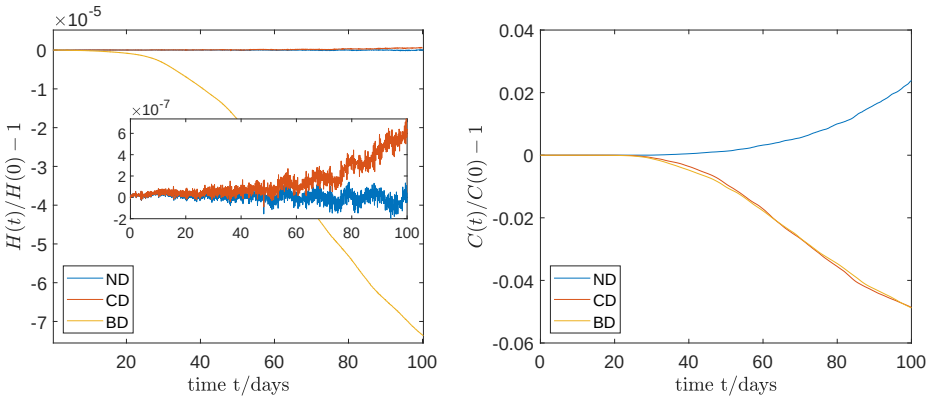


Figure 5.10.: Flow over mountain test case: Comparison of the relative errors in the energy (left) and potential enstrophy (right) for a simulation without dissipation (blue), with Casimir dissipation (red), and with biharmonic dissipation (yellow).

5.6. Conclusions

The development of high-fidelity numerical simulation tools for weather and climate prediction is limited by the competing goals of achieving energy conservation while preserving long-term stability of the time-integration scheme. To address this challenge, we consider a potential-enstrophy dissipation scheme that conserves energy, building on existing variational integrators for the rotating shallow water equations [8, 20]. In particular, the scheme combines the variational discretization framework of [65] with the selective decay proposed in [37]. The resulting semi-discrete scheme is shown to conserve energy, suggesting this is a viable framework for long-term climate simulations.

Numerical results are presented comparing the new scheme with the variational integrator without dissipation and with a standard dissipation approach using a biharmonic eddy viscosity term. These simulations are carried out on both the f-plane and sphere, and we observe that the simulation with no dissipation becomes noisy and loses physical fidelity, in contrast to the simulations with dissipation. When analysing the conservation properties, we find that the enstrophy dissipating scheme conserves the energy as well as the scheme without dissipation, while

the biharmonic dissipation leads to a substantial loss of energy. Additionally, by computing energy and enstrophy spectra, we see that simulations using enstrophy dissipation better resolve small-scale features than those using biharmonic dissipation.

Natural extensions of this work are to more realistic models for geophysical flows. In particular, the primitive equations are a common first step in developing accurate simulations of climate dynamics. Since the framework in [37] also applies to 3D flows, combining the variational discretization framework with Casimir selective decay would lead to a discretization methodology for the primitive equations that would enable stabilized long-term integration schemes.

5.7. Acknowledgement

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, the NSERC Discovery Grant program, and the InnovateNL CRC Leverage R&D program. WB would like to acknowledge funding from NERC NE/R008795/1

5.8. Detailed computations

5.8.1. Continuous functional Casimir derivative

The variational derivative of the Casimir on a doubly periodic plane domain is computed as follows

$$\begin{aligned}\int_M \frac{\delta C}{\delta \mathbf{m}} \cdot \delta\mathbf{m}\,\mathrm{d}\mathbf{x} &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} C(\mathbf{m}+\varepsilon\delta\mathbf{m}, h) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{1}{2}\int_M hq(\mathbf{m}+\varepsilon\delta\mathbf{m}, h)^2 \mathrm{d}\mathbf{x} \\ &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{1}{2}\int_M \frac{\mathbf{z}\cdot\left(\mathrm{curl}(\frac{\mathbf{m}+\varepsilon\delta\mathbf{m}}{h})\right)^2}{h}\mathrm{d}\mathbf{x} \\ &= \int_M \frac{\mathbf{z}\cdot\mathrm{curl}(\frac{\mathbf{m}}{h})}{h}\cdot\mathrm{curl}\left(\frac{\delta\mathbf{m}}{h}\right)\mathrm{d}\mathbf{x} = \int_M q\mathbf{z}\cdot\mathrm{curl}\left(\frac{\delta\mathbf{m}}{h}\right)\mathrm{d}\mathbf{x} \\ &= \int_M \mathrm{curl}(q\mathbf{z})\cdot\frac{\delta\mathbf{m}}{h}\mathrm{d}\mathbf{x}.\end{aligned}$$

In the final step, we use the identity $\nabla\cdot(A\times B) = (\nabla\times A)\cdot B - A\cdot(\nabla\times B)$, noting that $\int_M \mathrm{div}\,\mathbf{u}\,\mathrm{d}\mathbf{x} = \int_{\partial M}\mathbf{u}\cdot\mathbf{n}\,\mathrm{d}S = 0$, since M is doubly periodic and, thus, has no boundary. This gives

$$\frac{\delta C}{\delta\mathbf{m}} = \frac{1}{h}\mathrm{curl}(q\mathbf{z}) = -\mathbf{z}\times(\nabla q)/h = (\partial_y q, -\partial_x q, 0)^\top/h, \tag{5.39}$$

where we first use the identity $\nabla\times(\psi A) = \psi(\nabla\times A) + \nabla\times A$ and, then, use the fact that $\mathbf{z}$ is the canonical unit vector in the z-direction.

Similarly, when M is a two dimensional Riemannian manifold, we compute

$$\begin{aligned}\int_M \frac{\delta C}{\delta\mathbf{m}}\cdot\delta\mathbf{m}\,\mathrm{d}\sigma &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\frac{1}{2}\int_M hq(\mathbf{m}+\varepsilon\delta\mathbf{m}, h)^2\mathrm{d}\sigma = \int_M hq\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} q(\mathbf{m}+\varepsilon\delta\mathbf{m}, h)\mathrm{d}\sigma \\ &= \int_M q\mathbf{d}\frac{\delta\mathbf{m}}{h}\mathrm{d}\sigma = \int_M \mathbf{d}\left(\frac{\delta\mathbf{m}}{h}q\right) + \int_M \frac{\delta\mathbf{m}}{h}\wedge\mathbf{d}q \\ &= -\int_M \frac{\delta\mathbf{m}}{h}\cdot(\star\mathbf{d}q)^\sharp\mathrm{d}\sigma,\end{aligned}$$

which gives $\frac{\delta C}{\delta\mathbf{m}} = -\frac{1}{h}(\star\mathbf{d}q)^\sharp$. In the computation above, we have used Stokes'

theorem on M, $\int_M \mathbf{d}\alpha = \int_{\partial M} \alpha = 0$ (since $\partial M = \varnothing$), and the identity that $\alpha \wedge \star\beta = (\alpha \cdot \beta^\sharp)\mathrm{d}\sigma$, for one-forms α, β on M, with $\star$ and $\sharp$ denoting the Hodge star and sharp operators associated with the Riemannian metric.

5.8.2. Discrete functional Casimir derivative

The semi-discrete variational derivative of the enstrophy is given by

$$
\begin{aligned}
\left\langle \frac{\delta \mathcal{C}}{\delta M}, \delta M \right\rangle_1 &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{C}(M + \varepsilon \delta M, h) \\
\mathrm{Tr}\left(\frac{\delta \mathcal{C}}{\delta M}^\top \Omega \delta M \right) &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \frac{1}{2} \sum_\zeta \frac{1}{h\zeta} \left(\sum_{\tilde{e}_{nm} \in \partial\zeta} \frac{M_{nm} + \varepsilon \delta M_{nm}}{|\zeta| h_{nm}} \right)^2 |\zeta| \\
\sum_i \sum_j \frac{\delta \mathcal{C}}{\delta M}^\top_{ij} (\Omega \delta M)_{ji} &= \sum_\zeta \frac{1}{h_\zeta} \left(\sum_{\tilde{e}_{nm} \in \partial\zeta} \frac{M_{nm}}{|\zeta| h_{nm}} \cdot \sum_{h_{nm} \in \partial\zeta} \frac{\delta M_{nm}}{|\zeta| h_{nm}} \right) |\zeta| \\
\sum_{ij} \frac{\delta \mathcal{C}}{\delta M}_{ij} \Omega_{ii} \delta M_{ij} &= \sum_\zeta \frac{1}{h_\zeta} \left(\sum_{\tilde{e}_{nm} \in \partial\zeta} \frac{M_{nm}}{|\zeta| h_{nm}} \cdot \sum_{h_{nm} \in \partial\zeta} \frac{\delta M_{nm}}{h_{nm}} \right) \\
\sum_i \sum_{j \in N(i)} \frac{\delta \mathcal{C}}{\delta M}_{ij} \Omega_{ii} \delta M_{ij} &= \sum_\zeta q_\zeta \sum_{h_{nm} \in \partial\zeta} \frac{\delta M_{nm}}{h_{nm}} \\
&= \sum_{ij} \frac{\delta M_{ij}}{h_{ij}} q_{\zeta_+} + \frac{\delta M_{ji}}{h_{ji}} q_{\zeta_-} \\
&= \sum_{ij} \frac{\delta M_{ij}}{h_{ij}} (q_{\zeta_+} - q_{\zeta_-}).
\end{aligned}
$$

In the second-to-last step, we use the property that each edge e_{ij} has 2 neighbouring vertices, denoted by ζ_+ and ζ_-. In the last step, we use the fact that the matrix M is anti-symmetric, as is δM, while h is symmetric.

6. Conclusion and outlook

In this thesis, to address the lack of existing research on the development of numerical variational integrators for the application on geophysical fluid dynamics, we have provided an extension of the variational discretization of the SWE to the sphere and advanced the model.

In particular, we have constructed a variational discretization in space and time for the SWE on a rotating sphere, following the discretization technique proposed in [65]. To verify the accuracy and conservation properties, we carried out standard benchmark tests proposed in [90]. Using reference solutions, we verified that the discretization converges, with order of accuracy between 0.5 and 1. The error in the conservation of the energy oscillates around a long-term mean, which reduces first order with the time step.

To improve the physics of the model, we introduced a stochastic version of the SWE under location uncertainty. We derived the new model using a stochastic Reynolds transport theorem, based on decomposing the fluid flow into a large-scale component and noise term modelling the unresolved scales. The resulting model preserves total energy for any realization. Thus, we combine the variational discretization of the SWE with approximations of the stochastic terms based on finite volume operators. We demonstrated that this approach preserves (spatially) the total energy, when performing an f-plane simulation with a homogeneous noise.

Using an inhomogeneous noise, we demonstrated that the proposed model better captures the structure of a reference flow than a comparable deterministic model. Furthermore, the forecast system of the proposed model is more reliable than a PIC system.

To be able to carry out long term simulations with coherent physical structures, we discretized the selective decay proposed in [37] using the variational framework proposed by [65]. With this new discrete framework, we obtained a potential enstrophy dissipating, but energy conserving, discretization of the SWE. We carried out standard benchmarks to demonstrate that the stabilized scheme produces coherent physical structures while the un-stabilized scheme becomes noisy and unphysical. When compared to a biharmonic dissipation, which dissipates energy and potential enstrophy, we observed that more small-scales are resolved. Hence, the simulation is more physical.

Future research is proposed in the following areas:

1. The variational integrator on the sphere proposed in chapter 3 conserves the lake-at-rest steady-state solution exactly over arbitrary bottom topography. This, and the fact that mass is conserved to machine precision, are important in tsunami propagation models. Thus, it would be interesting to extend the variational integrator with a wetting and drying interface to handle the inundation process.

2. The variational discretization framework has not been applied to more complicated geophysical fluid dynamics. A first step would be to extend the variational integrator for the shallow water equations to a multi-layer shallow water model. Moreover, other extensions are possible such as the thermal shallow water equations.

3. The stochastic shallow water model proposed in chapter 4 has an accurate

ensemble spread with an efficient rank histogram. This suggests that the model could be used for data assimilation.

4. The location uncertainty model (see chapter 4) conserves the total energy but not the enstrophy. In chapter 5, we designed a selective dissipation that conserves energy while dissipating enstrophy. Thus, it is relevant to combine the selective enstrophy decay with the RSW-LU model proposed in chapter 4.

5. To explore a fully energy conserving time integrator, the shallow water equations can be derived in the variational finite element framework [35]. The resulting model could overcome the energy drift of the selective enstrophy dissipation (see chapter 5).

A. GFD terms

In a general form, the equations of motion of a fluid on a domain M can be written as

$$\frac{D\rho}{Dt} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{A.1}$$

$$\rho \frac{D\boldsymbol{v}}{Dt} = -\nabla p + F, \tag{A.2}$$

where $\boldsymbol{v} = (u, v, w)$ is the velocity of the fluid, p its pressure, ρ its density, and F is the force per unit volume. Eq. (A.1) is the mass continuity equation and Eq. (A.2) is the momentum equation.

The vertical component of the momentum equation is of the form $\frac{Dw}{Dt} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g$, where g denotes the gravity. We say a fluid is in **hydrostatic balance** if $\frac{Dw}{Dt} = 0$, which gives $\frac{\partial p}{\partial z} = -\rho g$. Let a parcel of fluid in hydrostatic balance be displaced, then the sum of the gravity and vertical pressure gradient forces can either reinforce the initial balance or work against it. The frequency of that oscillation is called **buoyancy frequency**.

Now, we add the Coriolis acceleration f to the horizontal part of Eq. (A.2) and obtain

$$\frac{D\boldsymbol{v}}{Dt} + f \times \boldsymbol{v} = -\frac{1}{\rho}\nabla p, \qquad \boldsymbol{v} = (u, v). \tag{A.3}$$

The fluid is in **geostrophic balance** if the pressure term balances the rotation term: $f \times \boldsymbol{v} = -\frac{1}{\rho}\nabla p$.

We define the **potential vorticity** for the SWE as $\omega := \frac{\nabla \times \boldsymbol{v}}{\rho}$. The numerator of the fraction is the vorticity $\widetilde{\omega} = \nabla \times \boldsymbol{v}$. Another important quantity is the **enstrophy** $\int_M |\widetilde{\omega}|^2 \, dM$. By taking the curl of Eq. (A.3), and assuming geostrophic balance, we obtain the advection equation

$$\frac{D\omega}{Dt} = (\omega \cdot \nabla)\boldsymbol{v}.$$

The circulation around a fluid parcel bounded by the curve γ is defined as $\oint_\gamma v \, dr$. Using Stokes' theorem, we relate the circulation to the vorticity $\oint_\gamma v \, dr = \int_S \widetilde{\omega} \cdot \boldsymbol{n} \, dS$, here S is the area enclosed by γ. Kelvin's circulation theorem states that

$$0 = \frac{D}{Dt} \oint_\gamma v \, dr.$$

Next, we look at the **energy spectrum**. We decompose the velocity field into its Fourier components

$$v(x,t) = \sum_k \bar{v}(k,t) e^{ik \cdot x}.$$

Then, the kinetic energy E can be written as the energy spectrum $\mathcal{E}$ in the Fourier space:

$$E = \frac{1}{2}\int_M \boldsymbol{v}^2 = \frac{1}{2}\int |\bar{\boldsymbol{v}}|^2 = \int \mathcal{E}(k).$$

In the case of forced dissipative turbulence, the energy spectrum has the general form $\mathcal{E}(k) = \varepsilon^{\frac{2}{3}} k^{-\frac{5}{3}}$ multiplied by a constant, where ε is the energy flux (see [86] for details). Since $\mathcal{E}$ decreases by $k^{-5/3}$, large scale features (represented by small wave numbers) break into small scale features (represented by high wave numbers), see Fig. A.1. The energy spectrum is related to the enstrophy spectrum $\mathcal{Z}$ by

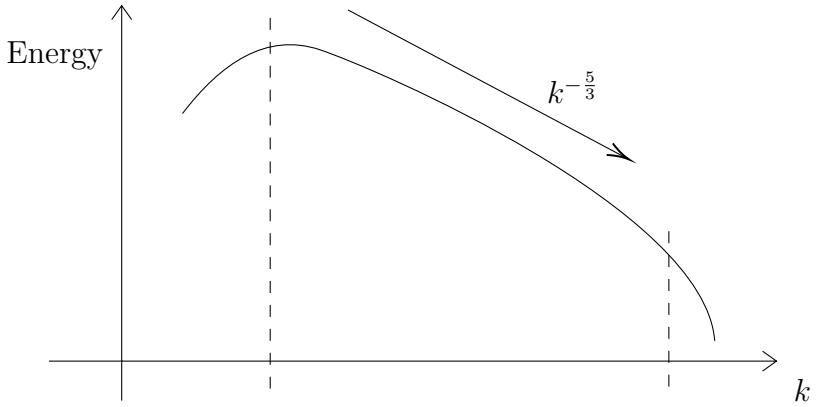


Figure A.1.: The energy spectrum. The first dashed line marks the energy input, and the second one the wave number where viscosity starts to act.

$k^2\mathcal{Z} = \mathcal{E}$, which shows that the enstrophy spectrum moves towards small wave numbers.

Bibliography

[1] M. Aechtner, N. K.-R. Kevlahan, and T. Dubos. "A conservative adaptive wavelet method for the shallow-water equations on the sphere". In: *Quarterly Journal of the Royal Meteorological Society* 141.690 (2015), pp. 1712–1726.

[2] JS Allen, JA Barth, and PA Newberger. "On intermediate models for barotropic continental shelf and slope flow fields. Part I: formulation and comparison of exact solutions". In: *Journal of Physical Oceanography* 20.7 (1990), pp. 1017–1042.

[3] J. Anderson and S. Anderson. "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts". In: *Monthly Weather Review* 127.12 (1999), pp. 2741–2758.

[4] A. Arakawa and Y.J.G. Hsu. "Energy conserving and potential-enstrophy dissipating schemes for the shallow water equations". In: *Monthly Weather Review* 118.10 (1990), pp. 1960–1969.

[5] A. Arakawa and V.R. Lamb. "A potential enstrophy and energy conserving scheme for the shallow water equations". In: *Monthly Weather Review* 109.1 (1981), pp. 18–36.

[6] B. K. Arbic, K. L. Polzin, R. B. Scott, J. G. Richman, and J. F. Shriver. "On eddy viscosity, energy cascades, and the horizonal resolution of gridded stallite altimeter products". In: *Journal of Physical Oceanography* 43.2 (2013), pp. 283–300.

[7] W. Bauer. “Toward goal-oriented r-adaptive models in geophysical fluid dynamics using a generalized discretization approach”. PhD thesis. Department of Geosciences, University of Hamburg, 2013.

[8] W. Bauer and F. Gay-Balmaz. “Towards a geometric variational discretization of compressible fluids: The rotating shallow water equations”. In: *Journal of Computational Dynamics* 6.1 (2017), pp. 1–37.

[9] W. Bauer and F. Gay-Balmaz. “Variational integrators for anelastic and pseudo-incompressible flows”. In: (2017). arXiv preprint arXiv:1701.06448.

[10] W. Bauer, P. Chandramouli, B. Chapron, L. Li, and E. Mémin. “Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach”. In: *Journal of Physical Oceanography* 50.4 (2020), pp. 983–1003.

[11] W. Bauer, P. Chandramouli, L. Li, and E. Mémin. “Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models”. In: *Ocean Modelling* 151 (2020), p. 101646.

[12] M. Beck, M. J. Gander, and F. Kwok. “B-Methods for the numerical solution of evolution problems with blow-up solutions. Part I: variation of the constant”. In: *SIAM Journal of Scientific Computing* 37.6 (2015), A2998–A3029.

[13] J. Berner, S-Y. Ha, J.P. Hacker, A. Fournier, and C. Snyder. “Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations”. In: *Monthly Weather Review* 139.6 (2011), pp. 1972–1995.

[14] A. Bihlo and R. O. Popovych. “Invariant discretization schemes for the shallow-water equations”. In: *SIAM Journal of Scientific Computing* 34.6 (2012), B810–B839.

[15] A. Bihlo and F. Valiquette. “Symmetry-preserving numerical schemes”. In: *Symmetries and Integrability of Difference Equations*. Springer, 2017, pp. 261–324.

[16] S. Blanes and F. Casas. *A concise introduction to geometric numerical integration*. Vol. 23. Boca Raton: CRC Press, 2016.

[17] L. Bonaventura and T. D. Ringler. “Analysis of discrete shallow-water models on geodesic Delaunay grids with C-type staggering”. In: *Monthly Weather Review* 133.8 (2005), pp. 2351–2373.

[18] N. Bou-Rabee and J. E. Marsden. “Hamilton–Pontryagin integrators on Lie groups. Part I: introduction and structure-preserving properties”. In: *Foundations of Computational Mathematics* 9 (2009), pp. 197–219.

[19] R. Brecht, L. Li, W. Bauer, and E. Mémin. “Rotating shallow water flow under location uncertainty with a structure-preserving discretization”. In: *arXiv preprint arXiv:2102.03783* (2021).

[20] R. Brecht, W. Bauer, A. Bihlo, F. Gay-Balmaz, and S. MacLachlan. “Variational integrator for the rotating shallow-water equations on the sphere”. In: *Quarterly Journal of the Royal Meteorological Society* 145.720 (2019), pp. 1070–1088.

[21] C. Budd and V. A. Dorodnitsyn. “Symmetry-adapted moving mesh schemes for the nonlinear Schrödinger equation”. In: *Journal of Physics A* 34.48 (2001), pp. 10387–10400.

[22] B. Chapron, P. Dérian, E. Mémin, and V. Resseguier. “Large-scale flows under location uncertainty: a consistent stochastic framework”. In: *Quarterly Journal of the Royal Meteorological Society* 144.710 (2018), pp. 251–260.

[23] Q. Chen, M. Gunzburger, and T. Ringler. "A scale-invariant formulation of the anticipated potential vorticity method". In: *Monthly Weather Review* 139.8 (2011), pp. 2614–2629.

[24] C. J. Cotter and A. T. T. McRae. "Compatible finite element methods for numerical weather prediction". In: *arXiv preprint arXiv:1401.0616v1* (2014).

[25] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2014.

[26] M. Desbrun, E. Gawlik, F. Gay-Balmaz, and V. Zeitlin. "Variational discretization for rotating stratified fluids". In: *Disc. Cont. Dyn. Syst. Series A* 34 (2014), pp. 479–511.

[27] C. Eldred and D. Randall. "Total energy and potential enstrophy conserving schemes for the shallow water equations using Hamiltonian methods-Part 1: derivation and properties". In: *Geoscientific Model Development* 10 (2017), pp. 791–810.

[28] N. Flyer and G. B. Wright. "A radial basis function method for the shallow water equations on a sphere". In: *Proceedings of the Royal Society A* 465.2106 (2009), pp. 1949–1976.

[29] N. Flyer, E. Lehto, S. Blaise, G.B. Wright, and A. St-Cyr. "A guide to RBF-generated finite differences for nonlinear transport: shallow water simulations on a sphere". In: *Journal of Computational Physics* 231.11 (2012), pp. 4078–4095.

[30] J. Frank, G. Gottwald, and S. Reich. "A hamiltonian particle-mesh method for the rotating shallow-water equations". In: *Meshfree Methods for Partial Differential Equations*. Ed. by M. Griebel, M.A. Schweitzer, T.J. Barth, M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose, and T. Schlick. Vol. 26.

Lecture Notes in Computational Science and Engineering. Berlin: Springer, 2002, pp. 131–142.

[31] C. E. Franzke and A. J. Majda. “Low-order stochastic mode reduction for a prototype atmospheric GCM”. In: *Journal of the Atmospheric Sciences* 63.2 (2006), pp. 457–479.

[32] C. E. Franzke, T. J. O’Kane, J. Berner, P. D. Williams, and V. Lucarini. “Stochastic climate theory and modeling”. In: *Wiley Interdisciplinary Reviews: Climate Change* 6.1 (2015), pp. 63–78.

[33] J. S. Frederiksen, T. J. O’Kane, and M. J. Zidikheri. “Subgrid modelling for geophysical flows”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1982 (2013), p. 20120166.

[34] J. Galewsky, R. Scott, and L. Polvani. “An initial-value problem for testing numerical models of the global shallow-water equations”. In: *Tellus Series A-dynamic Meteorology and Oceanography* 56.5 (2004), pp. 429–440.

[35] E.S. Gawlik and F. Gay-Balmaz. “A variational finite element discretization of compressible flow”. In: *Found Comput Math* (2020).

[36] E.S. Gawlik, P. Mullen, D. Pavlov, J.E. Marsden, and M. Desbrun. “Geometric, variational discretization of continuum theories”. In: *Physica D: Nonlinear Phenomena* 240.21 (2011), pp. 1724–1760.

[37] F. Gay-Balmaz and D.D. Holm. “Selective decay by Casimir dissipation in inviscid fluids”. In: *Nonlinearity* 26.2 (2013), p. 495.

[38] F. Gay-Balmaz and Darryl D. Holm. “A geometric theory of selective decay with applications in MHD”. In: *Nonlinearity* 27 (2014), pp. 1747–1777.

[39] G. Gottwald, D. T. Crommelin, and C. E. Franzke. “Stochastic climate theory”. In: *Nonlinear and Stochastic Climate Dynamics*. Cambridge University Press, 2017, pp. 209–240.

[40] G. Gottwald and J. Harlim. “The role of additive and multiplicative noise in filtering complex dynamical systems”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 469.2155 (2013), p. 20130096.

[41] I. Grooms and A. J. Majda. “Stochastic superparameterization in quasi-geostrophic turbulence”. In: *Journal of Computational Physics* 271 (2014), pp. 78–98.

[42] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*. Vol. 31. Springer Science & Business Media, 2006.

[43] T. M. Hamill. “Interpretation of rank histograms for verifying ensemble forecasts”. In: *Monthly Weather Review* 129 (2001), pp. 550–560.

[44] R. Heikes and D.A. Randall. “Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part I: basic design and results of tests”. In: *Monthly Weather Review* 123.6 (June 1995), pp. 1862–1880.

[45] R. Heikes and D.A. Randall. “Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part II. A detailed description of the grid and an analysis of numerical accuracy”. In: *Monthly Weather Review* 123.6 (June 1995), pp. 1881–1887.

[46] D. D Holm, J. E. Marsden, and T. S. Ratiu. “The Euler-Poincaré equations in geophysical fluid dynamics”. In: *arXiv:chao-dyn/9903035v1* (1999).

[47] D.D. Holm. “Applications of Poisson geometry to physical problems”. In: *Summer School and Conference on Poisson Geometry, ICTP*. Trieste, Italy, 4-22 July. 2005.

[48] D.D. Holm, J.E. Marsden, and T.S. Ratiu. “The Euler-Poincaré equations and semidirect products with applications to continuum theories”. In: *Advances in Mathematics* 137 (1998), pp. 1–81.

[49] S. Kadri Harouna and E. Mémin. “Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling”. In: *Computers and Fluids* 156 (2017), pp. 456–469.

[50] J. Kjellsson and L. Zanna. “The impact of horizontal resolution on energy transfers in global ocean models”. In: *Fluids* 2.3 (2017), p. 45.

[51] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Vol. 23. Springer-Verlag Berlin Heidelberg, 1992.

[52] R. H. Kraichnan. “Inertial ranges in two-dimensional turbulence”. In: *Physics of Fluids* 10.7 (1967), pp. 1417–1423.

[53] H. Kunita. *Stochastic flows and stochastic differential equations*. Vol. 24. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1997.

[54] J. M. Lee. *Riemannian manifolds: an introduction to curvature*. Vol. 176. Graduate Texts in Mathematics. Springer, New York, 1997.

[55] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*. Cambridge: Cambridge University Press, 2004.

[56] D. K. Lilly. “Numerical simulation of developing and decaying two-dimensional turbulence”. In: *Journal of Fluid Mechanics* 45.2 (1971), pp. 395–415.

[57] A. Majda, C. Franzke, and B. Khouider. “An applied mathematics perspective on stochastic modelling for climate”. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 366.1875 (2008), pp. 2427–2453.

[58] J. E. Marsden and M. West. “Discrete mechanics and variational integrators”. In: *Acta Numerica 2001* 10 (2001), pp. 357–514.

[59] W.H. Matthaeus and D. Montgomery. “Selective decay hypothesis at high mechanical and magnetic Reynolds numbers”. In: *New York Academy of Sciences, Annals* 357 (1980), pp. 203–222.

[60] A. T. T. McRae and C. J. Cotter. “Energy-and enstrophy-conserving schemes for the shallow-water equations, based on mimetic finite elements”. In: *Q. J. R. Meteorol. Soc.* 140 (2014), pp. 2223–2234.

[61] J.C. McWilliams. *Fundamentals of geophysical fluid dynamics*. Cambridge University Press, New York, 2006, 266 pages.

[62] E. Mémin. “Fluid flow dynamics under location uncertainty”. In: *Geophysical & Astrophysical Fluid Dynamics* 108.2 (2014), pp. 119–146.

[63] P.J. Olver. *Applications of Lie groups to differential equations*. Vol. 107. Springer Science & Business Media, 2012.

[64] N. Paldor and A. Sigalov. “A unified linear wave theory of the shallow water equations on a rotating plane”. In: *IUTAM Symposium on Hamiltonian Dynamics, Vortex Structures, Turbulence*. Springer. 2008, pp. 403–413.

[65] Dmitry Pavlov, Patrick Mullen, Yiying Tong, Eva Kanso, Jerrold E Marsden, and Mathieu Desbrun. “Structure-preserving discretization of incompressible fluids”. In: *Physica D: Nonlinear Phenomena* 240.6 (2011), pp. 443–458.

[66] J. Perot, D. Vidovic, and P. Wesseling. “Mimetic reconstruction of vectors”. In: *Compatible Spatial Discretizations*. Ed. by Douglas N. Arnold, Pavel B. Bochev, Richard B. Lehoucq, Roy A. Nicolaides, and Mikhail Shashkov. Vol. 142. The IMA Volumes in Mathematics and Its Applications. Springer New York, 2006, pp. 173–188. ISBN: 978-0-387-38034-6.

[67] S. Pope. *Turbulent flows*. Cambridge University Press, 2000.

[68] D. A. Randall. “Geostrophic Adjustment and the Finite-Difference Shallow-Water Equations”. In: *Monthly Weather Review* 122.6 (1994), pp. 1371–1377.

[69] V. Resseguier, E. Mémin, and B. Chapron. “Geophysical flows under location uncertainty, part I: Random transport and general models”. In: *Geophysical & Astrophysical Fluid Dynamics* 111.3 (2017), pp. 149–176.

[70] V. Resseguier, E. Mémin, and B. Chapron. “Geophysical flows under location uncertainty, part II: quasi-geostrophic models and efficient ensemble spreading”. In: *Geophysical & Astrophysical Fluid Dynamics* 111.3 (2017), pp. 177–208.

[71] V. Resseguier, E. Mémin, and B. Chapron. “Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence”. In: *Geophysical & Astrophysical Fluid Dynamics* 111.3 (2017), pp. 209–227.

[72] V. Resseguier, L. Li, G. Jouan, P. Derian, E. Mémin, and B. Chapron. “New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid computational fluid dynamics”. In: *Archives of Computational Methods in Engineering* (2020), pp. 1886–1784.

[73] T. D. Ringler and D. A. Randall. “A potential enstrophy and energy conserving numerical scheme for solution of the shallow-water equations on a geodesic grid”. In: *Monthly Weather Review* 130.5 (2002), pp. 1397–1410.

[74] T.D. Ringler, J. Thuburn, J.B. Klemp, and W.C. Skamarock. “A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured C-grids”. In: *Journal of Computational Physics* 229.9 (2010), pp. 3065–3090.

[75] P. Rípodas, A. Gassmann, J. Förstner, D. Majewski, M. Giorgetta, P. Korn, L. Kornblueh, H. Wan, G. Zängl, L. Bonaventura, and T. Heinze. “Icosahedral Shallow Water Model (ICOSWM): results of shallow water test cases and sensitivity to model parameters”. In: *Geoscientific Model Development* 2.2 (2009), pp. 231–251.

[76] R. Sadourny and C. Basdevant. “Parameterization of subgrid scale barotropic and baroclinic eddies in quasi-geostrophic models: anticipated potential vorticity method”. In: *Journal of Atmospheric Sciences* 42.13 (1985), pp. 1353–1363.

[77] J.-M. Sanz-Serna and M.-P. Calvo. *Numerical Hamiltonian problems*. Vol. 7. Applied Mathematics and Mathematical Computation. London: Chapman & Hall, 1994.

[78] L. Sirovich. “Turbulence and the dynamics of coherent structures, part I: Coherent structures.” In: *Quarterly of Applied Mathematics* 45.3 (1987), pp. 561–571.

[79] A. Staniforth and J. Thuburn. “Horizontal grids for global weather and climate prediction models: a review”. In: *Quarterly Journal of the Royal Meteorological Society* 138.662 (2012), pp. 1–26.

[80] O. Talagrand, R. Vautard, and B. Strauss. “Evaluation of probabilistic prediction systems”. In: Workshop on Predictability. ECMWF, 1997.

[81] J. Thuburn. “Numerical wave propagation on the hexagonal C-grid”. In: *Journal of Computational Physics* 227 (May 2008), pp. 5836–5858.

[82] J. Thuburn and C. J. Cotter. “A primal–dual mimetic finite element scheme for the rotating shallow water equations on polygonal spherical meshes”. In: *Journal of Computational Physics* 290 (2015), pp. 274–297.

[83] J. Thuburn and C.J. Cotter. “A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids”. In: *SIAM Journal on Scientific Computing* 34.3 (2012), B203–B225.

[84] J. Thuburn, T. D. Ringler, W. C. Skamarock, and J. B. Klemp. “Numerical representation of geostrophic modes on arbitrarily structured C-grids”. In: *Journal of Computational Physics* 228 (Dec. 2009), pp. 8321–8335.

[85] P.A. Ullrich, C. Jablonowski, and B. van Leer. “High-order finite-volume methods for the shallow-water equations on the sphere”. In: *Journal of Computational Physics* 229.17 (2010), pp. 6104–6134.

[86] G.K. Vallis. *Atmospheric and oceanic fluid dynamics*. Cambridge University Press, 2017.

[87] A. T. S. Wan, A. Bihlo, and J.-C. Nave. “Conservative methods for dynamical systems”. In: *SIAM Journal on Numerical Analysis* 55.5 (2017), pp. 2255–2285.

[88] A. T. S. Wan and J.-C. Nave. “On the arbitrarily long-term stability of conservative methods”. In: (2016).

[89] A. P. Weigel. “Ensemble forecasts”. In: *Forecast Verification*. John Wiley and Sons, Ltd, 2012. Chap. 8, pp. 141–166.

[90] D.L. Williamson, J.B. Drake, J.J. Hack, R. Jakob, and P.N. Swarztrauber. “A standard test set for numerical approximations to the shallow water equations in spherical geometry”. In: *Journal of Computational Physics* 102.1 (1992), pp. 211–224.