Research Article Open Access

Federica Gugole* and Christian L. E. Franzke

# Numerical Development and Evaluation of an Energy Conserving Conceptual Stochastic Climate Model

https://doi.org/10.1515/mcwf-2019-0004
Received November 8, 2018; accepted April 18, 2019

**Abstract:** In this study we aim to present the successful development of an energy conserving conceptual stochastic climate model based on the inviscid 2-layer Quasi-Geostrophic (QG) equations. The stochastic terms have been systematically derived and introduced in such a way that the total energy is conserved. In this proof of concept study we give particular emphasis to the numerical aspects of energy conservation in a high-dimensional complex stochastic system and we analyze what kind of assumptions regarding the noise should be considered in order to obtain physical meaningful results. Our results show that the stochastic model conserves energy to an accuracy of about 0.5% of the total energy; this level of accuracy is not affected by the introduction of the noise, but is mainly due to the level of accuracy of the deterministic discretization of the QG model. Furthermore, our results demonstrate that spatially correlated noise is necessary for the conservation of energy and the preservation of important statistical properties, while using spatially uncorrelated noise violates energy conservation and gives unphysical results. A dynamically consistent spatial covariance structure is determined through Empirical Orthogonal Functions (EOFs). We find that only a small number of EOFs is needed to get good results with respect to energy conservation, autocorrelation functions, PDFs and eddy length scale when comparing a deterministic control simulation on a $512 \times 512$ grid to a stochastic simulation on a $128 \times 128$ grid. Our stochastic approach has the potential to seamlessly be implemented in comprehensive weather and climate prediction models.

**Keywords:** stochastic parameterization, energy conservation, projection operator, spatial noise structure, Empirical Orthogonal Functions

**MSC:** 65C20,68U20

## 1 Introduction

The dynamics of the atmosphere and the oceans are by nature complex. Processes with different time and length scales interact with each other affecting the system as a whole. While climate and ocean models have considerably improved over the last few decades, we still cannot resolve all important physical scales and processes, see for instance [20, 6, 22]. The discretization of the continuous governing equations of motion is limited by the model resolution, which determines the size of the smallest resolvable scale. Despite the continued increase of computer power and, thus, of resolution, there are still many important processes in the atmosphere and in the oceans that cannot be explicitly resolved. These include turbulent motions with scales ranging from a few centimeters to the size of the model grid box, as well as processes that occur at a molecular

***Corresponding Author: Federica Gugole:** Meteorological Institute and Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany; E-mail: federica.gugole@uni-hamburg.de
**Christian L. E. Franzke:** Meteorological Institute and Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany

scale, like condensation and evaporation. Any numerical forecaster or modeller has to make a decision, based on the targeted objectives, regarding the spatial and temporal scales to resolve. As a consequence of this decision, each numerical scheme inevitably fails to resolve subgrid-scale processes.

These unresolved processes and scales cause many of the observed differences between models and observations. In order to represent these unresolved processes, so-called parameterizations are necessary which take into account the influence that the unresolved have on the resolved processes, if they would be resolved in high-resolution simulations [20]. Most parameterizations, however, are damping and do not take account of the energy and momentum fluxes from the unresolved to the resolved scales [34, 57]. This is a likely source of many of the observed biases in climate and ocean models. Without the added dissipation, however, energy and enstrophy would accumulate at the truncation scale and lead to a blow up in finite time. Hence, it is of fundamental importance to find systematic ways to parameterize the unresolved scales and processes of models, and to improve the model performance and reduce the model biases at coarser resolutions.

As suggested already in 1976 by Hasselmann [31], fast waves can be considered as noise with respect to the slowly evolving large-scale modes and, therefore, can be parameterized by stochastic processes [23]. Hence, to ameliorate this problem of too large damping, stochastic parameterization schemes have been developed (see recent reviews [20, 6, 22, 27]). There are several advantages in using stochastic parameterizations; the most important are: gain in computational time compared to higher resolved simulations; reduction of model errors and systematic representation of uncertainties and model errors. Most operational stochastic parameterization schemes are rather *ad hoc* developments [7] and do not conserve energy or momentum. As a consequence, current schemes have the disadvantage of the forfeiture of conservation laws and a likely loss of important symmetries in the model equations. For climate simulations conservation properties are of importance because they are leading to stable and realistic climate simulations, and should be considered also for stochastic parameterizations, not only for stability reasons, but also to respect the underlying dynamics of geophysical fluid flows.

From a mathematical perspective, there has been a growing interest over the last few decades in finding appropriate techniques to develop systematic methods to accurately and efficiently represent fast variables in multi-scale systems. Much fundamental work has already been done, e.g. Mémin [37] derived energy conserving geophysical fluid equations assuming that the velocity can be written as a mean state plus some perturbations, while Holm [32] used stochastic variational principles to obtain new stochastic fluid equations conserving helicity and the Kelvin circulation theorem. Numerical models following these theoretical approaches have been developed and show good performance and improved results with respect to the deterministic counterpart at the same resolution, see for example [51, 52] for numerical implementations and results of [37] and [12, 11] for applications of [32].

Furthermore, Majda, Timofeyev and Vanden-Eijnden [43, 42, 44, 39, 38] developed a systematic strategy for stochastic mode reduction starting from the assumption that the explicit nonlinear self-interaction of the fastest scales involved can be represented by a linear stochastic operator. This procedure is mathematically justified only for large time scale separation but showed good performances also in case of a less pronounced time scale separation. For its application to more complex atmospheric models see also [26, 19, 18, 50, 41]. A less theoretical, but still efficient, approach is given by the stochastic kinetic energy backscatter scheme (SKEBS) which is often used to represent model uncertainty arising from unresolved subgrid-scale processes and their interactions with larger scales [56, 5, 22, 47, 21]. Connected to the idea of backscatter, different deterministic and stochastic parameterizations aiming at representing the upscale turbulent cascades in eddy-permitting simulations have been developed. Among others, noticeable examples are given by [33, 49, 62, 28, 15], which also showed that the stochastic backscatter is, in general, a more efficient eddy parameterization than its deterministic counterpart.

In this paper we systematically further develop the projector approach outlined by Frank and Gottwald [16]. Here the stochastic noise is projected onto the energy manifold. This approach has the main advantage that it can be straight forwardly implemented in existing models while the above approaches [37, 32, 43] derive new equations of motions which are harder to implement in already existing and operational models. Frank and Gottwald [16] tested their scheme with a 4-dimensional toy-model for the large-scale dynamics of the atmosphere by means of a Lagrangian description of the dynamics. Here instead we apply it to a high-

dimensional conceptual climate model, i.e. the inviscid 2-layer QG model, in a Eulerian framework. The purpose of our study is a proof of concept whether this scheme can also be applied to high-dimensional models. Thus, the main research questions we aim to address in this study is: (i) Can we accurately conserve energy in a high-dimensional QG model? and (ii) What conditions do we need to impose on the spatial noise covariance matrix for this? Hence, in this work we focus on the technical aspects of this approach, analyzing potential issues due to the discretization of the continuous equations or to the numerical implementation in general. With this intention, we choose to apply this projection operator scheme to an energy conserving QG model as a hard numerical test case. While our particular set up might not seem interesting from a geophysical fluid dynamics point of view, we still consider it numerically challenging and hence a valuable benchmark in testing the numerical aspects and accuracy of our stochastic system in a high-dimensional geophysical model.

The remainder of this paper is organized as follows: in Section 2 we present the inviscid 2-layer energy conserving QG model both in its deterministic and stochastic formulations. Details on the numerics, like the choice of the numerical solvers and the definition of the spatio-temporal noise structure are provided in Section 3. Section 4 displays and discusses the outcomes of our stochastic model experiments. Finally in Section 5, we give a brief summary of our findings and outlook of future research directions.

# 2 The 2-Layer QG Model

## 2.1 The deterministic equations

As mentioned above, we start from the non-dimensional inviscid 2-layer QG equations presented in [59] on a $\beta$-plane with double-periodic boundary conditions

$$\frac{\partial q_B}{\partial t} = -J(\psi_B, q_B) - J(\psi_T, q_T)\,, \tag{1a}$$

$$\frac{\partial q_T}{\partial t} = -J(\psi_T, q_B) - J(\psi_B, q_T)\,, \tag{1b}$$

$$q_B = \nabla^2\psi_B + \beta y\,, \tag{2a}$$

$$q_T = \nabla^2\psi_T - k_d^2\psi_T\,, \tag{2b}$$

where $q_B = q_B(\mathbf{x}, t)$, $\psi_B = \psi_B(\mathbf{x}, t)$ and $q_T = q_T(\mathbf{x}, t)$, $\psi_T = \psi_T(\mathbf{x}, t)$ represent, respectively, potential vorticity (PV) and streamfunction of the barotropic and baroclinic mode on the horizontal plane $\mathbf{x} \in \mathbb{R}^2$ at time $t \in \mathbb{R}$, $\nabla$ is the horizontal gradient, $\nabla^2$ the Laplacian and $J$ the Jacobian operator

$$J(A, B) = \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x}\,.$$

Since we employ a non-dimensional description, the domain has been rescaled to a $2\pi \times 2\pi$ square. Double periodic QG models on a $\beta$-plane have been widely used in theoretical studies [8, 9, 29, 45]. Here we consider layers of equal thickness and the parameter $k_d$ (given by the relation $k_d^2/2 = \left(2f_0/Nh\right)^2$ where $N = 1.2 \cdot 10^{-2}$ is the Brunt-Väisälä frequency, $h$ the mean depth of the layers and $f \approx f_0 + \beta y$ the approximate Coriolis term where $f_0 = 1$ and $\beta = 0.509$) determines the strength of the shear between the two layers and hence also the intensity of the baroclinic instability. For most simulations in this study we are using a deformation radius of about 0.042 non-dimensional units and, thus, are in an ocean like eddy-permitting regime. In this setting, one non-dimensional time unit corresponds to roughly 2.5 days.

Since we want to focus on the energy conservation properties of our numerical scheme we do not introduce terms accounting for external forcing or eddy dissipation, and instead consider an initial value problem. We want to stress, that we want to focus in this study on the numerical and accuracy aspects of energy conservation in a stochastic setting and not on geophysical flow realism (see section 2.2 below).

The system given by equations (1)-(2) conserves its total energy $E$ and enstrophy $Z$:

$$E(q_B, q_T) = \frac{1}{2}\int_A \left[(\nabla\psi_B)^2 + (\nabla\psi_T)^2 + k_d^2{\psi_T}^2\right] dA \ , \quad Z(q_B, q_T) = \frac{1}{2}\int_A \left(q_B^2 + q_T^2\right) dA \ .$$

Details about conservation properties and how to derive them can be found in [59].

The Hamiltonian $H$ of the system is given by its total energy, thus it reads

$$H(q_B, q_T) = \frac{1}{2}\int_A \left[(\nabla\psi_B)^2 + (\nabla\psi_T)^2 + k_d^2{\psi_T}^2\right] dA \ . \tag{3}$$

It can be shown that

$$\begin{aligned}
\delta H &= +\int_A \left(\nabla\psi_B\cdot\delta\nabla\psi_B + \nabla\psi_T\cdot\delta\nabla\psi_T + k_d^2\psi_T\cdot\delta\psi_T\right) dA \\
&= -\int_A \left(\psi_B\cdot\delta\nabla^2\psi_B + \psi_T\cdot\delta\nabla^2\psi_T - \psi_T\cdot k_d^2\delta\psi_T\right) dA \\
&= -\int_A (\psi_B\cdot\delta q_B + \psi_T\cdot\delta q_T)\, dA \ ,
\end{aligned}$$

which implies

$$\frac{\partial H}{\partial q_B} = -\psi_B \ , \quad \frac{\partial H}{\partial q_T} = -\psi_T \ .$$

For a general review of Hamiltonian mechanics and its application to geophysical fluid dynamics see [2] and [55]. The following notation will be employed

$$A : B = a_{ij}b_{ij} = Tr(AB^T) \ .$$

## 2.2 The stochastic formulation

In this section we derive a stochastic energy conserving version of the 2-layer QG equations. In our formulation we include unresolved fast sub-grid processes by means of a stochastic forcing, modeled as an Ornstein-Uhlenbeck process, which we assume to act first on the baroclinic mode and then, because of the coupling between the two modes, to affect also the slower barotropic mode. For this choice we relate to the idea of backscatter, where energy goes from the smaller scales back into the larger processes. Therefore, we add a simple 2-dimensional stochastic field only to the fast baroclinic mode. The source terms so introduced would lead the dynamics to leave the manifold of constant energy on which the deterministic model (1)-(2) evolves. In order to balance the stochastic fluctuations that would affect the aforementioned manifold, we introduce an auxiliary 2-dimensional stochastic process $dY_t$. This procedure follows the method introduced by Frank and Gottwald in [16], with the difference that we consider a high-dimensional system in a Eulerian framework, and not a 4-dimensional system with a Lagrangian description. The following set of equations is therefore proposed:

$$dq_B = \left(-J(\psi_B, q_B) - J(\psi_T, q_T)\right) dt \ , \tag{4a}$$

$$dq_T = \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt - \Gamma q_T dt + \Sigma dW_t + dY_t \ , \tag{4b}$$

$$dY_t = S_t dW_t + B_t dt \ , \tag{4c}$$

where $dW_t$ denotes a 2-dimensional Wiener process, $\Gamma, \Sigma, S_t \in \mathbb{R}^{2\times 2}$ and $B_t \in \mathbb{R}^2$. The choice of adding the stochastic terms on the equation of the baroclinic PV not only reconnects to the concept of backscatter, but allows potentially also the application of stochastic mode reduction, as Frank and Gottwald did in their work

[16]. For instance, one could derive a reduced order stochastic model for just the barotropic modes [43, 42, 44, 19, 18].

Instead of dealing with two different stochastic processes, we want to write $S_t$ and $B_t$ as functions of $\Sigma$ and $\Gamma$. For that purpose, we write the increment of $H$ as a sum of two parts, a deterministic part $\mu_H$ including all the terms multiplied by $dt$, and a stochastic part $\sigma_H$ containing those with the Wiener process. By Ito's theorem we have

$$\begin{aligned} dH &= \frac{\partial H}{\partial q_B} \cdot dq_B + \frac{\partial H}{\partial q_T} \cdot dq_T + \frac{1}{2} \frac{\partial^2 H}{\partial q_T \partial q_T} : dq_T dq_T^T \\ &= -\psi_B \cdot \left(-J(\psi_B, q_B) - J(\psi_T, q_T)\right) dt \\ &\quad - \psi_T \cdot \left(-J(\psi_B, q_T) - J(\psi_T, q_B) + B_t\right) dt \\ &\quad + \left(\frac{1}{2} \frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma + S_t)(\Sigma + S_t)^T\right) dt \\ &\quad - \psi_T \cdot (-\Gamma q_T)\, dt - \psi_T \cdot (\Sigma + S_t)\, dW \\ &= \mu_H dt + \sigma_H dW_t \,, \end{aligned}$$

where the transposed is denoted by the superscript $^T$, and

$$\begin{aligned} \mu_H &= \psi_B \cdot \left(J(\psi_B, q_B) + J(\psi_T, q_T)\right) + \psi_T \cdot \left(J(\psi_T, q_B) + J(\psi_B, q_T) - B_t\right) \\ &\quad + \psi_T \cdot \Gamma q_T + \frac{1}{2} \frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma + S_t)(\Sigma + S_t)^T \\ &= -\nabla_{q_B} H \cdot \left(J(\psi_B, q_B) + J(\psi_T, q_T)\right) - \nabla_{q_T} H \cdot \left(J(\psi_T, q_B) + J(\psi_B, q_T)\right) \\ &\quad + \nabla_{q_T} H \cdot B_t - \nabla_{q_T} H \cdot \Gamma q_T + \frac{1}{2} \frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma + S_t)(\Sigma + S_t)^T \end{aligned} \tag{5}$$

$$\begin{aligned} \sigma_H &= -\psi_T \cdot (\Sigma + S_t) \\ &= \nabla_{q_T} H \cdot (\Sigma + S_t) \,. \end{aligned}$$

Since we want to conserve the total energy, $dH$ has to be zero. Therefore we impose both $\mu_H$ and $\sigma_H$ to be zero. Following the reasoning outlined in [16], the auxiliary stochastic process $dY_t$ should not perturb the dynamics on the tangent space and should be constructed only to counterbalance those components of the OU process which are orthogonal to the manifold of constant energy, thus we define a projection operator $\mathbb{P} \in \mathbb{R}^{2\times 2}$. Since the Wiener process affects only the evolution equation of the baroclinic PV, it will be sufficient to project with respect to the manifold of constant baroclinic energy:

$$\begin{aligned} \mathbb{P} &= \mathrm{I} - \frac{1}{|\nabla_{q_T} H|^2} \nabla_{q_T} H (\nabla_{q_T} H)^T \\ &= \mathrm{I} - \frac{1}{|\psi_T|^2} \psi_T \psi_T^T \,, \end{aligned}$$

where $\mathrm{I} \in \mathbb{R}^{2\times 2}$ stands for the identity matrix. Since $\mathbb{P}(\nabla_{q_T} H) = 0$, $\mathbb{P}$ projects onto the tangent space of the baroclinic kinetic energy surface. Consequently, we want $S_t$ and $B_t$ to satisfy

$$\mathbb{P} S_t = 0 \,, \quad \mathbb{P} B_t = 0 \,. \tag{6}$$

From the assignment $\sigma_H = 0$ we can easily determine $S_t$. In fact, since $\nabla_{q_T} H$ is in the kernel of $\mathbb{P}$, imposing $\sigma_H = 0$ is equivalent to requiring $\Sigma + S_t = \mathbb{P}(\Sigma + S_t)$. Thus, using Eq. (6), we obtain

$$S_t = -(\mathrm{I} - \mathbb{P})\, \Sigma \,. \tag{7}$$

Substituting Eq. (7) into Eq. (5) and considering only the terms arising from the inclusion of the stochastic processes into the deterministic set of equations (since the deterministic model conserves energy, the other terms do not contribute to the variation of total energy), we can determine $B_t$ from the requirement $\mu_H = 0$:

$$B_t = (\mathrm{I} - \mathbb{P})\,\Gamma q_T + \frac{1}{2\,|\psi_T|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T\,. \tag{8}$$

Placing Eq. (7) and Eq. (8) into Eq. (4), after some manipulations, we get the following set of equations

$$dq_B = \left(-J(\psi_B, q_B) - J(\psi_T, q_T)\right)dt\,, \tag{9a}$$

$$dq_T = \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right)dt + \mathbb{P}\Sigma dW_t - \left(\mathbb{P}\Gamma q_T - \frac{1}{2\,|\psi_T|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T\right)dt\,. \tag{9b}$$

Equations (2)-(9) constitute our stochastic energy conserving 2-layer QG system. A detailed derivation of equations (8) and (9) is reported in Appendix A. As can be seen, the resulting set of equations contains multiplicative noise and nonlinear damping, due to the specific definition of the projection operator. The multiplicative noise is in fact a correlated additive multiplicative (CAM) noise [38, 53, 17].

## 3 Numerical implementation

Since we aim to analyze possible applications of this approach to climate and ocean models, which are typically formulated in terms of finite volumes or finite differences, we discretize equations (2)-(9) in terms of finite differences in the framework of a grid-point based code. Our discretization of the QG model is based on the energy and enstrophy conserving discretization scheme by Arakawa [1]. This scheme ensures that energy and enstrophy are conserved for all truncations. Especially this scheme does not require any numerical diffusion or dissipation for numerical stability. This property affects the energy and enstrophy cascades by automatically redistributing the energy and enstrophy at the truncation scales, making the model simulations unrealistic. However, using this discretization scheme will allow us to focus on the accuracy of the energy conservation of our stochastic approach.

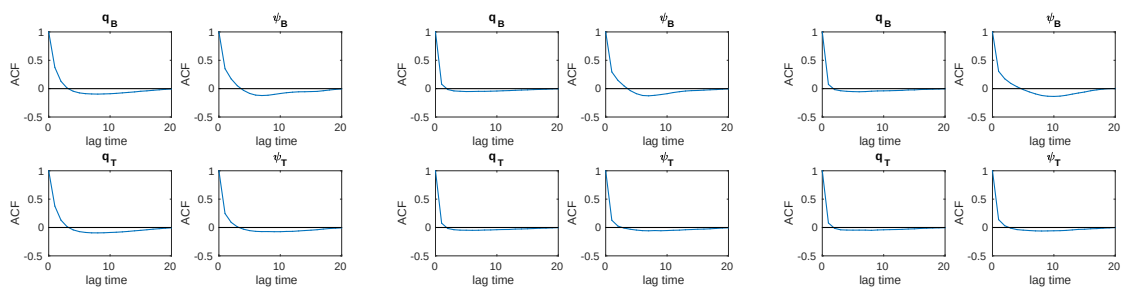
For the time stepping we employ explicit Runge-Kutta (RK) methods (whose order will be a matter of discussion in the following section), and we use a Fast Fourier Transform (FFT) to invert the Laplacian operator and obtain the streamfunctions from the corresponding PV. Since FFT is an exact numerical method and the Arakawa scheme is designed to conserve energy and enstrophy for any truncation, the only spurious effect on the energy due to the numerics is given by the RK method, which is known to be to a small extent dissipative in time. When dealing with the stochastic terms, we integrate them using either the Euler-Maruyama or Milstein schemes [48]. Finally, the initial distributions of the barotropic and baroclinic streamfunctions are generated using a pseudo-random number generator, i.e. no a-priori structure is given as input, and we define the corresponding PV by equations (2). We do not change the seed when defining the initial condition, thus all simulations at resolution $128 \times 128$ start from the same initial condition. Once the initial condition is defined, we set a new seed for the noise generator. Even though we do not have any forcing the model does not settle into a barotropic regime; the baroclinic modes are still active with active barotropic and baroclinic mode interactions (not shown). Furthermore, the probability density functions of the barotropic-baroclinic energy transfer terms are symmetric (not shown), suggesting an active interaction between barotropic and baroclinic modes.

## 3.1 Deterministic model

Before moving to the stochastic set of equations and related results, we test different orders of accuracy of our numerical scheme in the implementation of the deterministic 2-layer QG model (equations (1)-(2)) in order to find the optimal balance between accuracy and computational time.

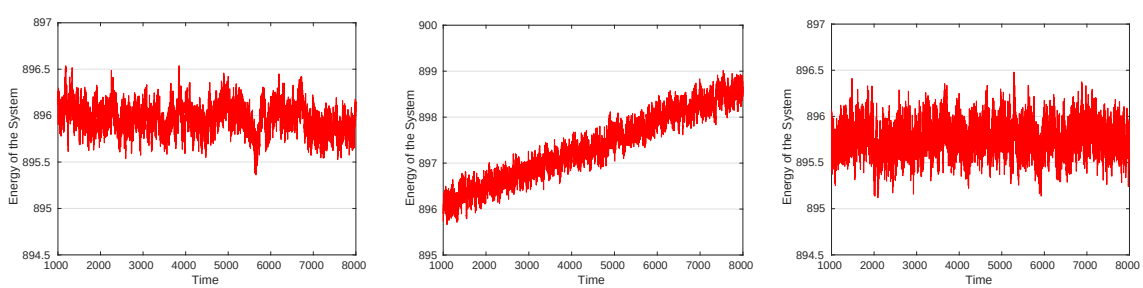
In our code, we solve the above evolution equations (1) for the PVs and then we compute the corresponding streamfunctions through equations (2). We use explicit Runge-Kutta $2^{nd}$ and $4^{th}$ order methods for the time integration, Arakawa $2^{nd}$ and $4^{th}$ order discretizations of the Jacobian [1] and a Fast Fourier Transform to invert the Laplacian operator. While performing our tests, we also consider different values of the mean depth of the fluid $h$; in particular we consider the cases $h = 1, 10, 100$. These tests are performed over a $128 \times 128$ spatial grid with a time step of $\Delta t = 10^{-3}$.

We do not report here all the statistics and energy graphs obtained with the different combinations of solvers, but show only those motivating our choice to employ RK4 and Arakawa $4^{th}$ order in the following.




**(a)** $h = 100$; RK2 & Arakawa $2^{nd}$ order. **(b)** $h = 100$; RK2 & Arakawa $4^{th}$ order. **(c)** $h = 100$; RK4 & Arakawa $4^{th}$ order.

**Figure 1:** ACF for the case with $h = 100$ and different combinations of deterministic numerical solvers. By using the second order method both for RK and Arakawa schemes, the baroclinic mode and the barotropic PV decorrelate more slowly.




**(a)** $h = 1$; RK2 & Arakawa $2^{nd}$ order. **(b)** $h = 1$; RK2 & Arakawa $4^{th}$ order. **(c)** $h = 1$; RK4 & Arakawa $4^{th}$ order.

**Figure 2:** Total energy graphs with $h = 1$ and different combinations of deterministic numerical solvers. It can be seen that also after the initial spin up period, which has been here neglected, energy increases in time when solving with RK2 and $4^{th}$ order Arakawa.

Figure 1 shows that the lower order combination of RK2 with Arakawa $2^{nd}$ order does not reproduce accurately the autocorrelation function (ACF) in the case $h = 100$ and that just increasing the order of Arakawa's discretization is enough to capture correctly the ACF. However, when combined with RK2 it does not conserve energy in the case $h = 1$ also after the initial spin up period (Fig. 2, where the spin up period has been neglected). More generally, we found that RK4 with Arakawa $4^{th}$ order is more reliable and that the scenario with $h = 10$ has a less discernible spin up period and it is the best reproduced case study with all the con-

sidered solvers, therefore we decided to employ this higher order numerical implementation and we fixed $h = 10$. As a consequence of this choice, the Rossby deformation radius $1/k_d$ is approximately 0.042 and we are in a ocean-like regime with small eddies. Furthermore, while the energy fluctuates around a mean value the fluctuations are relatively small; the energy fluctuation amplitude is less than 1% of the total energy. With the chosen numerical solvers also enstrophy is conserved by the system to a similar accuracy, as in the continuous scenario (not shown).

## 3.2 Stochastic equations

As in the deterministic case, we first solve the stochastic evolution equations (9) for the PVs and then we get the corresponding streamfunctions through equations (2). As a consequence of the analyses in the previous paragraph, we use an explicit RK $4^{th}$ order method for the time integration, Arakawa $4^{th}$ order discretization of the Jacobian [1] and FFT to invert the Laplacian operator. The stochastic terms are integrated using either the Euler-Maruyama or Milstein schemes [48]. Later we will analyze differences in the outcomes due to the stochastic solver. In the stochastic simulations we employ a $128 \times 128$ spatial grid with a time step of size $\Delta t = 10^{-3}$. As a consequence $\Delta x \approx 0.049$ and the model is in the eddy permitting regime.

As we will demonstrate below, for the dynamical consistency between deterministic and stochastic models it is crucial to consider spatially correlated noise. To demonstrate this, we consider two scenarios: in the first we assume that the noise on each grid point behaves as independent and identically distributed (iid) random variables, while in the second we allow for correlation between different grid points. In the following subsection a more detailed description of how the correlation matrix of the noise is constructed can be found. Finally, in order to generate the noise, we first produce uniformly distributed random numbers using the Mersenne-Twister algorithm [46], and then we reshape them through the Box-Muller procedure in such a fashion that they are sampled from a Gaussian distribution with the desired mean and variance, which in our case is $\mathcal{N}\left(\mu = 0, \sigma^2 = \Delta t\right)$. We compare the outcomes with a reference solution given by a deterministic simulation with $512 \times 512$ grid points and $\Delta t = 10^{-4}$. For a fair comparison, we project the fine grid data onto a grid with the same resolution as for the stochastic simulations.

## 3.3 Spatial noise structure

For allowing spatial correlations among different grid points, we need to determine the elements of the matrix $\Sigma$. We do this using eigenvectors obtained from a dimension reduction technique. Here we employ Empirical Orthogonal Functions (EOFs) [58]. We derive the EOFs from the high-dimensional deterministic control simulation.

Once the eigenvectors and the corresponding eigenvalues are computed, we select a number of EOFs and define $\Sigma$ as a convex combination of the chosen eigenvectors $E_i$. Such technique has been applied already in [25, 24]. A more general linear combination could be used and it would be easy to modify this constraint in order to attribute a stronger (or weaker) amplitude to the noise. The weights $\omega_i$ are selected as uniformly distributed random numbers, i.e.,

$$\Sigma = \sum_i \omega_i E_i \,, \quad \sum_i \omega_i = 1 \,, \quad \omega_i \sim \mathcal{U}\{0, 1\} \,,$$

where the eigenvectors are matrices with dimensions equal to the grid size. Coefficients related to the redundant eigenvectors are set to be zero. In view of the fact that the noise is only in the equation of the baroclinic mode, we use $\psi_T$-EOFs computed with respect to the $\mathrm{L}^2$ norm using the data of the high resolution run projected onto the coarser grid. Considering that the weights $\omega_i$ are chosen randomly in each simulation, no particular direction is preferred with respect to the others reducing possible biases in the results due to how the EOFs have been computed.

# 4 Results of the stochastic simulations

## 4.1 Space-time independent noise

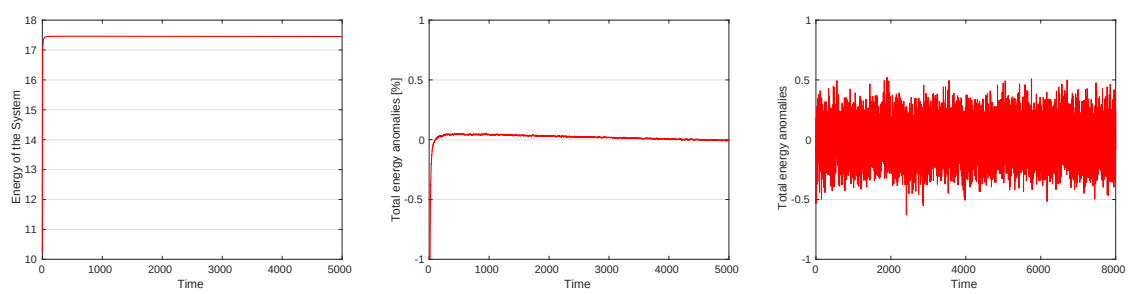
For reasons that will become evident later, in this scenario we perform the stochastic integration only with the Euler-Maruyama scheme. As assessment criteria, we first look at the conservation of energy and then at other statistical properties like the autocorrelation function (ACF) and probability density function (PDF). Here we consider iid white noise with zero mean and variance equal to the time step. Therefore, the matrices $\Sigma$ and $\Gamma$ can be written as

$$\Sigma = \sigma \mathrm{I}\,, \quad \Gamma = \gamma \mathrm{I}$$

where $\mathrm{I} \in \mathbb{R}^{2\times 2}$ stands for the identity matrix and $\sigma, \gamma \in \mathbb{R}$. We show the results for the case $\sigma = 1$ and $\gamma = 1$.

### 4.1.1 Numerical results

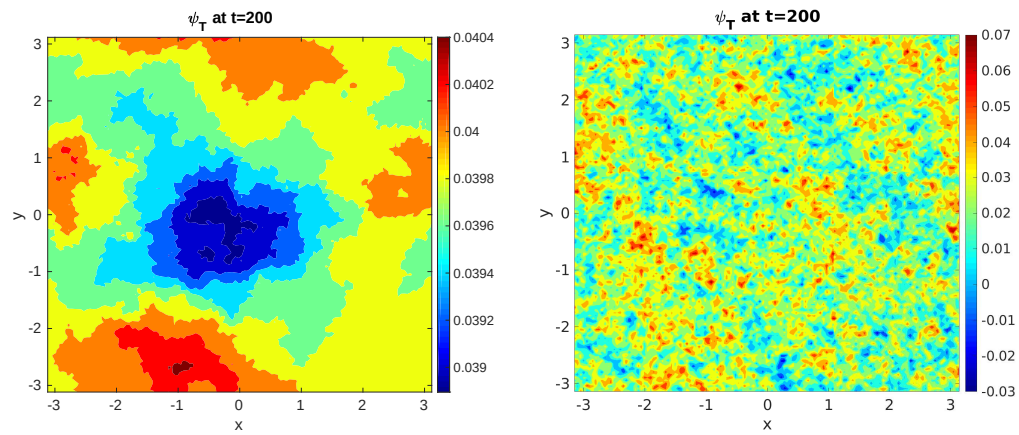
In this case, after a positive jump at the beginning of the simulation (because of the forcing, the system moves to the closest stable state, which has a different amount of total energy), energy is conserved (see Figure 3a) with very small fluctuations in time (see Figure 3b). In fact, when compared to a deterministic run at the same resolution (see Figure 3c), quite unexpectedly, variations in the total energy graph are smaller in the stochastic simulation. Still, looking at Figure 3b, it is possible to notice a small decreasing trend suggesting that the parameterization is damping in the long run which is likely due to the Runge-Kutta scheme, which is known to be damping for.




**(a)** Total energy of the system. **(b)** Total energy anomalies of the system. **(c)** Total energy anomalies of a low resolution deterministic run.

**Figure 3:** Total energy graph for the 2-layer stochastic QG model, with iid white noise and $\sigma = 1$ , $\gamma = 1$. After an initial positive jump, energy is conserved. Graph (c) shows the total energy anomalies of a deterministic run at the same resolution. It can be noticed that, surprisingly enough, oscillations are smaller in the case of the stochastic system.

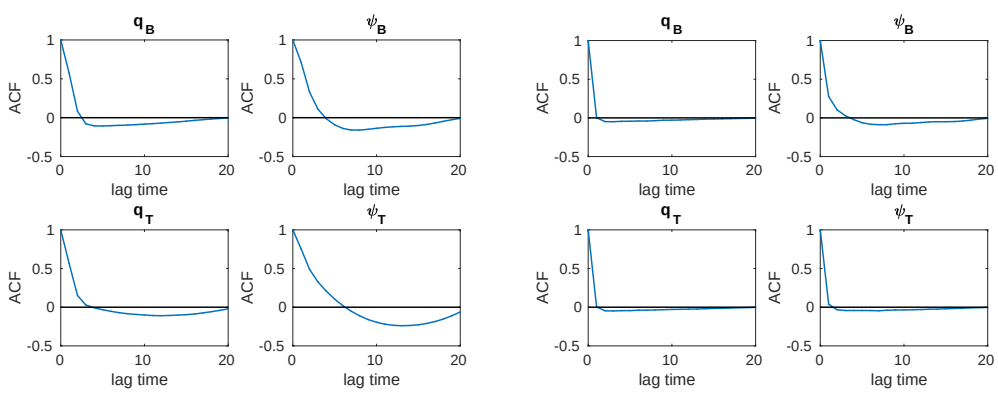
A less reassuring result is given by the contour plot of the baroclinic mode. In Figure 4 we show the baroclinic streamfunction (but a very similar result can be observed also for the baroclinic PV) at time $t = 200$ (left) and the plot of the same field at the same time given by our reference solution (right). What immediately stands out is the different pattern presented by the two figures. Furthermore, at a closer look it can also be noticed that the different colors in Figure 4a represent differences in the order of $10^{-4}$ and they become even smaller when looking at the contour plot for later times (not shown), meaning that the field is moving towards a constant state in space. This explains the smaller amplitude of the fluctuations in the energy graph with respect to the deterministic scenario, and it is reflected also in the ACF and in the PDF (see Figures 5-6). The former displays longer decorrelation times in general and, more specifically, the baroclinic streamfunction




**(a)** Baroclinic streamfunction. **(b)** Baroclinic streamfunction: reference solution.

**Figure 4:** Contour plot for the baroclinic streamfunction given by the stochastic system with iid white noise, $\sigma = 1$, $\gamma = 1$ (left) and by the high resolution deterministic simulation (right). The reader will immediately notice the different patterns displayed by the two pictures and, at a closer look, that the colors in the left graph represent differences in the order of $10^{-4}$, which get even smaller with the developing of the simulation, implying that the field is moving towards a constant state in space.

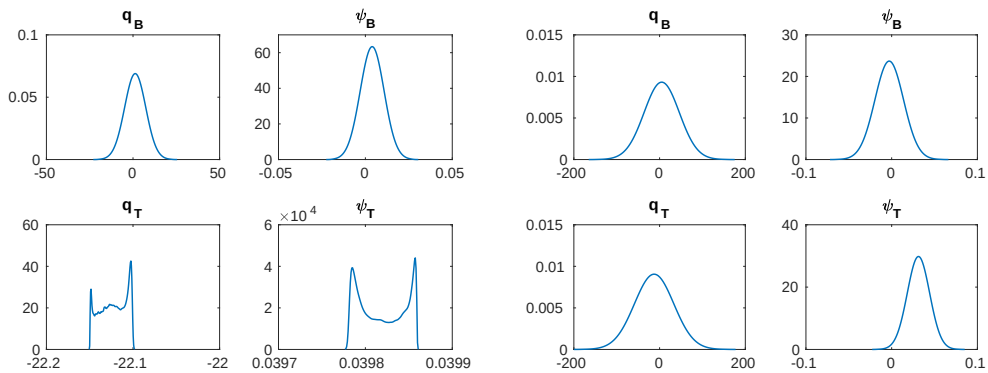
seems to require a longer decorrelation time with respect to the barotropic streamfunction which is in contrast with the physics. The latter instead shows no Gaussian distribution for the baroclinic mode and smaller variance for the barotropic mode.




**(a)** Auto-correlation functions. **(b)** Auto-correlation functions: reference solution.

**Figure 5:** ACF resulting from the stochastic system with iid white noise, $\sigma = 1$, $\gamma = 1$ (left) and from the high resolution deterministic simulation (right). The stochastic simulation does not well reproduce the ACF of the reference solution but displays longer decorrelation times.

In their paper [16] Frank and Gottwald used iid noise obtaining conservation of energy and physically reasonable outcomes. As already stated earlier, they employ a Lagrangian discretization of the system, while we consider the dynamics from a Eulerian point of view. Since Lagrangian descriptions of motion follow the trajectories of the single particles, and not the fluid as a whole in a fixed domain, in this frame the main purpose of the noise is simply to perturb the trajectory while remaining on the manifold of constant energy. Hence any spatial information added to the noise is not strictly necessary. On the other hand, Eulerian descriptions focus on what happens inside a well-defined domain and do not care about the behavior of the single particles. Thus in this framework spatial iid noise means that each grid point does not feel the influence of its




**(a)** Probability density functions. **(b)** Probability density functions: reference solution.

**Figure 6:** PDF rising from the stochastic system with iid white noise, $\sigma = 1$ , $\gamma = 1$ (left) and from the high resolution deterministic simulation (right). The baroclinic mode of the stochastic set of equations does not have a Gaussian distribution. Moreover the barotropic mode displays less variance.

neighbors, hence the noise would represent phenomena which fully evolve and decorrelate inside the cell; the gap between the large resolved scales and such small phenomena is too big to be correctly resolved by the numerics. Therefore it appears crucial to define a spatial structure of the noise in order to characterize how the noise should behave inside the domain and interact with the deterministic dynamics. The next section discusses in more detail our results.

## 4.2 Space-time correlated noise

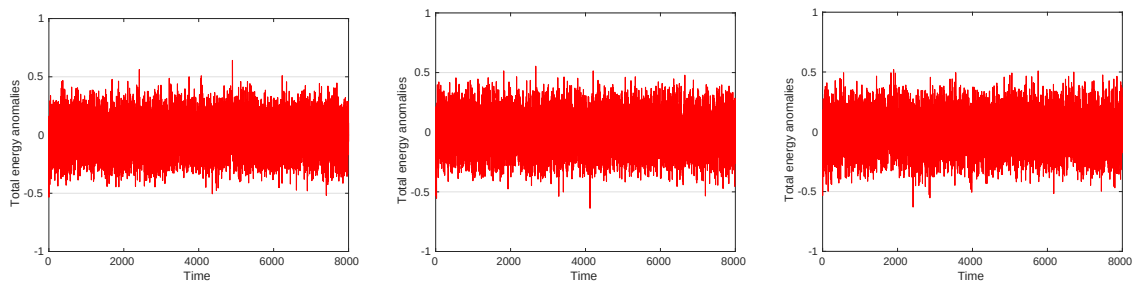
To ease computations, we neglected the $\Gamma$ term in equation 9b, i.e. $\Gamma = \mathbf{0}$. For both the Euler-Maruyama and Milstein schemes, we run an ensemble of 40 simulations using a convex combination of the first two EOFs to build the covariance matrix $\Sigma$. We tried also combinations with a different number of EOFs. When considering up to the first 10 EOFs, similar results to those we report here are obtained. With 20 or more EOFs we noted slightly worse performances of the scheme. Because of the constraint $\sum_i \omega_i = 1$, when considering relatively many EOFs, each of them has a small amplitude and then the patterns contrast with each other resulting in a not well-defined structure. On the other hand, a combination of a smaller number of EOFs can still maintain the individual patterns while allowing interaction with each other. In what follows, we opted for using only the first two main patterns.

As evaluation criteria, we employ the same analyses as before. Regarding the PDF, we also compute the first and second moments of the centers in order to investigate the ensemble variance. In addition we compare to the reference solution: the total variance and eddy length (computed through space correlations as presented in [3]).

### 4.2.1 Energy conservation

In each simulation the total energy fluctuates in time around a constant value. Differently from the previous case, there is no jump to a different stable state at the beginning of the time integration, meaning that our stochastic system keeps its evolution on the manifold defined by the initial condition. In both ensembles, if we compare the amplitude of the anomalies $A_{Anom}$ with respect to the mean value of the energy $\mu_{En}$, we see that $A_{Anom}$ is, for most of the running time, around 0.5 % of $\mu_{En}$ with spikes no greater than 0.7 %. We would like to point out that, even though for each individual simulation the evolution of the field variables is different (and this is shown by the fact that the PDFs of each individual run are centered in different locations),

the total energy of the system is almost the same at each time step, with differences in the order of $10^{-6}$, for each ensemble member. This implies that, starting from the same initial condition the model is exploring different possible configurations available with the defined amount of total energy. In Figure 7 we show the time-evolution of the energy anomalies for an individual Euler-Maruyama (Milstein) ensemble member together with the total energy anomalies graph of a deterministic simulation at the same resolution. It can be seen that fluctuations are roughly of the same amplitude for both the deterministic and the stochastic system. Thus, the energy fluctuations are mainly a result of the deterministic numerical scheme and not of the used stochastic scheme. This shows that the projection operator works very well in high-dimensional models and suggests that one should improve the discretization of the deterministic part in order to ameliorate energy conservation of the stochastic system.




**(a)** Euler-Maruyama. **(b)** Milstein. **(c)** Deterministic run.

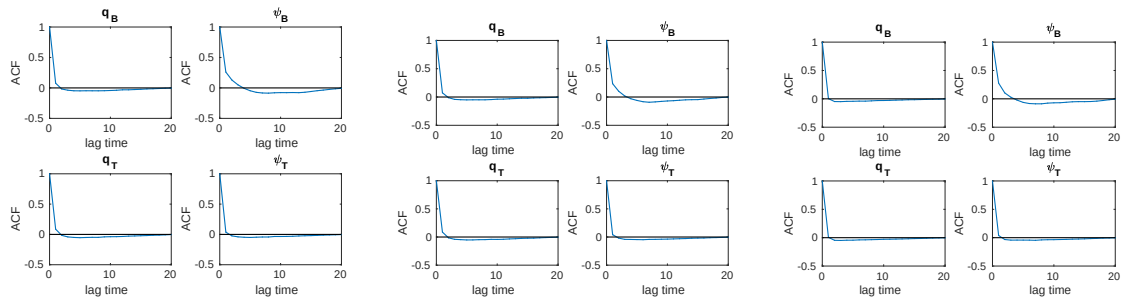
**Figure 7:** Total energy anomalies, displayed as a percentage of the mean total energy value, of an individual member of the Euler-Maruyama (Milstein) ensemble. On the right we also show the total energy anomalies graph for a deterministic run at the same resolution. The reader can notice that the fluctuations have roughly the same amplitude.

### 4.2.2 ACF

Every ensemble member shows roughly the same ACF pattern, independent of the stochastic solver. Differently from the previous case with iid white noise, in both ensembles we obtain decorrelation times very close to the reference. The barotropic streamfunction displays a longer decorrelation time with respect to the baroclinic streamfunction, suggesting that in future work a stochastic mode reduction might be performed for eliminating the baroclinic modes and having a stochastic barotropic model as in [43, 42, 44, 19, 18]. In Figure 8 we show the ACF for one stochastic simulation of the Euler-Maruyama (Milstein) ensemble together with the reference solution.

### 4.2.3 PDF

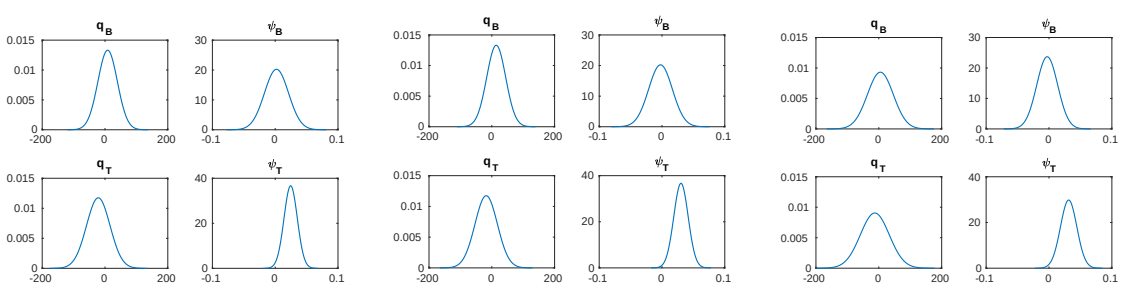
In contrast to the case with iid noise, in each run of both ensembles we recover the Gaussian behavior of the baroclinic mode displayed by the reference solution and more variance for the barotropic, see Figure 9 for the PDF graph of an individual Euler-Maruyama (Milstein) ensemble member and the reference solution. On the other hand, the PDF of an individual stochastic run shows less variance with respect to the reference solution except for the barotropic streamfunction which, in the comparison, shows more (this can be noticed in Figure 9). Hence we decided to investigate the variance of the ensemble by looking at the first and second moment of the center of the PDFs of the ensemble members. While we are not too much interested in the exact value taken by the first moment, due to the chaotic nature of the system, we would like to point out that, the Milstein ensemble displays more variance with respect to the Euler-Maruyama scheme. In Table 1 we report the 95%




**(a)** Euler-Maruyama. **(b)** Milstein. **(c)** Reference solution.

**Figure 8:** ACF of a particular Euler-Maruyama (Milstein) ensemble member and reference solution. Both ensembles well catch and reproduce the ACF shown by the high resolution deterministic simulation.

confidence interval (CI) of the centers of the PDFs; those related to the Milstein scheme span a wider range of values. We would like to remind the reader that the Euler-Maruyama and Milstein schemes have the same order of weak convergence (i.e. it is 1 for both schemes), but different order of strong convergence (0.5 for Euler-Maruyama and 1 for Milstein). Hence, since we are considering long time simulations, statistical properties of the field variables are more sensitive to weak convergence, while the evolution of trajectory paths is more sensitive to strong convergence. This explains why both ensembles catch the right shape of the PDF but at the same time the Milstein ensemble displays more variance. It could be argued that 0.5 might not be a meaningful difference; on the other side, high order stochastic integration methods include complicated correcting terms which might be hard to implement, see for instance [54] for an example of necessary conditions that have to be satisfied by a class of stochastic integration methods with (strong) order 1.5. Hence here we tried to analyze two of the most likely employed methods in complex climate models. We also checked the total variance of the stochastic ensemble and compared to a deterministic run at the same resolution and to the high resolution deterministic simulation. Unfortunately our stochastic parameterization is still not able to mend for the variance lost with the coarsening of the grid (see Table 1).




**(a)** Euler-Maruyama. **(b)** Milstein. **(c)** Reference solution.

**Figure 9:** PDF of an individual member of Euler-Maruyama (Milstein) ensemble and reference solution. The careful reader might notice that an individual ensemble run displays less variance with respect to reference for almost all field variables, inducing us to investigate the ensemble variance (see Table 1).

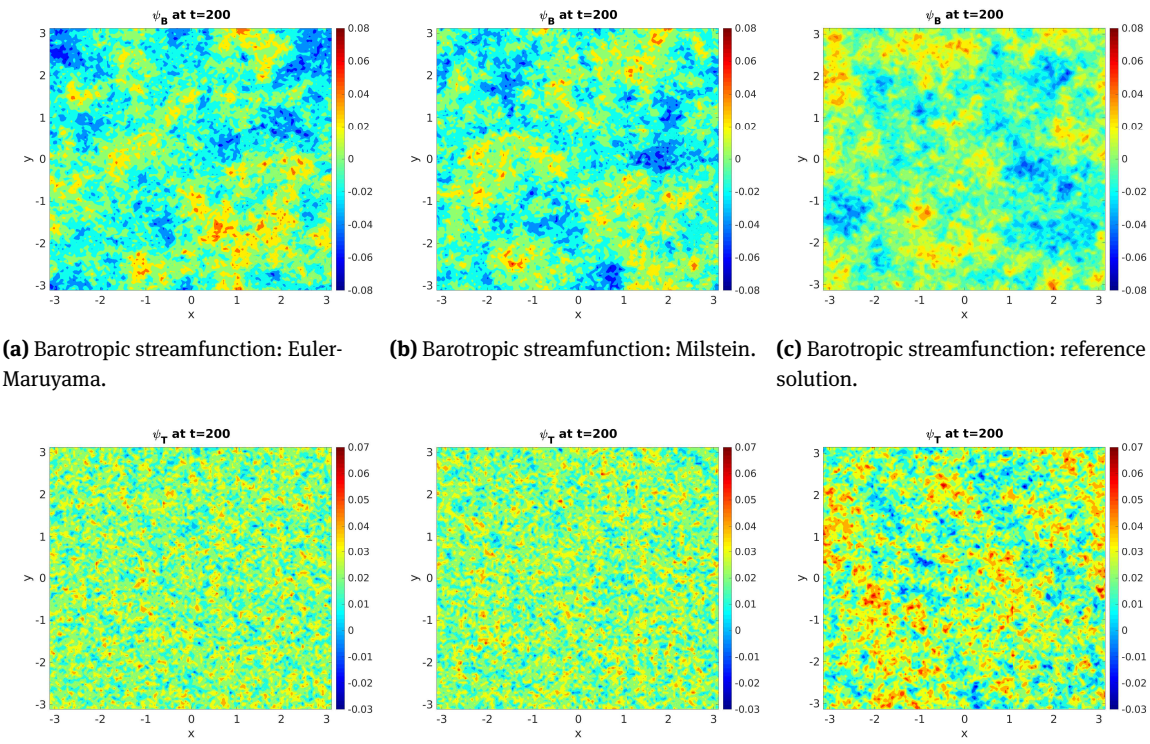
#### 4.2.4 Eddy length

When looking at the contour plots for the baroclinic (but also the barotropic) mode, we can see that in this scenario the stochastic simulations reproduce a similar pattern as for the case of the reference solution (see Figure 10 for a representation of the barotropic and baroclinic streamfunctions in the different setups). In order to have a more objective comparison criterion, we computed the eddy length for the two streamfunctions as a measure of the correlation among different grid points at a fixed time and looked at the e-folding scale; see for instance [3] for a more detailed description. In Table 1 we report the outcomes for the reference solution, a coarse deterministic simulation, Euler-Maruyama and Milstein ensemble. When considering the baroclinic mode, our stochastic parameterization does not improve the eddy length. In both ensembles it is about half of the baroclinic eddy length displayed by the high resolution simulation, remaining close to the outcome of a coarse deterministic run (see Table 1). A different conclusion is valid for the barotropic mode. In fact, in the case of the low resolution deterministic run we get an eddy length of circa $0.524$ both in zonal and meridional direction, while in the stochastic Euler-Maruyama (Milstein) ensemble it can vary between $\approx 0.511$ ($\approx 0.510$) and $\approx 0.570$ ($\approx 0.560$). Comparing these results with the reference solution, we can notice that we still did not manage to reproduce the high resolution eddy length ($\approx 0.714 - 0.716$) but we obtained an improvement of circa 9% with respect to the low resolution deterministic simulation (see Table 1). This result suggests that the perturbation induced by the noise is still not strong enough, but we are heading in the right direction. Furthermore, as has already been shown in [60] and references therein, the error dynamics considerably depends on the specific scale at which it is introduced, with a faster growth when located at small spatial scales. Here the noise structure is built with the first two EOFs, hence it can be regarded as a perturbation on the large spatial scales, which is in good agreement with the improved eddy length for the slower mode. Nevertheless, as stated in [58], if the first EOF can be associated with a definite physical process, this is more difficult already with the second (and even harder for higher-order) EOF because of the orthogonality constraint. On the other hand, real-world processes might not have orthogonal patterns. In fact, the patterns that most efficiently represent variance do not necessarily have anything to do with the underlying dynamical structure.

## 5 Conclusions and perspectives

We described the numerical implementation and evaluation of an energy conserving high-dimensional stochastic conceptual climate model. Our main focus here is the proof of concept whether the projection operator approach [16] can be applied to high-dimensional complex geophysical flow systems in a Eulerian setting ([16] developed this approach in a Lagrangian setting for a low-order model). Furthermore we investigate which assumptions regarding the noise should be considered in order to obtain not only energy conservation but also dynamically consistent results. For this purpose we used our QG model without forcing, dissipation and hyperdiffusion. Even though, the resulting circulations are less realistic when compared with the atmosphere and the oceans, this setup provides an ideal test bed for the numerical evalution of energy conservation of our numerical scheme.

In their paper, Frank and Gottwald [16] derived an energy conserving stochastic formulation for a 4-dimensional multi-scale toy model of the atmosphere. In order to preserve the conservation of energy they projected the noise with respect to the manifold of constant energy in such a fashion that those components of the noise, which would lead the trajectory to leave this manifold, are eliminated. In this paper we brought forward this approach and applied it to the high-dimensional 2-layer QG model through its Hamiltonian formulation. With the idea of analyzing the applicability of this procedure, not just to simple models but to a wider range of models with different degrees of complexity, we discretized the evolving equations in a Eulerian framework by means of finite differences, i.e. in a similar setup as most climate and ocean models. We could also introduce a time-scale separation parameter $\varepsilon$, depending on the different time scales of barotropic and baroclinic modes, to account for the time scale separation between the two modes. Even though here we




**(a)** Barotropic streamfunction: Euler-Maruyama.

**(b)** Barotropic streamfunction: Milstein.

**(c)** Barotropic streamfunction: reference solution.

**(d)** Baroclinic streamfunction: Euler-Maruyama.

**(e)** Baroclinic streamfunction: Milstein.

**(f)** Baroclinic streamfunction: reference solution.

**Figure 10:** Contour plot for the barotropic (top) and baroclinic (bottom) streamfunctions for an individual member of Euler-Maruyama (Milstein) ensemble and for the reference simulation. The patterns are similar, even though the stochastic simulations still do not have the same eddy length as the high resolution deterministic run (see Table 1).

focused on other issues and did not consider any time scale separation in our numerical simulations, stochastic mode reduction is a possible research direction to be followed.

In particular, we investigated the delicate step from a continuous to a discrete formulation and found that the numerics can be sensitive to the mean depth of the fluid, and hence chose solvers that reproduce correctly the properties of the system, e.g. conservation of energy, in the most general scenario. Once this aspect had been settled, we analyzed the effects on the system dynamics and statistics due to the introduction of the stochastic process. Mainly we compared the results for two different scenarios: in the first, the noise of each grid point behaves like an iid random variable while, in the second, we considered spatio-temporal correlations. We found that employing iid noise leads to either that energy is not conserved or to unphysical results and hence defining a spatio-temporal structure is important to respect the underlying dynamics of geophysical flows and for the conservation of energy and the preservation of important statistical properties, e.g. PDF. This is due to the Eulerian nature of our implementation. Frank and Gottwald employed a Lagrangian description which follows the trajectories of the single elements. Hence in their model the noise had the unique purpose of perturbing the trajectories while remaining on the manifold of constant energy. On the other hand, a Eulerian point of view looks at a well-defined domain and considers the fluid as a whole. Therefore in this frame the noise should perturb the dynamics while conserving energy and preserving the main properties of the fluid.

In the present work a convex combination of the first two EOFs, computed on the data of a high resolution deterministic run, have been used to define the spatio-temporal correlations consistent with the behavior of the deterministic system; other dimension reduction techniques, such as [4, 10, 14], could be used too. We did not recover the same amount of variance as in the high resolution simulation, but the eddy length in the barotropic mode is improved. This suggests that the stochastic perturbations are not strong or spatially

**Table 1:** Summary of some of the previously discussed analyses in the cases of, from left to right: the reference solution, a deterministic run at the coarse resolution, Euler-Maruyama and Milstein ensembles. In particular we report: location of the center of the PDF and, in case of the ensembles, its 95% CI, total variance and eddy length (the latter computed only for barotropic and baroclinic streamfunctions).

| Field variable | 512 × 512 det. | 128 × 128 det. | Euler-Maruyama | Milstein |
|---|---|---|---|---|
| **Mean** | | | | |
| $\psi_B$ | $-3.5687 \cdot 10^{-3}$ | $-2.4223 \cdot 10^{-3}$ | $-6.3485 \cdot 10^{-4}$ | $-4.4880 \cdot 10^{-4}$ |
| $\psi_T$ | $2.6234 \cdot 10^{-2}$ | $1.987 \cdot 10^{-2}$ | $2.5108 \cdot 10^{-2}$ | $2.4466 \cdot 10^{-2}$ |
| $q_B$ | 1.0412 | −5.1523 | 1.1737 | −0.2762 |
| $q_T$ | −14.6414 | −32.4428 | −14.9759 | −12.0254 |
| **95% CI** | | | | |
| $\psi_B$ | | | [−0.0078, 0.0066] | [−0.0070, 0.0061] |
| $\psi_T$ | | | [0.0212, 0.0290] | [0.0199, 0.0291] |
| $q_B$ | | | [−9.5180, 11.8654] | [−13.8563, 13.3039] |
| $q_T$ | | | [−28.8209, −1.1308] | [−28.0849, 4.0342] |
| **Total variance** | | | | |
| $\psi_B$ | $2.6302 \cdot 10^{-4}$ | $3.5852 \cdot 10^{-4}$ | $3.6169 \cdot 10^{-4}$ | $3.6100 \cdot 10^{-4}$ |
| $\psi_T$ | $1.6564 \cdot 10^{-4}$ | $1.0931 \cdot 10^{-4}$ | $1.0931 \cdot 10^{-4}$ | $1.0931 \cdot 10^{-4}$ |
| $q_B$ | 1698.6409 | 830.2853 | 829.9909 | 829.9071 |
| $q_T$ | 1796.2310 | 1066.7904 | 1067.1422 | 1067.1993 |
| **Eddy length** | | | | |
| $\psi_B$ zonal | $7.1428 \cdot 10^{-1}$ | $5.2422 \cdot 10^{-1}$ | [0.51124, 0.56958] | [0.50976, 0.56011] |
| $\psi_B$ merid. | $7.1632 \cdot 10^{-1}$ | $5.2406 \cdot 10^{-1}$ | [0.51151, 0.56854] | [0.50994, 0.56006] |
| $\psi_T$ zonal | $1.4918 \cdot 10^{-1}$ | $7.4597 \cdot 10^{-2}$ | [0.07452, 0.07459] | [0.07451, 0.07459] |
| $\psi_T$ merid. | $1.4918 \cdot 10^{-1}$ | $7.4577 \cdot 10^{-2}$ | [0.07451, 0.07460] | [0.07451, 0.07460] |

coherent enough. Another possible explanation is the lack of temporal memory in our scheme [27]. Memory terms have been included in many techniques, such as multi-level regression models [36, 35, 40, 30], showing encouraging results. However they might be rather complicated to implement and might lead to unstable and diverging simulations as reported in [13] in the case of the Wouters and Lucarini parameterization [61]. Investigating the impact of spatial coherence and memory in the noise will be part of our future research.

Two basic arguments are that the constraint $\sum_i \omega_i = 1$ for the noise amplitude was arbitrary and it could be changed in order to attribute a stronger (or weaker) amplitude to the noise; moreover we computed the EOFs with respect to the Euclidian norm and not to the total energy norm. A more philosophical discussion regards instead the usage of the EOF technique itself. In fact it is sensible that using the first two EOFs improves the dynamics of the large scales since the first EOFs can be easily associated to large scale dynamics. Going down the ladder, because of the orthogonality constraint, it becomes harder and harder to associate EOFs to well-defined physical phenomena and hence also to the smaller scales [58]. As has already been shown in [60] and references therein, the error dynamics is considerably dependent on the specific scale at which it is introduced, with a faster growth when located at small spatial scales. In spectral models this obstacle is easily resolved, since the wavenumber where the noise should be introduced (choosing therefore its spectral properties) can be selected directly. In a grid-point framework this is not the case. Further studies in this direction will be done in order to gain this ability also when using a grid-point discretization since most climate and ocean models are based on this type of numerics and will be reported elsewhere.

**Author contributions:** F.G. and C.F. developed the stochastic model formulation and wrote the manuscript. F.G. developed the model code and performed the simulations and data analysis.

**Supplementary material:** The model code and data are available upon request.

**Funding:** This paper is a contribution to the project M2 of the Collaborative Research Center TRR181 *Energy Transfers in Atmosphere and Ocean* funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 274762653.

**Conflict of interest:** The authors declare no conflicts of interest.

**Acknowledgement:** We thank Profs. M. Oliver and J. Rademacher for useful and constructive discussions throughout the preparation of this work. We thank also two anonymous reviewers who helped us to improve the quality of the paper.

# A Derivation of the stochastic 2-layer QG model

We report here a more detailed description of the procedure to derive equations (8) and (9) starting from the system (4), and making use of (7) and of the energy conservation property of the deterministic system (1)-(2).

As already mentioned, once (7) is obtained we substitute it into Eq. (5). Now we notice that the increment of total energy of the system (1)-(2) is

$$\begin{aligned} dH_{det} &= \frac{\partial H}{\partial q_B} \cdot dq_B + \frac{\partial H}{\partial q_T} \cdot dq_T \\ &= -\psi_B \cdot \big(-J(\psi_B, q_B) - J(\psi_T, q_T)\big) - \psi_T \cdot \big(-J(\psi_T, q_B) - J(\psi_B, q_T)\big) \\ &= 0 \,, \end{aligned}$$

due to energy conservation of the deterministic equations. Hence we need to consider only the remaining terms arising from the inclusion of the stochastic processes in the deterministic equations; we impose them to equal zero. Computations proceed as follows

$$\begin{aligned} -\psi_T \cdot B_t + \psi_T \cdot \Gamma q_T + \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma + S_t)(\Sigma + S_t)^T &= 0 \\ -\psi_T \cdot B_t + \psi_T \cdot \Gamma q_T + \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma - \Sigma + \mathbb{P}\Sigma)(\Sigma - \Sigma + \mathbb{P}\Sigma)^T &= 0 \\ -\psi_T \cdot B_t + \psi_T \cdot \Gamma q_T + \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P} &= 0 \\ -\psi_T^T B_t + \psi_T^T \Gamma q_T + \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P} &= 0 \,. \end{aligned}$$

At this stage we multiply from the left by $\frac{1}{|\psi_T|^2}\psi_T$ obtaining

$$-\frac{1}{|\psi_T|^2}\psi_T\psi_T^T B_t + \frac{1}{|\psi_T|^2}\psi_T\psi_T^T \Gamma q_T + \frac{1}{|\psi_T|^2}\psi_T\left(\frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right) = 0 \,.$$

Here we note that $\frac{1}{|\psi_T|^2}\psi_T\psi_T^T = \mathrm{I} - \mathbb{P}$ and that the term $\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}$ is a scalar. Hence we can rewrite the previous equation as

$$-(\mathrm{I} - \mathbb{P})B_t + (\mathrm{I} - \mathbb{P})\Gamma q_T + \frac{1}{|\psi_T|^2}\left(\frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T = 0 \,.$$

Now, making use of properties (6) and solving for $B_t$ we recover

$$B_t = (\mathrm{I} - \mathbb{P})\,\Gamma q_T + \frac{1}{2\,|\psi_T|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T \,,$$

which is Eq. (8). Once relations (7)-(8) have been derived, we can include Eq. (4c) in the evolution equation of the baroclinic PV (4b) and drop it, since drift and diffusion of the process $dY$ can now be written as a function of $\Sigma$ and $\Gamma$. Then, replacing $S$ and $B$ with the corresponding expressions as functions of $\Sigma$ and $\Gamma$, manipulations of Eq. (4b) proceed as follows:

$$\begin{aligned}
dq_T &= \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt - \Gamma q_T dt + \Sigma dW + S_t dW + B_t dt \\
&= \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt - \Gamma q_T dt + \Sigma dW - (\mathrm{I} - \mathbb{P})\, \Sigma dW \\
&\quad + \left( (\mathrm{I} - \mathbb{P})\, \Gamma q_T + \frac{1}{2|\psi_T|^2} \left( \frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P} \right) \psi_T \right) dt \\
&= \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt - \Gamma q_T dt + \Sigma dW - \Sigma dW + \mathbb{P}\Sigma dW \\
&\quad + \Gamma q_T dt - \mathbb{P}\Gamma q_T dt + \frac{1}{2|\psi_T|^2} \left( \frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P} \right) \psi_T dt \\
&= -\left(J(\psi_T, q_B) + J(\psi_B, q_T)\right) dt + \mathbb{P}\Sigma dW_t \\
&\quad - \left( \mathbb{P}\Gamma q_T - \frac{1}{2\,|\psi_T|^2} \left( \frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P} \right) \psi_T \right) dt\,,
\end{aligned}$$

while the equation of the barotropic PV (4a) remains unchanged. The above derived stochastic evolution equation of the baroclinic PV, together with Eq. (4a) and the corresponding streamfunctions described by Eq. (2), defines our stochastic 2-layer QG system.

# References


[1] A. Arakawa. "Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I." In: *Journal of Computational Physics* 1.1 (1966), pp. 119–143.

[2] G. Badin and F. Crisciani. *Variational formulation of fluid and geophysical fluid dynamics - Mechanics, symmetries and conservation laws*. Springer Berlin, 2018.

[3] E. A. Barnes and D. L. Hartmann. "The Global Distribution of Atmospheric Eddy Length Scales." In: *Journal of Climate* 25.9 (2012), pp. 3409–3416.

[4] M. Belkin and P. Niyogi. "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation." In: *Neural Computation* 15.6 (2003), pp. 1373–1396.

[5] J. Berner, G. Shutts, M. Leutbecher, and T. Palmer. "A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system." In: *J. Atmos. Sci.* 66 (2009), pp. 603–626.

[6] J. Berner, U. Achatz, L. Batte, L. Bengtsson, A. Camara, H. Christensen, M. Colangeli, D. Coleman, D. Crommelin, S. Dolaptchiev, C. Franzke, P. Friederichs, P. Imkeller, H. Jarvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. Palmer, C. Penland, M. Sakradzija, J. von Storch, A. Weisheimer, M. Weniger, P. Williams, and J. Yano. "Stochastic parameterization: Toward a new view of weather and climate models." In: *Bull. Amer. Meteorol. Soc.* 98.3 (2017), pp. 565–588.

[7] R. Buizza, M. Miller, and T. Palmer. "Stochastic representation of model uncertainties in the ECMWF ensemble prediction system." In: *Quart. J. R. Meteorol. Soc.* 125 (1999), pp. 2887–2908.

[8] G. F. Carnevale. "Statistical features of the evolution of two-dimensional turbulence." In: *Journal of Fluid Mechanics* 122 (1982), 143–153.

[9] G. F. Carnevale and J. S. Frederiksen. "Nonlinear stability and statistical mechanics of flow over topography." In: *Journal of Fluid Mechanics* 175 (1987), 157–181.

[10] R. R. Coifman and S. Lafon. "Diffusion maps." In: *Applied and Computational Harmonic Analysis* 21 (2006), pp. 5–30.

[11] C. J. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko. "Modelling uncertainty using circulation-preserving stochastic transport noise in a 2-layer quasi-geostrophic model." In: *arXiv:1802.05711* (2018).

[12] C. J. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko. "Numerically modelling stochastic Lie transport in fluid dynamics." In: *arXiv:1801.09729* (2018).

[13] J. Demaeyer and S. Vannitsem. "Comparison of stochastic parameterizations in the framework of a coupled ocean–atmosphere model." In: *Nonlinear Processes in Geophysics* 25.3 (2018), pp. 605–631.

[14] H. A. Dijkstra, A. Tantet, J. Viebahn, E. Mulder, M. Hebbink, D. Castellana, H. van den Pol, J. Frank, S. Baars, F. Wubs, M. Chekroun, and C. Kuehn. "A numerical framework to understand transitions in high-dimensional stochastic dynamical systems." In: *Dynamics and Statistics of the Climate System* 1.1 (2016).

[15] S. Dwivedi, C. L. E. Franzke, and F. Lunkeit. "Energetically Consistent Stochastic and Deterministic Kinetic Energy Backscatter Schemes for Atmospheric Models." In: *Q. J. Roy. Meteorol. Soc.* (2019), submitted.
[16] J. E. Frank and G. A. Gottwald. "Stochastic homogenization for an energy conserving multi-scale toy model of the atmosphere." In: *Physica D* 254 (2013), pp. 46–56.
[17] C. L. E. Franzke. "Extremes in dynamic-stochastic systems." In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.1 (2017), p. 012101.
[18] C. L. E. Franzke and A. J. Majda. "Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM." In: *J. Atmos. Sci.* 63 (2006), pp. 457–479.
[19] C. L. E. Franzke, A. J. Majda, and E. Vanden-Eijnden. "Low-order stochastic mode reduction for a realistic barotropic model climate." In: *J. Atmos. Sci.* 62 (2005), pp. 1722–1745.
[20] C. L. E. Franzke, T. J. O'Kane, J. Berner, P. D. Williams, and V. Lucarini. "Stochastic Climate Theory and Modelling." In: *WIREs Climate Change* 6 (2015), pp. 63–78.
[21] J. S. Frederiksen, T. J. O'Kane, and M. J. Zidikheri. "Stochastic subgrid parameterizations for atmospheric and oceanic flows." In: *Physica Scripta* 85.6 (2012), p. 068202.
[22] J. S. Frederiksen, V. Kitsios, T. J. O'Kane, and M. J. Zidikheri. "Stochastic Subgrid Modelling for Geophysical and Three-Dimensional Turbulence." In: *Nonlinear and Stochastic Climate Dynamics*. Ed. by C. L. E. Franzke and T. J.Editors O'Kane. Cambridge University Press, 2017, 241–275.
[23] C. W. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Vol. 4. Springer Berlin, 2009.
[24] D. Giannakis and A. J. Majda. "Comparing low-frequency and intermittent variability in comprehensive climate models through nonlinear Laplacian spectral analysis." In: *Geophysical Research Letters* 39.10 ().
[25] D. Giannakis and A. J. Majda. "Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability." In: *Proceedings of the National Academy of Sciences* 109.7 (2012), pp. 2222–2227.
[26] D. Givon, R. Kupferman, and A. Stuart. "Extracting macroscopic dynamics: model problems and algorithms." In: *Nonlinearity* 17.6 (2004), R55.
[27] G. A. Gottwald, D. Crommelin, and C. L. E. Franzke. "Stochastic Climate Theory." In: *Nonlinear and Stochastic Climate Dynamics*. Ed. by C. L. E. Franzke and T. O'Kane. Cambridge: Cambridge University Press, 2017.
[28] I. Grooms and L. Zanna. "A note on 'Toward a stochastic parameterization of ocean mesoscale eddies'." In: *Ocean Modelling* 113 (2017), pp. 30 –33.
[29] M. J. Grote, A. J. Majda, and C. Grotta Ragazzo. "Dynamic Mean Flow and Small-Scale Interaction through Topographic Stress." In: *Journal of Nonlinear Science* 9.1 (Feb. 1999), pp. 89–130.
[30] J. Harlim, A. Mahdi, and A. J. Majda. "An ensemble Kalman filter for statistical estimation of physics constrained nonlinear regression models." In: *Journal of Computational Physics* 257 (2014), pp. 782 –812.
[31] K. Hasselmann. "Stochastic climate models part I. Theory." In: *Tellus* 28.6 (1976), pp. 473–485.
[32] D. D. Holm. "Variational principles for stochastic fluid dynamics." In: *Proc. R. Soc.* A 471 (2015), p. 20140963.
[33] M. F. Jansen and I. M. Held. "Parameterizing subgrid-scale eddy effects using energetically consistent backscatter." In: *Ocean Modelling* 80 (2014), pp. 36–48.
[34] E. Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, 2003.
[35] D. Kondrashov, S. Kravtsov, and M. Ghil. "Empirical mode reduction in a model of extratropical low-frequency variability." In: *J. Atmos. Sci.* 63.7 (2006), pp. 1859–1877.
[36] S. Kravtsov, D. Kondrashov, and M. Ghil. "Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability." In: *J. Climate* 18 (2005), pp. 4404–4424.
[37] E. Mémin. "Fluid flow dynamics under location uncertainty." In: *Geophysical and Astrophysical Fluid Dynamics* 108 (2014), pp. 119–46.
[38] A. J. Majda, C. L. E. Franzke, and D. Crommelin. "Normal forms for reduced stochastic climate models." In: *Proc. Natl. Acad. Sci. USA* 106 (2009), pp. 3649–3653.
[39] A. J. Majda, C. L. E. Franzke, and B. Khouider. "An Applied Mathematics Perspective on Stochastic Modelling for Climate." In: *Phil. Trans. R. Soc. A* 366 (2008), pp. 2429–2455.
[40] A. J. Majda and J. Harlim. "Physics constrained nonlinear regression models for time series." In: *Nonlinearity* 26.1 (2013), p. 201.
[41] A. J. Majda and D. Qi. "Strategies for Reduced-Order Models for Predicting the Statistical Responses and Uncertainty Quantification in Complex Turbulent Dynamical Systems." In: *SIAM Review* 60.3 (2018), pp. 491–549.
[42] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden. "A mathematical framework for stochastic climate models." In: *Communications on Pure and Applied Mathematics* 54.8 (2001), pp. 891–974.
[43] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden. "Models for stochastic climate prediction." In: *Proceedings of the National Academy of Sciences* 96.26 (1999), pp. 14687–14691.
[44] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden. "Systematic strategies for stochastic mode reduction in climate." In: *J. Atmos. Sci.* 60.14 (2003), pp. 1705–1722.
[45] A. J. Majda and X. Wang. *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*. Cambridge University Press, 2006.

[46] M. Matsumoto and T. Nishimura. “Mersenne Twister:A 623-Dimensionally Equidistribuited Uniform Pseudo-Random Number Generator.” In: *ACM Transactions on Modelling and Computer Simulation* 8.1 (1998), pp. 3–30.

[47] T. J. O’Kane and J. S. Frederiksen. “Statistical dynamical subgrid-scale parameterizations for geophysical flows.” In: *Physica Scripta* 2008.T132 (2008), p. 014033.

[48] G. A. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

[49] P. Porta Mana and L. Zanna. “Toward a stochastic parameterization of ocean mesoscale eddies.” In: *Ocean Modelling 79* (2014), pp. 1 –20.

[50] D. Qi and A. J. Majda. “Low-Dimensional Reduced-Order Models for Statistical Response and Uncertainty Quantification: Two-Layer Baroclinic Turbulence.” In: *Journal of the Atmospheric Sciences* 73.12 (2016), pp. 4609–4639.

[51] 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.

[52] V. Resseguier, E. Mémin, and B. Chapron. “Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading.” In: *Geophysical & Astrophysical Fluid Dynamics* 111.3 (2017), pp. 177–208.

[53] P. D. Sardeshmukh and P. Sura. “Reconciling non-Gaussian climate statistics with linear dynamics.” In: *J. Climate* 22.5 (2009), pp. 1193–1207.

[54] M. J. Senosiain and A. Tocino. “Two-step strong order 1.5 schemes for stochastic differential equations.” In: *Numerical Algorithms* 75.4 (Aug. 2017), pp. 973–1003.

[55] T. G. Shepherd. “Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics.” In: *Advances in Geophysics* 32 (1990), pp. 287–338.

[56] G. Shutts. “A kinetic energy backscatter algorithm for use in ensemble prediction systems.” In: *Quart. J. Roy. Meteorol. Soc.* 131 (2005), pp. 3079–3102.

[57] D. J. Stensrud. *Parameterization schemes*. Cambridge University Press, 2007.

[58] H. von Storch and F. W. Zwiers. *Statistical analysis in climate research*. Cambridge University Press, 2003.

[59] G. K. Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2006.

[60] S. Vannitsem. “Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics.” In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 032101.

[61] J. Wouters and V. Lucarini. “Disentangling multi-level systems: averaging, correlations and memory.” In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.03 (2012), P03003.

[62] L. Zanna, P. Porta Mana, J. Anstey, T. David, and T. Bolton. “Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction.” In: *Ocean Modelling* 111 (2017), pp. 66 –80.