Spatial Covariance Modeling for Stochastic Subgrid-Scale Parameterizations Using Dynamic Mode Decomposition

Federica Gugole[1] and Christian Franzke[1]

[1]Meteorological Institute and Center for Earth System Research and Sustainability

November 24, 2022

Abstract

Stochastic parameterizations are broadly used in climate modeling to represent subgrid scale processes. While different parameterizations are being developed considering different aspects of the physical phenomena, less attention is given to the technical and numerical aspects. In particular, the use of Empirical Orthogonal Functions (EOFs) is well established whenever a spatial structure is required, without considering its possible drawbacks. By applying an energy consistent parameterization to the 2-layer Quasi-Geostrophic (QG) model, we investigate the model sensitivity to the \emph{a priori} assumptions made on the parameterization. In particular, we consider here two methods to prescribe the spatial covariance of the noise. First, by using climatological variability patterns provided by EOFs, and second, by using time-varying dynamics-adapted Koopman modes, approximated by Dynamic Mode Decomposition (DMD). The performance of the two methods are analyzed through numerical simulations of the stochastic system on a coarse spatial resolution, and the outcomes compared to a high-resolution simulation of the original deterministic system. The comparison reveals that the DMD based noise covariance scheme outperforms the EOF based. The use of EOFs leads to a significant increase of the ensemble spread, and to a meridional misplacement of the bi-modal eddy kinetic energy (EKE) distribution. On the other hand, using DMDs, the ensemble spread is confined and the meridional propagation of the zonal jet stream is accurately captured. Our results highlight the importance of the systematic design of stochastic parameterizations with dynamically adapted spatial correlations, rather than relying on statistical spatial patterns.

Spatial Covariance Modeling for Stochastic Subgrid-Scale Parameterizations Using Dynamic Mode Decomposition

F. Gugole[1] and C. L. E. Franzke[1]

[1]Meteorological Institute and Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany

Key Points:

- The structure of the spatial noise covariance matrix of stochastic parameterizations is important for flow dynamics and energy consistency
- Our results show that a noise covariance matrix based on Dynamic Mode Decomposition produces better results then a typically used Empirical Orthogonal Function based scheme
- Our Dynamic Mode Decomposition scheme is flow adaptive

Corresponding author: Federica Gugole, `federica.gugole@uni-hamburg.de`

Abstract

Stochastic parameterizations are broadly used in climate modeling to represent subgrid scale processes. While different parameterizations are being developed considering different aspects of the physical phenomena, less attention is given to the technical and numerical aspects. In particular, the use of Empirical Orthogonal Functions (EOFs) is well established whenever a spatial structure is required, without considering its possible drawbacks. By applying an energy consistent parameterization to the 2-layer Quasi-Geostrophic (QG) model, we investigate the model sensitivity to the *a priori* assumptions made on the parameterization. In particular, we consider here two methods to prescribe the spatial covariance of the noise. First, by using climatological variability patterns provided by EOFs, and second, by using time-varying dynamics-adapted Koopman modes, approximated by Dynamic Mode Decomposition (DMD). The performance of the two methods are analyzed through numerical simulations of the stochastic system on a coarse spatial resolution, and the outcomes compared to a high-resolution simulation of the original deterministic system. The comparison reveals that the DMD based noise covariance scheme outperforms the EOF based. The use of EOFs leads to a significant increase of the ensemble spread, and to a meridional misplacement of the bi-modal eddy kinetic energy (EKE) distribution. On the other hand, using DMDs, the ensemble spread is confined and the meridional propagation of the zonal jet stream is accurately captured. Our results highlight the importance of the systematic design of stochastic parameterizations with dynamically adapted spatial correlations, rather than relying on statistical spatial patterns.

1 Introduction

Geophysical flows involve a multitude of phenomena with vastly different spatial and temporal scales (e.g. Franzke, Oliver, Rademacher, & Badin, 2019; Vallis, 2006). Due to the underlying nonlinear equations of motion all these scales interact with each other. In order to obtain dynamically consistent and stable long time simulations, geophysical models need, in principle, to cover the whole range of scales. This poses great computational challenges: processes occurring on spatial scales smaller than the prescribed numerical grid scale and processes occurring on temporal scales faster than the prescribed numerical time step cannot be resolved. These unresolved subgrid-scale processes nevertheless may be energetically important, such as, for example, convection processes which are not resolved by current climate models, and may significantly affect the dynamics on the large resolved scales. To capture the effects of the subgrid-scale processes, parametrizations are typically introduced, whereby the unresolved scales are conditioned on the resolved scales (Stensrud, 2007).

Further complications, caused by the inevitable distinction between resolved and unresolved spatial scales, in numerical schemes occur for nonlinear fluid systems which exhibit energy and enstrophy cascades. For atmospheric dynamics it is well known that enstrophy is transferred from larger to smaller scales, until it is dissipated at the dissipation scale, whereas energy is transported from smaller to larger scales (Dubrulle, 2019; Vallis, 2006). For the majority of models, as for instance for general circulation models, the numerical resolution is not fine enough to resolve the dissipation processes. Subsequently, the enstrophy piles up at the truncation level, making the numerical model unstable and subject to numerical blow up. In order to guarantee numerical stability, artificial hyper-viscosity is introduced, leading to an increased viscosity of the fluid, which dissipates also the kinetic energy. Furthermore, the injection of energy from the unresolved subgrid-scales leads to an unphysical grid-size dependent representation of the kinetic energy.

In recent years there has been an extensive interest in the development of stochastic parameterizations for sub-grid scale processes (e.g. Berner et al., 2017; Franzke, O'Kane, Berner, Williams, & Lucarini, 2015; Gottwald, Crommelin, & Franzke, 2017; Imkeller & von Storch, 2001; Palmer & Williams, 2010). To mitigate possible damaging effects on the predictability by artificial energy dissipation, there has been a growing interest in designing energy-conserving and energy-consistent stochastic parametrizations (e.g. Dwivedi, Franzke, & Lunkeit, 2019; Frank & Gottwald, 2013; Gugole & Franzke, 2019; Jansen & Held, 2014; Jansen, Held, Adcroft, & Hallberg, 2015; Mémin, 2014; Resseguier, Mémin, & Chapron, 2017b). Broadly speaking, energy-consistent parametrizations fall into two different categories. The first approach is to derive expressions for additional terms to augment current deterministic fluid equations, such as done for kinetic backscatter (Dwivedi et al., 2019; Jansen & Held, 2014; Juricke, Danilov, Kutsenko, & Oliver, 2019; Zurita-Gotor, Held, & Jansen, 2015). The second strategy is to instead derive new stochastic expressions of the geophysical flow equations such that they still conserve, for instance, energy (Mémin, 2014; Resseguier, Mémin, & Chapron, 2017a; Resseguier et al., 2017b; Resseguier, Pan, & Fox-Kemper, 2019) or the Kelvin circulation theorem (Cotter, Crisan, Holm, Pan, & Shevchenko, 2018, 2019; Cotter, Gottwald, & Holm, 2017; Holm, 2015).

We consider here a forced and damped 2-layer Quasi-Geostrophic (QG) model, and as stochastic parameterization we employ the projection operator approach introduced in Frank and Gottwald (2013). The energy-consistent parametrization developed in Frank and Gottwald (2013) had been devised only for a low-dimensional Hamiltonian ordinary differential equation. Subsequently it was successfully adapted for an unforced inviscid QG model in Gugole and Franzke (2019). However, the spatial covariance of the stochastic parametrization is not specified by the methodology suggested in Frank and Gottwald (2013), and in Gugole and Franzke (2019) it was shown to be crucial for the system to have physically meaningful results. Our aim is to further investigate the sensitivity of the model dynamics with respect to the definition of the noise covariance. Such a noise covariance is usually determined *a priori* and is not representative of some specific scale-dynamics. Very often Empirical Orthogonal Functions (EOFs) (von Storch & Zwiers, 2003)) are employed for this purpose.

In addition to using EOFs, which capture the climatological dominant patterns of the variability, we will investigate spatial covariances based on Dynamic Mode Decomposition (DMD) (Kutz, Brunton, Brunton, & Proctor, 2016; Schmid, 2010; Schmid, Li, Juniper, & Pust, 2011). DMD is a computationally cost-effective algorithm attempting to compute a finite-dimensional approximation of the Koopman operator. The infinite-dimensional Koopman operator encodes the dynamics of a dynamical system and propagates observables in time (Lasota & Mackey, 1994). The intimate relationship between DMD modes and the eigenfunctions of the Koopman operator was established in Rowley, Mezić, Bagheri, Schlatter, and Henningson (2009). The patterns extracted by the DMD method, the so called DMD or Koopman modes, describe the dominant dynamical structures, and their corresponding eigenvalues characterize their temporal oscillation periods and their growth rates. In contrast to EOFs, DMD decomposes the dynamics according to its local in time oscillatory behavior. Connections between DMD and other model reduction techniques such as EOF or linear inverse modeling are discussed in Penland (1989); Penland and Magorian (1993); Schmid et al. (2011); Tu, Rowley, Luchtenburg, Brunton, and Kutz (2014). By projecting the full system onto the subspace spanned by the leading DMD modes, the governing equations may be approximated by a low-dimensional dynamical system allowing to study flow stability and bifurcations among other characteristics (Bagheri, 2013; Jovanovic, Schmid, & Nichols, 2014; Noack, Stankiewicz, Morzyaski, & Schmid, 2016; Schmid, 2010; Schmid et al., 2011; Schmid, Meyer, & Pust, 2009). Here we shall use DMDs to construct the spatial structure of the noise covariance matrix. DMDs have the same numerical complexity as EOFs, and have the advantage of using information of the system on the fly, with no additional information. For more details on DMDs and its limitations in approximating Koopman modes the interested reader is referred to Tu et al. (2014); Williams, Kevrekidis, and Rowley (2015)).

In contrast to approaches attempting to determine subgrid-scale information from highly resolved simulations (e.g. Berloff, 2005; Franzke, Majda, & Vanden-Eijnden, 2005; Hermanson, Hoskins, & Palmer, 2009; Porta Mana & Zanna, 2014), our approach using DMDs has the potential to seamlessly adapt to any grid resolution and is, hence, scale-adaptive. Our results show that the use of a dynamically adapted noise covariance keeps the ensemble spread confined and the meridional propagation of the zonal jet is better captured than with EOFs.

The remainder of this paper is structured as follows: In Section 2 we introduce the forced and damped 2-layer QG model. Section 3 describes the energy-consistent stochastic parameterization scheme. The spatial covariance of the noise is determined in Section 4 using EOF and DMD analysis. Section 5 presents results from numerical simulations exploring the effect of employing either climatological or dynamically adapted spatial covariances. We conclude with a discussion in Section 6.

2 The QG model

We consider the non-dimensional forced and damped 2-layer QG equations on a β-plane with double periodic boundary conditions (Vallis, 2006). This model represents synoptic-scale atmospheric dynamics around the mid-latitudes based on the quasi-geostrophic approximation, and simulates a jet-like zonal flow when suitable values for the parameters are chosen. A vertical structure of two discrete layers, which we assume to have equal depth, is the minimal vertical resolution that allows the representation of baroclinic processes (Holton, 2004).

Subgrid-scale eddies and bottom friction are modeled by biharmonic viscosity, while in the upper layer (i.e. $i = 1$) large-scale forcing is provided by a prescribed background-flow $U = 0.6$ as, for instance, in Cotter et al. (2018); Jansen and Held (2014). The external forcing leads to the formation of a jet stream with non-trivial meridional structure whose location experiences meridional shifts - a prominent feature of the observed atmospheric jet stream (Feldstein, 1998; James & Dodd, 1996; Riehl, Yeh, & La seur, 1950). Since we consider a non-dimensional description, the horizontal extensions have been rescaled to a $2\pi\times2\pi$ square. Finally the evolution equations for the potential vorticities (PVs)

$$q_i(\mathbf{x},t) = \nabla^2\psi_i + (-1)^i \frac{k_d^2}{2}(\psi_1 - \psi_2) + \beta y \quad i \in \{1,2\}$$

on the horizontal plane $\mathbf{x} = (x\ y)^T \in \mathbb{R}^2$, where x and y denote the zonal and the meridional directions respectively, read

$$\frac{\partial q_1}{\partial t} = -J(\psi_1 - Uy, q_1) - \nabla^2\left(\nu_1\nabla^4\psi_1\right) , \tag{1a}$$

$$\frac{\partial q_2}{\partial t} = -J(\psi_2, q_2) - \nabla^2\left(\nu_2\nabla^4\psi_2\right) - \tau_f^{-1}\nabla^2\psi_2 , \tag{1b}$$

where $\psi_i(\mathbf{x},t)$ $i \in \{1,2\}$ are the corresponding streamfunctions and $\tau_f = 10$ the frictional time-scale. The term $k_d^2/2 = (2f_0/Nh)^2$ quantifies the strength of the shear between the two layers and, hence, also the intensity of the baroclinic instability ($N = 1.2\cdot 10^{-2}$ being the Brunt-Väisälä frequency, $h = 200$ the mean depth of the layers and $f \approx f_0 + \beta y$ the approximate Coriolis term with $f_0 = 1$ and $\beta = 0.509$). These values imply a Rossby deformation radius $k_d^{-1} \approx 0.85$. The strength of the effective damping of the subgrid-scale eddies is quantified by $\nu_i = \nu(\psi_i)$. We follow Jansen and Held (2014); Leith (1996) and set

$$\nu_i(\mathbf{x}) = C_{Leith}\Delta^6\left|\nabla^4\psi_i\right| \quad i \in \{1,2\}$$

where $C_{Leith} = 0.005$ is an empirical constant and Δ is the size of the numerical grid-spacing. ∇ and ∇^2 denote, respectively, the horizontal gradient and the Laplacian operator, while the Jacobian operator J is defined as

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

In order to have a better defined distinction between slow and fast modes, we rewrite equations (1) as barotropic and baroclinic modes by assuming that barotropic modes evolve more slowly then baroclinic modes. Barotropic and baroclinic streamfunctions, ψ_B and ψ_T, can be defined as:

$$\psi_B = \frac{1}{2}(\psi_1 + \psi_2) \ , \quad \psi_T = \frac{1}{2}(\psi_1 - \psi_2) \ ;$$

which lead to the corresponding barotropic and baroclinic potential vorticities, q_B and q_T,

$$q_B = \nabla^2\psi_B + \beta y \ , \quad q_T = \nabla^2\psi_T - k_d^2\psi_T \ . \tag{2}$$

It can easily be shown that barotropic and baroclinic PVs can also be written as

$$q_B = \frac{1}{2}\left(q_1 + q_2\right) \ , \quad q_T = \frac{1}{2}\left(q_1 - q_2\right) \ ,$$

and we can use these relations to determine the evolution equations for q_B and q_T from (1). After some manipulations we obtain

$$\begin{aligned}\frac{dq_B}{dt} = & - J(\psi_B - \frac{1}{2}Uy, q_B) - J(\psi_T - \frac{1}{2}Uy, q_T) - \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) \\ & - \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) + \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) \ ,\end{aligned} \tag{3a}$$

$$\begin{aligned}\frac{dq_T}{dt} = & - J(\psi_T - \frac{1}{2}Uy, q_B) - J(\psi_B - \frac{1}{2}Uy, q_T) + \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) \\ & - \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) - \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) \ ,\end{aligned} \tag{3b}$$

where the derivative operator d is only with respect to time, and the biharmonic viscosity coefficient has been decomposed in its constant and non-constant parts. The unforced inviscid part of system (3) is Hamiltonian with the Hamiltonian H given by

$$H(q_B, q_T) = \frac{1}{2}\iint \left[(\nabla\psi_B)^2 + (\nabla\psi_T)^2 + k_d^2{\psi_T}^2\right] d\mathbf{x} \ , \tag{4}$$

corresponding to the total energy. The Hamiltonian allows for the following relationships, which we will use in the next Section,

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

For a general review of Hamiltonian mechanics and its application to geophysical fluid dynamics see for example Badin and Crisciani (2018); Salmon (1988); Shepherd (1990).

The numerical truncation affects deeply the dynamics by introducing a larger error at coarser resolutions. In particular, since smaller scales are not represented, the reinjection of kinetic energy from the unresolved into the resolved scales is reduced. This implies that the kinetic energy is dependent on the grid resolution (Dwivedi et al., 2019; Jansen & Held, 2014) leading, for instance, to a misrepresentation of the eddy kinetic energy at coarser resolutions (Juricke et al., 2019; Porta Mana & Zanna, 2014). Since the computational cost of high resolution simulations is often prohibitive, we aim at recovering the large scale variability induced by the faster modes, and hence increase the eddy kinetic energy at lower resolutions, by correcting the numerical error through the introduction of a stochastic parameterization for the sub-grid scales. In the next Section we present a stochastic parametrization which ensures that the stochastic noise does not break the inherent energy balance of the system.

3 Energy consistent stochastic parameterization

Our underlying model assumption is that there are many fast baroclinic modes which drive both the resolved and the large-scale barotropic modes, and which can be efficiently represented by a stochastic Ansatz. Since barotropic modes are mainly large-scale, its spectra are dominated by the large-scales, and the noise forcing can effectively affect just the baroclinic modes. Hence, as in Gugole and Franzke (2019), we represent the unresolved fast sub-grid processes by means of a stochastic forcing, which we assume to act directly on the baroclinic mode and only indirectly on the barotropic mode. In order to introduce only dynamically consistent perturbations, we employ the projection operator method proposed in Frank and Gottwald (2013) to construct a stochastic forcing such that the energy of the unforced inviscid core of the 2-layer QG model is preserved. This choice allows to retain the balance between the external forcing and the dissipation, while redistributing the energy among the scales. The approach by Frank and Gottwald (2013) also introduces seamlessly state-dependent noise and dissipation. This potentially also allows for a realistic representation of subgrid-scale effects as in previous studies (Berner, Shutts, Leutbecher, & Palmer, 2009; Dwivedi et al., 2019; Franzke et al., 2015; Jansen et al., 2015) since this approach also predicts the corresponding nonlinear damping. In previous approaches the damping needed to be tuned in order to ensure numerical stability (Whitaker & Sardeshmukh, 1998; Zhang & Held, 1999). Our approach avoids any empirical tuning of the damping.

Since Gaussian white noise exists only as a distribution, stochastic evolution equations should be interpreted as integral equations (Gardiner, 2009; Pavliotis & Stuart, 2008). Hence we slightly change notation towards this interpretation, where we dropped the integral symbol in order to have a not too heavy notation. In this work we adopt Itô's interpretation of the stochastic integrals (Gardiner, 2009). We propose the following stochastically forced modification of the 2-layer QG system (1)

$$
\begin{aligned}
dq_B =& -\left(J(\psi_B - \frac{1}{2}Uy, q_B) + J(\psi_T - \frac{1}{2}Uy, q_T)\right)dt - \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right)dt \\
&- \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) + \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right)dt\ , \qquad (5a) \\
dq_T =& -\left(J(\psi_T - \frac{1}{2}Uy, q_B) + J(\psi_B - \frac{1}{2}Uy, q_T)\right)dt + \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right)dt \\
&- \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) - \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right)dt \\
&+ \Sigma(\mathbf{x},t)\,dW_t + dY_t\ , \qquad (5b) \\
dY_t =& Bdt + SdW_t, \qquad (5c)
\end{aligned}
$$

where W_t denotes a Wiener process. The auxiliary stochastic process Y_t, which is parametrized by $B = B(\mathbf{x},t)$ and $S = S(\mathbf{x},t)$, is determined to ensure that the stochastic forcing $\Sigma(\mathbf{x},t)dW_t$ preserves the energy given by the Hamiltonian (4) (Frank & Gottwald, 2013). Using Itô's formula (Gardiner, 2009) the change in the energy is given by

$$
\begin{aligned}
dH &= \frac{\partial H}{\partial q_B}\cdot dq_B + \frac{\partial H}{\partial q_T}\cdot dq_T + \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : dq_T dq_T^T \\
&= \mu_H dt + \sigma_H dW_t\ ,
\end{aligned}
$$

where the matrix inner product is defined as $A : B = a_{ij}b_{ij} = Tr(AB^T)$, and where

$$\begin{aligned}
\mu_H =& + \psi_B \cdot \left(J(\psi_B - \frac{1}{2}Uy, q_B) + J(\psi_T - \frac{1}{2}Uy, q_T) + \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right)\right) \\
&+ \frac{C_{Leith}\Delta^6}{2}\psi_B \cdot \nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) + \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) dt \\
&+ \psi_T \cdot \left(J(\psi_T - \frac{1}{2}Uy, q_B) + J(\psi_B - \frac{1}{2}Uy, q_T) - \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) - B_t\right) \\
&+ \frac{C_{Leith}\Delta^6}{2}\psi_T \cdot \nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) + \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) dt \\
&+ \frac{1}{2}\frac{\partial^2 H}{\partial q_T \partial q_T} : (\Sigma + S_t)(\Sigma + S_t)^T \ , \\
\sigma_H =& - \psi_T \cdot (\Sigma + S_t) \\
=& \nabla_{q_T} H \cdot (\Sigma + S_t) \ .
\end{aligned}$$

Our aim is to control the stochastic forcing in order to preserve the energetic balance between the external forcing and the dissipation. In order to guarantee the total energy not to be affected by the stochastic forcing, we set σ_H and the sum of those terms in μ_H due to the stochastic processes to be zero. The auxiliary process must be constructed to force the deviations from the manifold of constant energy, caused by the stochastic forcing $\Sigma(\mathbf{x}, t)dW_t$, back onto the manifold. It should therefore only have components orthogonal to the manifold of constant energy. Thus we define a projection operator $\mathbb{P}$, which projects onto the tangent space of the energy manifold, and we require $\mathbb{P}S = \mathbb{P}B = 0$. Since the Wiener process affects only the evolution equation of the baroclinic mode, it is sufficient to project onto the manifold of constant baroclinic energy, and we define the projection operator $\mathbb{P}$ as

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

where I stands for the identity operator. Using $\mathbb{P}\left(\nabla_{q_T} H\right) = 0$, the condition $\sigma_H = 0$ provides an expression for S, while it is possible to determine B by considering only the terms of μ_H due to the introduction of the stochastic processes:

$$\begin{aligned}
S &= -\left(\mathrm{I} - \mathbb{P}\right)\Sigma \ , \\
B &= +\frac{1}{2\left|\psi_T\right|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T \ .
\end{aligned}$$

We can now finally express our stochastic forced and damped 2-layer QG model (5) as

$$\begin{aligned}
dq_B =& -\left(J(\psi_B - \frac{1}{2}Uy, q_B) + J(\psi_T - \frac{1}{2}Uy, q_T)\right) dt - \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) dt \\
&- \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) + \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) dt \ ,
\end{aligned} \tag{6a}$$

$$\begin{aligned}
dq_T =& -\left(J(\psi_T - \frac{1}{2}Uy, q_B) + J(\psi_B - \frac{1}{2}Uy, q_T)\right) dt + \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) dt \\
&- \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B+\psi_T)\right|\nabla^4(\psi_B+\psi_T) - \left|\nabla^4(\psi_B-\psi_T)\right|\nabla^4(\psi_B-\psi_T)\right) dt \\
&+ \mathbb{P}\Sigma dW_t + \frac{1}{2\left|\psi_T\right|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T dt \ .
\end{aligned} \tag{6b}$$

The stochastic forced and damped 2-layer QG model (6) 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 (Majda, Franzke,

& Crommelin, 2009; Sardeshmukh & Sura, 2009). The interested reader may find more details about the necessary steps for the derivation of (6) in Frank and Gottwald (2013); Gugole and Franzke (2019).

In equations (6) the noise strength $\Sigma(\mathbf{x}, t)$, which specifies the spatial covariance of the noise, is still unspecified. In Gugole and Franzke (2019), it was shown that the choice of a dynamically consistent spatial structure of the noise covariance is crucial for a stochastic parametrization to be reliable. We propose in the next Section ways to prescribe the spatial structure.

4 The spatial covariance structure of the noise

We prescribe the spatial covariance of the noise by expressing $\Sigma(\mathbf{x}, t)$ through p dynamically relevant patterns of the large-scale dynamics $\phi_i(\mathbf{x}, t)$, $i = 1, \ldots, p$. In particular, we write

$$\Sigma(\mathbf{x}, t) = \sum_{i=1}^{p} \gamma_i \phi_i(\mathbf{x}, t) \tag{7}$$

where the $\gamma_i \in \mathbb{R}$ are weights associated with each pattern.

We shall discuss here two choices of patterns ϕ_i: first, Empirical Orthogonal Functions (EOFs), which capture time-invariant climatological patterns, and, second, patterns obtained by means of Dynamical Mode Decomposition (DMD), which describe time-varying, dynamically adapted dominant patterns.

4.1 Empirical Orthogonal Functions

4.1.1 Theory

EOF is a multivariate statistical analysis technique that derives the dominant patterns of variability from a n-dimensional field, usually indexed by location in space (von Storch, 1995; von Storch & Zwiers, 2003). Let $\mathbf{X}$ be an n-dimensional random vector, whose mean is assumed to be zero; otherwise the anomalies of the field with respect to the mean should be considered. At its first stage the EOF analysis computes the vector ϕ_1 with $\|\phi_1\| = 1$ such that

$$\epsilon_1 = \mathbb{E}\left(\left\|\mathbf{X} - \langle \mathbf{X}, \phi_1 \rangle \phi_1\right\|^2\right) \tag{8}$$

is minimized, where we denoted with $\mathbb{E}$ the expectation operator, the vector norm by $\|\cdot\|$ and the inner product with $\langle \cdot, \cdot \rangle$. Equation (8) describes the projection of the field $\mathbf{X}$ onto a 1-dimensional subspace spanned by the vector ϕ_1. Minimizing ϵ_1 is equivalent to maximizing the variance of $\mathbf{X}$ contained in this subspace, in fact it can be shown that

$$\epsilon_1 = \mathrm{Var}(\mathbf{X}) - \mathrm{Var}(\langle \mathbf{X}, \phi_1 \rangle) \ ,$$

where the variance of $\mathbf{X}$ is defined to be the sum of the variances of its elements. Let Γ denote the covariance matrix of $\mathbf{X}$. It can be shown that ϕ_1 is an eigenvector of Γ with corresponding eigenvalue λ_1. Therefore, the minimum of equation (8) is achieved by the vector associated to the largest eigenvalue of Γ, i.e. vector ϕ_1.

The same procedure is repeated to find the second EOF, which is the vector ϕ_2 with $\|\phi_2\| = 1$ minimizing

$$\epsilon_2 = \mathbb{E}\left(\left\|(\mathbf{X} - \langle \mathbf{X}, \phi_1 \rangle \phi_1) - \langle \mathbf{X}, \phi_2 \rangle \phi_2\right\|^2\right) \ ,$$

and corresponding to the second largest eigenvalue λ_2 of Γ. Finally we remark that Γ is an Hermitian matrix, hence its eigenvectors are orthogonal to one another. Moreover in case of translationally invariant systems they correspond to Fourier modes.

4.1.2 Constructing Σ using EOF

EOFs are computed on a time series of the baroclinic streamfunction (after the dynamics settled on the attractor) of the deterministic system (3) over a spatial grid with 128×128 elements. To construct the spatial structure of the noise, we compute a linear combination of the first p EOF patterns ϕ_i^{EOF} for $i = 1, \dots, p$ with weights given by the square roots of their corresponding eigenvalues λ_i^{EOF} $i = 1, \dots, p$ writing (7) as

$$\Sigma(\mathbf{x}) = \sum_{i=p}^{l} \sqrt{\lambda_i^{EOF}} \phi_i^{EOF}(\mathbf{x}) \ . \tag{9}$$

Hence, Σ is constant in time and $\Lambda = \Sigma\Sigma^T$ corresponds to the variance of the QG model's baroclinic stream function as approximated by the first p EOFs.

EOF patterns As in the majority of cases, the spectrum of the EOF eigenvalues rapidly decay, and the first 5 EOFs carry circa 95% of the variance. Higher EOFs do not carry significant variance and hence might be considered as numerical noise (see Figure 1a). EOFs 1-2 (Figures 1b-1c) represent the predominant traveling Rossby wave supported by the 2-layer QG model. EOF 3 (Figure 1d) does not represent any wave but captures the spatial dominant pattern associated with the jet stream. EOFs 4-5 (Figures 1e-1f) capture again dominant wave patterns. In our numerical simulations, we use either only the first two EOFs, corresponding to $\Lambda \approx 0.36$, or the first five EOFs, i.e. $\Lambda \approx 0.47$.

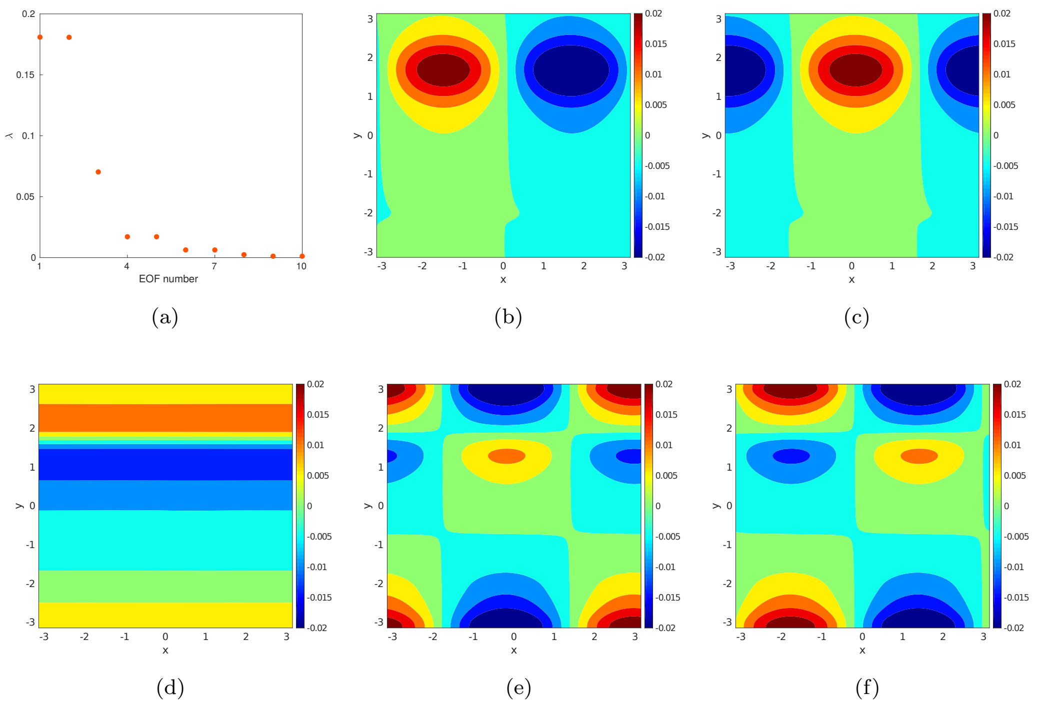

Figure 1. EOF singular values spectrum of the first 10 eigenvectors and first 5 EOF patterns. From left to right: top row; eigenvalues spectrum, EOF-1 and EOF-2; bottom row; EOF-3, EOF-4 and EOF-5.

EOFs are widely used in the climate science, thanks to their robust computability given a large available data set. Nonetheless EOFs have known limitations. In particular, their physical interpretation is restricted. While it is possible to associate the first EOF with observed physical features, this becomes increasingly complicated for higher-order EOFs, because of the orthogonality constraint (von Storch & Zwiers, 2003). We therefore introduce in the next Section DMDs, which capture relevant modes, adapted to the prevailing dynamics.

4.2 Dynamic Mode Decomposition

4.2.1 DMD and the Koopman operator

Here we briefly present the Koopman operator and its connection with Dynamic Mode Decomposition (DMD). Detailed reviews about the Koopman operator can be found, for instance, in (Budišić, Mohr, & Mezić, 2012; Mezić, 2013), while theory and applications of DMD are provided, among others, in (Kutz et al., 2016; Schmid, 2010; Tu et al., 2014).

Let $\dot{x} = f(x)$ denote a general continuous-time dynamical system with initial condition $x(0) = x_0 \in \mathbb{R}^n$. On the assumption that there exists a unique solution of this initial value problem, it is possible to introduce the flow map φ_t such that $x(t) = \varphi_t(x_0)$. Define an arbitrary observable $\psi(x)$. The value of this observable ψ, which the system sees starting in x_0 at time t, is

$$\psi(t, x_0) = \psi(\varphi_t(x_0)) \ .$$

The Koopman operator is a semigroup of operators $\mathcal{K}_t$, acting on the space of observables parameterized by time t

$$\mathcal{K}_t \psi(x_0) = \psi(\varphi_t(x_0)) \ .$$

It is important to underline that the operator $\mathcal{K}_t$ is linear also in case of non-linear dynamics f, thus it makes sense to consider its spectral properties, but the eigenfunctions of the Koopman operator are not necessarily linear. Dynamic mode decomposition is a data-driven technique for computing an approximation of the Koopman modes. Consider a dynamical system as above, and two sets of data, either of the state variables or of any observable of them,

$$\mathbf{X} = \begin{pmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \\ | & | & & | \end{pmatrix} \qquad \mathbf{X}' = \begin{pmatrix} | & | & & | \\ \mathbf{x}'_1 & \mathbf{x}'_2 & \cdots & \mathbf{x}'_m \\ | & | & & | \end{pmatrix}$$

such that

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}(t_k) \in \mathbb{R}^n \ , & \mathbf{x}'_k &= \mathbf{x}(t_k + \delta t) = \mathcal{K}_{\delta t} \mathbf{x}_k \ , \\ \mathbf{x}_k &= \mathbf{x}(t_{k-1} + \Delta t) = \mathcal{K}_{\Delta t} \mathbf{x}_{k-1} \ , & \mathbf{x}'_k &= \mathbf{x}(t_{k-1} + \delta t + \Delta t) = \mathcal{K}_{\Delta t} \mathbf{x}'_{k-1} \ , \end{aligned}$$

where $m\Delta t$ defines the time window, and $\delta t \leq \Delta t$ determines the accuracy of the reconstructed dynamics. It is important to mention that matrices $\mathbf{X}$ and $\mathbf{X}'$ are assumed to be *tall and skinny*, i.e. it is assumed that the size n of a snapshot is larger than the number $m-1$ of snapshots. In the DMD algorithm the Koopman operator is approximated by means of a least square fit operator $\mathbf{K}_{\delta t}$ relating data $\mathbf{X}' \approx \mathbf{K}_{\delta t}\mathbf{X}$. The numerically stable algorithm, based on a singular value decomposition and outlined for the first time in Schmid (2010) and improved in Tu et al. (2014), allows for a low-rank $r \leq m$ representation of the operator $\mathbf{K}_{\delta t}$ onto the first r EOF modes of matrix $\mathbf{X}$. Details about the algorithm, as well as a MATLAB© function, are provided in Kutz et al. (2016). The DMD modes ϕ_i are the (complex) eigenvectors of $\mathbf{K}_{\delta t}$, and they are not orthogonal. Furthermore, they represent dynamically relevant structures, the so called Koopman modes, whose temporal oscillation periods and their growth rates are provided by their associated (complex) eigenvalues λ_i. There exists a real eigenvalue $\lambda_0 = 1$ with eigenvector ϕ_0 corresponding to the mean of the observable $\mathbf{x}$. Whereas EOF decomposes the dynamics according to dominant stationary patterns, DMD decomposes the dynamics according to its local in time oscillatory behavior.

We remark that there exists an intimate relationship between the DMD matrix $\mathbf{K}_{\delta t}$ and the Koopman operator, first realized in Rowley et al. (2009). However, it is well established that DMD provides a good approximation of the actual Koopman operator - and hence constitutes a good representation of the underlying dynamics - only in case of sufficiently rich and diverse observations (Budišić et al., 2012; Tu et al., 2014; Williams et al., 2015). The least square approximation of the Koopman operator suggests that a good approximation is guaranteed for sufficiently small δt and for sufficiently small time intervals $m\Delta t$ such that the dynamics is essentially linear.

4.2.2 Defining the noise covariance by means of DMD

As for EOFs, we choose the baroclinic stream function ψ_T to determine the DMD modes. In deterministic systems the eigenvalues of the Koopman operator lie on the complex unit circle and, apart from the eigenvalue λ_0^{DMD} corresponding to the mean mode, appear as complex conjugate pairs. In stochastic systems, however, eigenvalues inside or outside the unit circle may appear; see Figure 2 for an instance of the DMD eigenvalues for the stochastic QG model (6). Since we want to capture the dynamically relevant patterns of the deterministic QG system, we exclude all eigenmodes ϕ_i^{DMD} whose eigenvalues do not lie on the unit circle (within some tolerance to account for numerical noise). The eigenvectors and eigenvalues are sorted with decreasing real part according to $\lambda_0^{DMD} = 1 > \mathrm{Re}(\lambda_1^{DMD}) \geq \cdots \geq \mathrm{Re}(\lambda_r^{DMD})$. For each pair $(\lambda_i^{DMD}, \phi_i^{DMD})$ we choose also its complex conjugate pair. To give a graphical illustration, the blue dot in Figure 2 corresponds to λ_0^{DMD}, while the green and orange dots to λ_1^{DMD} and λ_2^{DMD}, respectively, and their complex conjugates. The eigenmodes corresponding to the eigenvalues marked in red in Figure 2 are neglected since they are away from the unit circle.

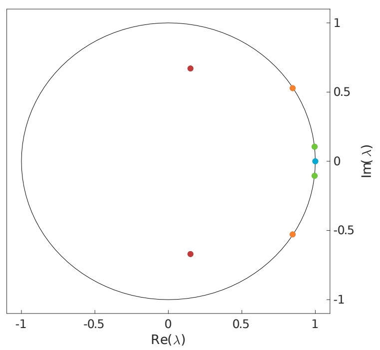

Figure 2. Example of DMD eigenvalues spectrum with parameters $m = 16$, $r = 7$, $\delta t = 0.1$, $\Delta t = 3\delta t$. The blue dot corresponds to λ_0^{DMD}, while the green, orange and red ones to λ_1^{DMD}, λ_2^{DMD} and λ_3^{DMD}, respectively, and their complex conjugates.

To construct the spatial structure $\Sigma(\mathbf{x}, t)$ of the noise, we choose the first $p = 2$ dominant DMD patterns $\phi_{1,2}^{DMD}$ obtained from the low-resolution simulation of the stochastic 2-layer QG system (6). Since the eigenvalues and the eigenfunctions are now complex, each mode is considered together with its complex conjugate, hence Σ reads

$$\begin{aligned}\Sigma(\mathbf{x},t) &= \frac{1}{2}\sum_{i=1}^{2}\left(\left(\mathrm{Re}\left(\lambda_i^{DMD}(t)\right)+\imath\mathrm{Im}\left(\lambda_i^{DMD}(t)\right)\right)\left(\mathrm{Re}\left(\phi_i^{DMD}(\mathbf{x},t)\right)+\imath\mathrm{Im}\left(\phi_i^{DMD}(\mathbf{x},t)\right)\right)+\text{c.c.}\right)\\ &= \sum_{i=1}^{2}\left(\mathrm{Re}\left(\lambda_i^{DMD}(t)\right)\mathrm{Re}\left(\phi_i^{DMD}(\mathbf{x},t)\right)-\mathrm{Im}\left(\lambda_i^{DMD}(t)\right)\mathrm{Im}\left(\phi_i^{DMD}(\mathbf{x},t)\right)\right)\ , \qquad (10)\end{aligned}$$

where $\imath^2=-1$ and c.c. denotes the complex conjugate. Finally we normalize $\Lambda=\Sigma\Sigma^T$ to be either $\Lambda=\lambda_1^{EOF}+\lambda_2^{EOF}\approx 0.36$ or $\Lambda=\sum_{i=1}^{5}\lambda_i^{EOF}\approx 0.47$. This is done to ensure that the noise has equal intensity both with EOFs and DMDs, and therefore have a fairer comparison of the results. To numerically estimate the first two complex conjugate DMD eigenpairs $(\lambda_i^{DMD},\phi_i^{DMD})$ for $i=1,2$, we choose a small time interval $\delta t=0.1$ (recall that δt needs to be chosen sufficiently small to allow for a reliable estimation of the DMD matrix $\mathbf{K}_{\delta t}$ which encodes the dynamics). Furthermore we choose a time window of $m\Delta t=4.8$ time units, which corresponds to roughly half an eddy turnover time for the parameters of our set-up (see Section 5.1 for details), and a separation of snapshots of $\Delta t=3\delta t$ (implying $m=16$). When numerically estimating singular value decompositions, only the first few singular vectors are reliable. An optimal truncation criterion was provided in Gavish and Donoho (2014) which, applied to our data, amounts to setting a low-rank approximation with $r=7$ eigenmodes. We have tested that for the selected values of the parameters, DMD provides a good reconstruction of the dynamics in a time window of length $m\Delta t$ time units, as can be seen in Figure 3, where the actual dynamic is shown alongside the DMD reconstruction. Other sets of parameters corresponding to different time windows spanning between 2 and 10 time units have been tested, but this particular choice was the only one among those tested which does not present two eigenvalues with null imaginary part and real part very close to 1. This second mean-mode cannot be excluded by our procedure since the module of its corresponding eigenvalue is still very close to 1, but by plotting and comparing it to the other modes it can be seen that it is numerically spurious and not dynamically meaningful. We tested also the case with $m=48,\ r=7,\ \Delta t\equiv\delta t=0.1$, i.e. we considered a time window of the same length and instead of sub-sampling - i.e. sampling consecutive snapshots in the same dataset every $\Delta t>\delta t$ - we chose a small value of r, but the results show that sub-sampling is more efficient in filtering out the numerical noise.

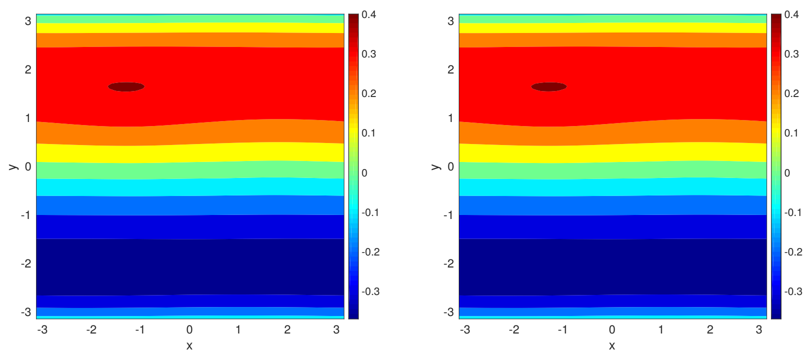

Figure 3. Comparison between the DMD reconstruction (left), and the true dynamics of ψ_T (right). Parameters of the DMD analysis were $m=16,\ r=7,\ \Delta t=3\delta t,\ \delta t=0.1$.

Contrary to EOFs, which require a long off-line simulation to be determined, the DMD pairs $\phi_{1,2}^{DMD}$ and $\lambda_{1,2}^{DMD}$ are computed on the fly after each $m\Delta t$ time units, hence in this case Σ is a function also of time. Since for the first $m\Delta t$ time units the DMDs are not available yet, Σ is initialized using the first two EOFs. For simplicity we do not propagate the DMD modes by means of the Koopman operator, but keep them constant for $m\Delta t$ time units. We have checked that our results do not change much when propagating the DMDs in time to evaluate Σ at each time step. In our setup the DMD modes do not move much away from the initial state in the selected $m\Delta t$ time window, and this might be a reason why we obtained similar outcomes. For more complex models, a computationally cheaper alternative might be to recompute the DMD modes less often and to propagate the DMD modes for longer times.

DMD patterns In Figure 4 we show real and imaginary parts of the first two DMD modes as computed with the aforementioned set of parameters. The mode representing the mean has been neglected and only one of the two modes corresponding to a complex conjugate pair of eigenvalues is displayed. Since the DMD analysis is repeated along the simulation, the resulting modes are not exactly the same for the entire run, but the eddies move in the zonal direction. Moreover the eddies in the first mode slowly shift towards higher latitudes because of the meridional jet movement (as detected by the third EOF eigenvector). Since DMD decomposes the dynamics according to its oscillatory behavior, the jet cannot be represented by a DMD eigenmode for the reason that it is not a wave. Hence in the DMD decomposition of the dynamics, the jet can be noticed only indirectly via its effect on the other modes. This is particularly evident when looking at the first mode as computed at the beginning (Figures 4a-4d) and at the end (Figures 4b-4e) of a simulation, when the difference in the meridional coordinate of the eddies is at its maximum. In this specific case only the first mode is affected by the jet, while the eddies in the other modes retain the same meridional coordinate while revolving in the zonal direction. Hence, for sake of simplicity, we display the second mode only as at the onset of a simulation (Figures 4c-4f).

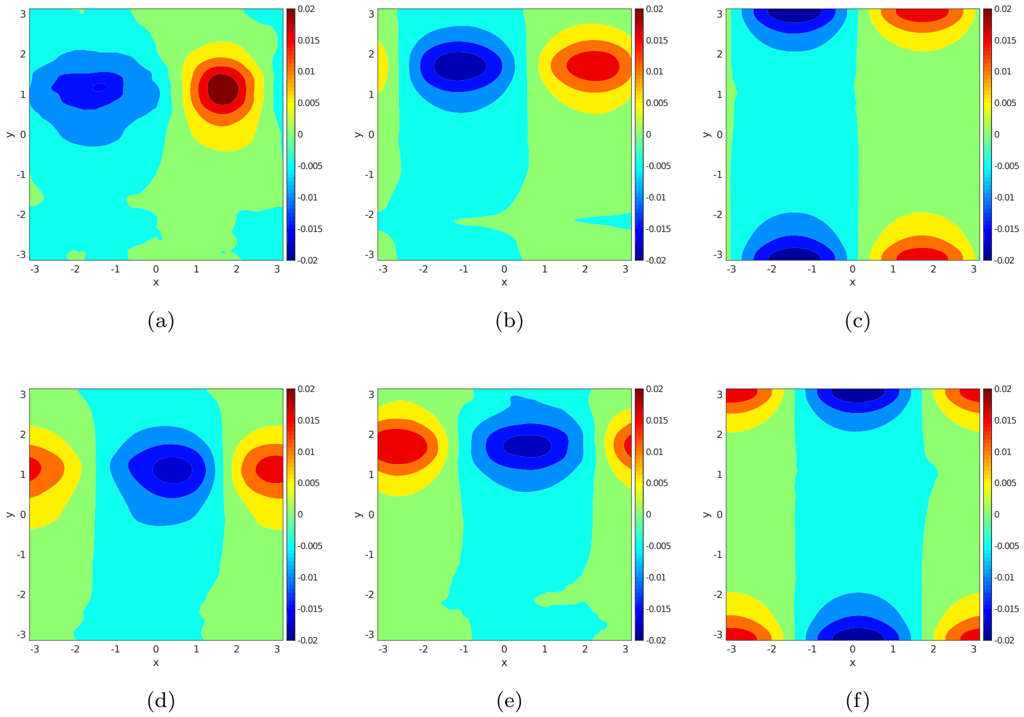

Figure 4. Real (top) and imaginary (bottom) parts of the first DMD mode at the beginning (left) and at the end (middle) of the simulation for $m = 16$, $r = 7$, $\Delta t = 3\delta t$, $\delta t = 0.1$. It can be noticed that in the course of the simulation, the eddies move in the zonal direction and shift towards high latitudes. This movement on the meridional axis is how DMD detects the jet. Real and imaginary parts of the second DMD mode (right) are also displayed. Differently from the eddies of the first mode, here they move only in zonal direction.

Real and imaginary parts of DMD mode number 1 resemble closely EOFs 1-2, although in the DMD mode the eddy patterns look smaller and less regular. Furthermore the eddies are centered in different meridional coordinates. This is likely due to the fact that EOFs capture directly the jet behavior, which is represented by EOF 3. EOFs 4-5 are the most comparable eigenvectors to the second DMD mode (Figure 4c-4f), but significant differences can be spotted for $y \in [0.8,\ 1.8]$, where some eddy structure is present in the EOF vectors but is absent in the DMD mode. This could be an artifact due to the orthogonality constraint of the EOF algorithm.

5 Results

We now present numerical results comparing outputs of a high-resolution simulation of the deterministic forced and damped 2-layer QG model (3) with those of a deterministic low-resolution simulation as well as with the energy-consistent stochastic parametrization (6) run at a low-resolution. Particular emphasis is given on comparing the effect of the respective prescribed spatial noise structures, using either (9) or (10) for EOFs and DMDs, respectively.

5.1 Model setup

As in most current ocean and climate models, we discretize equations (3)-(6) by means of finite differences in a grid-point based framework. The numerical discretization of the Jacobian operator in our QG model is based on the energy and enstrophy conserving scheme by Arakawa (1966). This scheme ensures that energy and enstrophy are conserved for all truncations in the inviscid case. In particular, this scheme does not require any numerical diffusion nor dissipation for numerical stability. For the time stepping of the deterministic part we employ a 4^{th} order explicit Runge-Kutta method, while we use the Euler-Maruyama scheme for the stochastic terms (Pavliotis & Stuart, 2008). The inversion of the Laplacian is achieved in spectral space using Fast Fourier Transforms.

The simulations of the stochastic system (6) are run with a spatial resolution of 128×128 grid points and a time step of $dt = 10^{-3}$. All simulations start from the same initial condition, which we have assured to lie on the attractor by having employed a preceding integration of the deterministic equations at resolution 128×128 for 8000 time units. For each setting of the stochastic system we run the analyses on an ensemble of 10 independent simulations, and compare the outcomes with those of an equivalent deterministic low resolution and with those of a deterministic high resolution simulation. The latter, which will be referred also as reference solution, has been obtained by running the deterministic model (3) on a finer grid of 512×512 grid points. For numerical stability reasons, the reference solution is run with $dt = 10^{-4}$, and its results are projected on the coarser 128×128 spatial grid, to allow for a direct comparison with the outcomes of the respective low-resolution simulations.

5.2 Total energy

Looking at the total energy graphs of the different realizations in the various setups with EOF (orange) and DMD (green) reported in Figure 5, it can be noticed that on average the energy is stable with both techniques, fluctuating by about 1-2% of its absolute value; which is about the same as in a inviscid setting (Gugole & Franzke, 2019). Although the EOF ensemble members show more variance, when only the first two EOFs are used, the system seems to be slightly dissipative in time. This is particularly evident when looking at the ensemble mean (blue line in Figure 5a). The inclusion of EOFs 3-4-5 reduces the dissipative effect, but realizations with a clear increasing trend can be present (Figure 5c). Furthermore, some ensemble members drift away from the high resolution simulation. This also raises questions about long-term stability of the simulations, which we cannot currently resolve because of a too great computational expense of carrying out the simulations for much longer time periods.

On the other hand, the spread of the DMD ensemble members has less variance but well encloses the energy graph of the reference solution. Individual runs are more energetically stable, stay close to the reference solution, and the system seems to be less dissipative compared with the EOF based simulations (Figures 5b-5d). This suggests that the usage of a dynamically adapted noise structure may help the numerical model to remain on the manifold of constant energy and in a dynamically consistent flow regime. In any case deviations from the mean are less than 2%. Hence they might be considered as negligible.

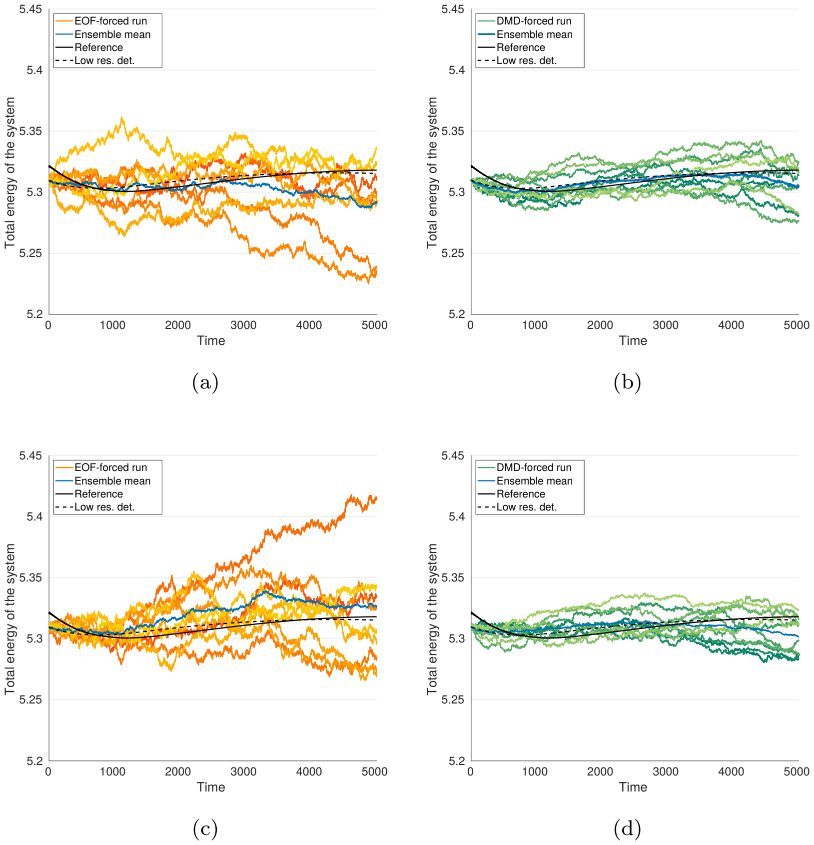

Figure 5. Total energy graphs for stochastic simulations using EOFs (left) or DMDs (right). The case $\Lambda \approx 0.36$, corresponding to the first two EOFs, is displayed in the top row, while the scenario $\Lambda \approx 0.47$ (first five EOFs) is shown in the bottom row. The parameters for DMD have been set as follows: $p = 2,\ m = 16,\ r = 7,\ \delta t = 0.1,\ \Delta t = 3\delta t$. Each stochastic ensemble contains 10 realizations.

5.3 Eddy kinetic energy

In order to compute the eddy kinetic energy (EKE), we first computed the horizontal velocities for the barotropic and baroclinic modes from the respective streamfunctions using

$$u = -\frac{\partial \psi}{\partial y}\ , \quad v = \frac{\partial \psi}{\partial x}\ ,$$

where u is the zonal and v the meridional velocity. Then we considered a time window of k time units to compute the temporal mean velocities, i.e. $\bar{u}_B$, $\bar{v}_B$ and $\bar{u}_T$, $\bar{v}_T$ for barotropic and baroclinic modes respectively. Afterwards for each time unit we computed the deviations from the mean, e.g. $u'_B(t) = u_B(t) - \bar{u}_B$, and used these quantities to compute the EKE for each grid point for all t. As a last step we averaged in time and then also in the zonal direction, therefore the EKE is displayed simply as a function of the meridional direction y (Figures 6-7); or we averaged only in the zonal direction and looked at the time evolution of the EKE projected on the meridional coordinate (Figure 8).

Since the system does not have any annual cycle or similar, we split the time series in windows of 1000 time units, and consider each window individually. Such a length of the time intervals ensures one not to be looking just at transient dynamics while considering small movements of the jet, due to its low frequency variability. Although the time-averaged EKE shows a bi-modal behavior in all windows, the meridional location of the peaks varies according to the jet movement. Hence we want to check how well the stochastic parameterization keeps track of the jet shift. The time-averaged EKE of the baroclinic mode for $t \in [1000,\ 2000]$ and for $t \in [3000,\ 4000]$ in the different stochastic setups with EOF (orange) and DMD (green) are reported in Figures 6 and 7, respectively. The EKE of the barotropic mode shows similar results as for the baroclinic mode, hence, it is not reported here. Figure 8 show the time evolution of the barotropic EKE for $t \in [3000,\ 4000]$ in case of the low resolution deterministic run, one stochastic simulation with EOFs 1-2-3-4-5 and $\Lambda \approx 0.47$, one realization with DMDs and $\Lambda \approx 0.47$, and for the reference solution. Similar conclusions hold also for the time evolution of the baroclinic PV.

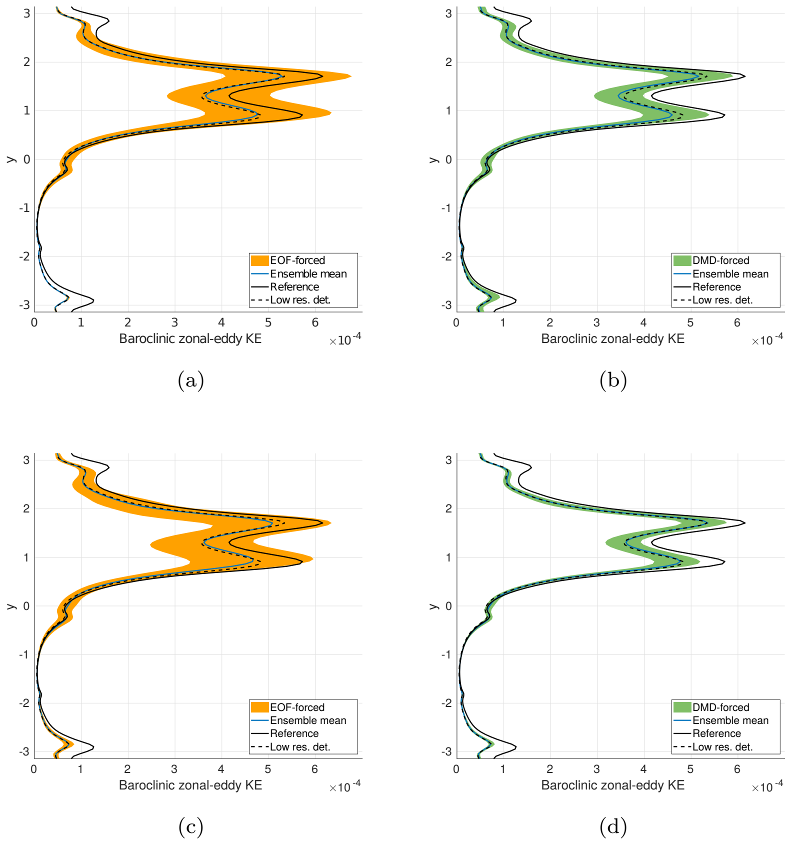

Figure 6. Baroclinic EKE for $t \in [1000,\ 2000]$ for stochastic simulations using EOFs (left) or DMDs (right). The case $\Lambda \approx 0.36$, corresponding to the first two EOFs, is displayed in the top row, while the scenario $\Lambda \approx 0.47$ (first five EOFs) is shown in the bottom row. The parameters for DMD have been set as follows: $p = 2$, $m = 16$, $r = 7$, $\delta t = 0.1$, $\Delta t = 3\delta t$. Each stochastic ensemble contains 10 realizations.

Both for $t \in [1000,\ 2000]$ and $t \in [3000,\ 4000]$ it can be seen that the ensemble forced by EOFs 1-2 has overshoots, which are compensated in the mean (blue line in Figures 6 and 7) by simulations with lower EKE. This is particularly evident at later times (Figure 7), where the uncertainties grow in time and the single members do not display a coherent behavior, i.e. different realizations have different meridional locations for the bi-modal structure and rather different EKE amplitudes. The introduction of EOFs 3-4-5 reduces the overshoots, but has also lower undershoots and does not help the ensemble members to maintain a coherent behavior for longer times. It can be further noticed in Figure 7 that, both with EOFs 1-2 and with EOFs 1-2-3-4-5, the EKE of the stochastic realizations is shifted to too high meridional positions. On the other hand the DMD forced ensembles have less variance and do not always enclose the reference solution, but they remain close to it and they follow quite well the meridional movement of the jet. Furthermore in the DMD ensembles, the uncertainties grow much more slowly in time, allowing the single members to display a coherent behavior also at later stages of the system evolution.

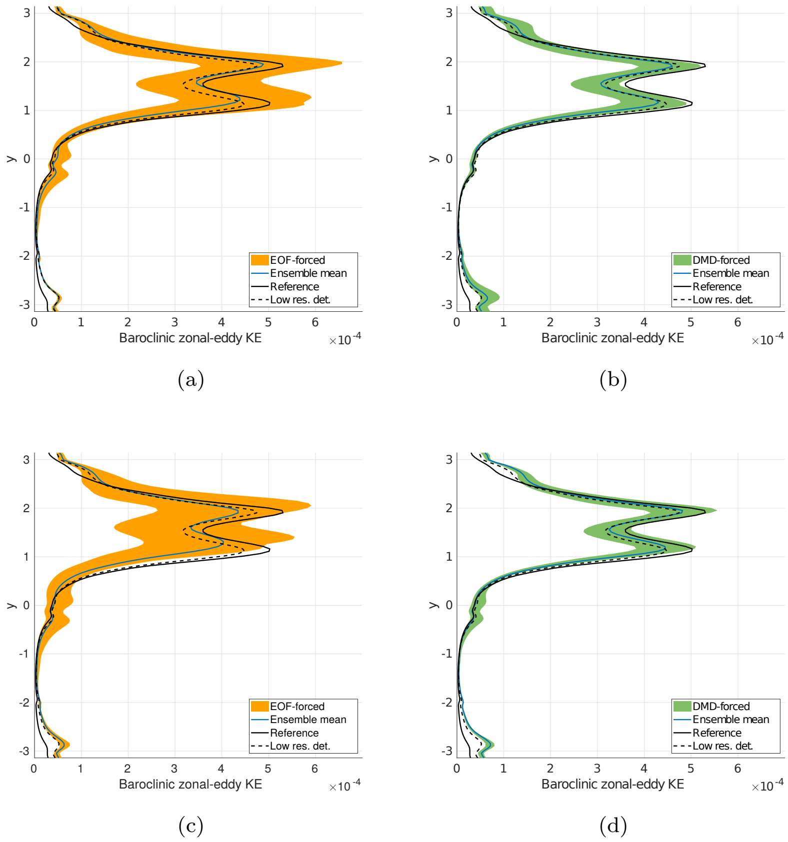

Figure 7. Baroclinic EKE for $t \in [3000,\ 4000]$ for stochastic simulations using EOFs (left) or DMDs (right). The case $\Lambda \approx 0.36$, corresponding to the first two EOFs, is displayed in the top row, while the scenario $\Lambda \approx 0.47$ (first five EOFs) is shown in the bottom row. The parameters for DMD have been set as follows: $p = 2$, $m = 16$, $r = 7$, $\delta t = 0.1$, $\Delta t = 3\delta t$. Each stochastic ensemble contains 10 realizations.

By looking at the time evolution of the barotropic EKE for $t \in [3000,\ 4000]$ (Figure 8) we can notice that while the deterministic low resolution captures the main features, like the meridional shift of the positive EKE at $y \approx 1$, other characteristics are in general underestimated, for instance the amplitude of the eddy kinetic energy at $y \approx -3$ or at $y \approx 2$ for $t \in [3000,\ 3200]$, see Figures 8a and Figure 8d. Some of these properties are indeed recovered by the DMD forced simulation, see for example the enforced EKE at $y \approx 2$ or at $y \approx -3$ in Figure 8c, while the EOFs induced stochastic forcing does not recover them, and creates spurious EKE, see for instance at $y \approx 3$ for $t \in [3400,\ 3600]$ in Figure 8b. Some spurious EKE at $y \approx 3$ can be noticed also in the DMD-forced simulation, but of smaller amplitude with respect to EOFs. Similar outcomes hold also in case of the baroclinic EKE, and are hence not reported here.

These results suggest that the use of a dynamically adapted noise covariance matrix in stochastic parameterizations might be better suited to model phenomena, which do not reach statistical equilibrium, while keeping track of the large scale dynamics. Moreover, considering the wide usage of DMD to detect dynamical features like instabilities and bifurcations (Bagheri, 2013; Budišić et al., 2012; Kutz et al., 2016), a dynamically adapted spatial correlation might more easily foster the system towards tipping points.

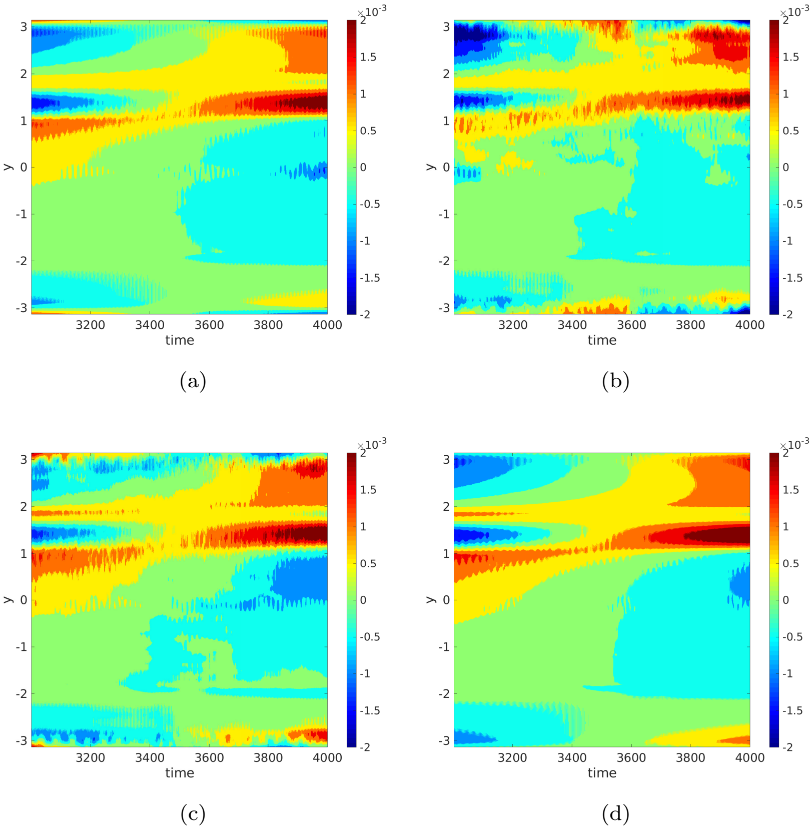

Figure 8. Time evolution of the barotropic EKE projected on the y-axis for the following cases: (a) deterministic low resolution simulation; (b) stochastic simulation with EOFs 1-2-3-4-5, $\Lambda \approx 0.47$; (c) DMD forced stochastic simulation $p = 2$, $m = 16$, $r = 7$, $\delta t = 0.1$ $\Delta t = 3\delta t$, $\Lambda \approx 0.47$; (d) reference solution.

5.4 Flow dynamics

In Figure 9 we show the time evolution of the projection over the zonal coordinate x of the barotropic potential vorticities for the time interval $t \in [1950,\ 2000]$. For a better comparison we removed the zonal mean and plot the resulting eddies. The projection of the baroclinic PV displays similar results and is thus not discussed here. The same graph is shown for the low resolution deterministic simulation (Figure 9a), one realization with EOFs 1-2-3-4-5, $\Lambda \approx 0.47$ (Figure 9b), one with DMD $p = 2,\ m = 16,\ r =$

7, $\delta t = 0.1$, $\Delta t = 3\delta t$, $\Lambda \approx 0.47$ (Figure 9c), and for the reference solution (Figure 9d). Although the eddy phase speed is correctly represented also by the low resolution simulation, the zonal extension and/or intensity of the eddies is underestimated; see for example around $t = 1965$ time units for $-3 < x < 0$. Both stochastic simulations maintain the correct eddy phase speed and help increasing the extension of the eddies, but DMD retains a stronger and less noisy signal. This result confirms the ability of DMD to include the sub-grid scales phenomena without weakening the signal of the larger scales.

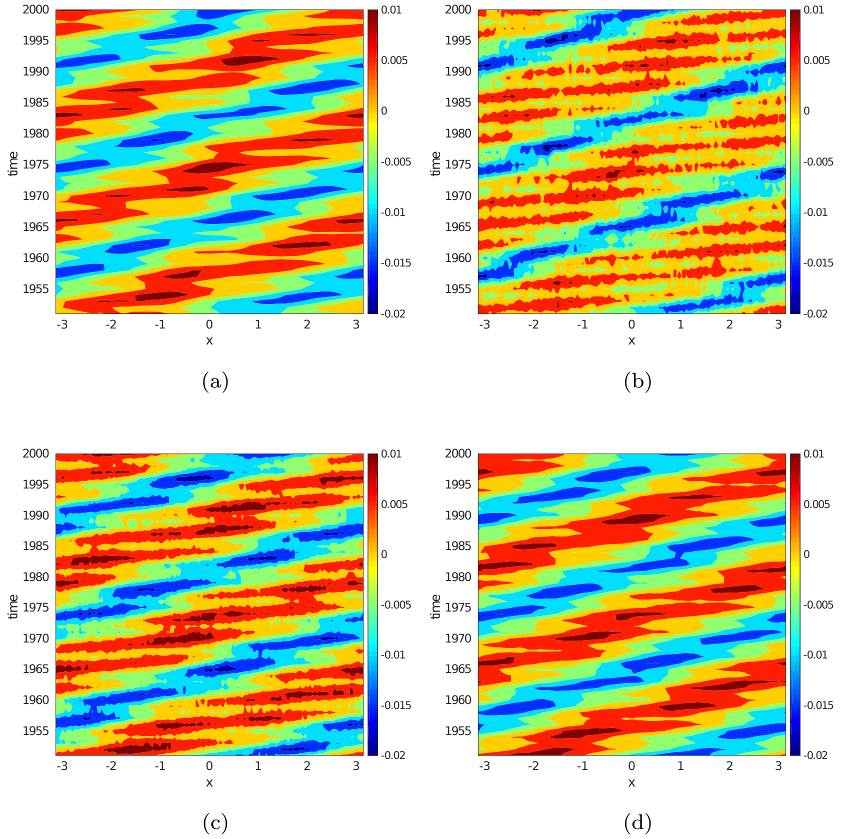

Figure 9. Time evolution in the interval $t \in [1950,\ 2000]$ of the barotropic PVs anomalies with respect to the zonal mean projected on the x-axis for the following cases: (a) deterministic low resolution simulation; (b) stochastic simulation with EOFs 1-2-3-4-5, $\Lambda \approx 0.47$; (c) DMD forced stochastic simulation $p = 2$, $m = 16$, $r = 7$, $\delta t = 0.1$ $\Delta t = 3\delta t$, $\Lambda \approx 0.47$; (d) reference solution.

6 Summary and discussion

In this study we develop a novel way to derive dynamically based noise covariance matrices which are flow-dependent. In the framework of the forced and damped 2-layer QG model we consider an energy-consistent stochastic parameterization based on the projection operator approach (Frank & Gottwald, 2013). As shown in Gugole and Franzke (2019), the definition of the noise spatial structure is of fundamental importance for this parameterization to return physically meaningful results, hence we analyze here two different procedures for its definition. In particular, we investigate a statistical and a dynamical approach by using two different dimension reduction techniques: Empirical Orthogonal Functions and Dynamic Mode Decomposition. The former looks at the variance field of the fluid, while the latter is strictly linked to the Koopman operator and hence to the generator of the dynamics. EOFs have been widely used in the literature, nevertheless there is in general no one-to-one correspondence between the EOF eigenvectors and physical modes (von Storch & Zwiers, 2003). Moreover, being a statistical technique, it requires long time series in order to obtain reliable patterns. In contrast, DMD is able to work with *tall and skinny* matrices (Kutz et al., 2016), hence also with very short time series, and it has oscillatory modes. Therefore the choice of the length of the time series, $m\Delta t$, and the temporal shift between the two input matrices, δt, are crucial and serve as scale selection. In our model set-up we use half a eddy turn-over time as a physically based time interval to recompute the DMD. Since DMD decomposes the dynamics according to its local in time oscillatory behavior, its modes and the noise covariance, have to be recomputed periodically. This is a new approach in stochastic parameterizations allowing the noise covariance to be a function of time, while typically a fixed noise covariance is used during the whole realization.

Total energy graphs reveal that the EOF ensembles are either more dissipative or might include realizations with a clear increasing trend. On the other hand DMD runs are individually more energetically consistent with the high-resolution control simulation, suggesting that a dynamically adapted noise structure might help the model to stay on the manifold of constant energy. This might suggest that this approach may have the potential to lead to more dynamically consistent simulations and long-term stability. By analyzing the eddy kinetic energy, it has been discovered that in case of EOFs the uncertainties grow faster, which induce the single ensemble members to display very different amplitudes of the EKE. Furthermore the location of the bi-modal structure of the EKE ensemble mean is not well defined among the individual realizations, and it is in general moved towards too high meridional locations. The DMD forced ensembles instead are able to follow the meridional jet shift and well catch the meridional location of the double-peak also at later times. Moreover the uncertainties grow more slowly, allowing the individual members to display a coherent behavior during the entire simulation and to stay close to the reference solution. Finally, the field dynamics time evolution in the DMD ensembles retain a stronger and less noisy signal.

As regards computational time, DMD is very cheap, can reliably deal with rather short time series and does not need extra computations beforehand, but can be run alongside the main code. These aspects allow the DMD algorithm to periodically reanalyze the dynamics and redefine the noise covariance accordingly. Hence it is a very good candidate to parameterize scales undergoing phase transitions, or which do not reach statistically stable profiles. This might also allow DMD to be used for scale-adaptive parameterization schemes. Moreover, due to its close link to the Koopman operator, and to its ability to detect instabilities and bifurcations within dynamical systems (Bagheri, 2013; Kutz et al., 2016), it might foster the system to reach tipping points or regime transitions.

Our results suggest that a dynamically adapted spatial structure should be considered in future developments of stochastic parameterizations. This is further motivated by the physics. Not only are the large scales affected by the small scales, but also the small-scale processes are influenced by the large-scale motions. Hence physically correct parameterizations of the unresolved scales should allow the sub-grid processes to be influenced by the resolved modes. Furthermore, the propagation of the DMD modes by means of the Koopman operator might be seen as a sort of memory term, which in turn has been shown to be important in parameterization schemes (Franzke et al., 2015; Gottwald et al., 2017; Hu & Franzke, 2017; Sakradzija, Seifert, & Heus, 2015). However, more detailed studies are required to establish what kind of relation, if any, exists between the propagation of the DMD modes and memory terms.

Acknowledgments

This work is a contribution to the project M2 of the Collaborative Research Center TRR181 *Energy Transfer in Atmosphere and Ocean* funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 274762653. Model scripts are publicly available on GitHub©: `https://github.com/FGugole/2L_QG_EOF_DMD`. Data have been generated by use of the aforementioned scripts. The authors thank G. A. Gottwald and S. Juricke for discussions related to this work.

References

Arakawa, A. (1966). Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. part i. *Journal of Computational Physics*, *1*(1), 119–143.

Badin, G., & Crisciani, F. (2018). *Variational formulation of fluid and geophysical fluid dynamics - mechanics, symmetries and conservation laws.* Springer Berlin.

Bagheri, S. (2013). Koopman-mode decomposition of the cylinder wake. *Journal of Fluid Mechanics*, *726*, 596623. doi: 10.1017/jfm.2013.249

Berloff, P. S. (2005). Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, *529*, 7195. doi: 10.1017/S0022112005003393

Berner, J., Achatz, U., Batte, L., Bengtsson, L., Camara, A., Christensen, H., ... Yano, J. (2017). Stochastic parameterization: Toward a new view of weather and climate models. *Bull. Amer. Meteorol. Soc.*, *98*(3), 565–588.

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

Budišić, M., Mohr, R., & Mezić, I. (2012). Applied Koopmanism. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *22*(4), 047510. Retrieved from `https://doi.org/10.1063/1.4772195` doi: 10.1063/1.4772195

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

Cotter, C. J., Crisan, D., Holm, D. D., Pan, W., & Shevchenko, I. (2019). Numerically modelling stochastic Lie transport in fluid dynamics. *Multiscale Model. Simul.*, *17*, 192–232.

Cotter, C. J., Gottwald, G. A., & Holm, D. D. (2017). Stochastic partial differential fluid equations as a diffusive limit of deterministic lagrangian multi-time dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2205), 20170388. Retrieved from `https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2017.0388` doi: 10.1098/rspa.2017.0388

Dubrulle, B. (2019). Beyond Kolmogorov cascades. *Journal of Fluid Mechanics*, *867*, P1. doi: 10.1017/jfm.2019.98

Dwivedi, S., Franzke, C. L. E., & Lunkeit, F. (2019). Energetically consistent stochastic and deterministic kinetic energy backscatter schemes for atmospheric models. *Q. J. Roy. Meteorol. Soc.*, *145*, 1–11.

Feldstein, S. B. (1998). An observational study of the intraseasonal poleward propagation of zonal mean flow anomalies. *Journal of the atmospheric sciences*, *55*(15), 2516–2529.

Frank, J. E., & Gottwald, G. A. (2013). Stochastic homogenization for an energy conserving multi-scale toy model of the atmosphere. *Physica D*, *254*, 46–56.

Franzke, C. L. E., Majda, A. J., & Vanden-Eijnden, E. (2005). Low-order stochastic mode reduction for a realistic barotropic model climate. *J. Atmos. Sci.*, *62*, 1722–1745.

Franzke, C. L. E., O'Kane, T., Berner, J., Williams, P., & Lucarini, V. (2015). Stochastic climate theory and modelling. *WIREs Climate Change*, *6*, 63–78.

Franzke, C. L. E., Oliver, M., Rademacher, J., & Badin, G. (2019). Systematic multi-scale methods for geophysical flows. In A. Iske & C. Eden (Eds.), *Energy transfers in atmosphere and ocean* (pp. 1–51). Springer.

Gardiner, C. W. (2009). *Stochastic methods: A handbook for the natural and social sciences* (Vol. 4). Springer Berlin.

Gavish, M., & Donoho, D. L. (2014, Aug). The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Transactions on Information Theory*, *60*(8), 5040-5053. doi: 10.1109/TIT.2014.2323359

Gottwald, G., Crommelin, D., & Franzke, C. L. E. (2017). Stochastic climate theory. In C. L. E. Franzke & T. O'Kane (Eds.), *Nonlinear and stochastic climate dynamics* (pp. 209–240). Cambridge: Cambridge University Press.

Gugole, F., & Franzke, C. L. E. (2019). Numerical development and evaluation of an energy conserving conceptual stochastic climate model. *Mathematics for Climate and Weather Forecasting*, *5*, 45-64.

Hermanson, L., Hoskins, B., & Palmer, T. (2009). A comparative method to evaluate and validate stochastic parametrizations. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *135*(642), 1095–1103.

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proc. R. Soc., A 471*, 20140963.

Holton, J. R. (2004). *An Introduction To Dynamic Meteorology* (Fourth Edi ed.).

Hu, G., & Franzke, C. L. E. (2017). Data assimilation in a multi-scale model. *Mathematics of Climate and Weather Forecasting*, *3*(1), 118–139.

Imkeller, P., & von Storch, J.-S. (Eds.). (2001). *Stochastic climate models* (Vol. 49). Basel: Birkhauser Verlag.

James, I., & Dodd, J. (1996). A mechanism for the low-frequency variability of the mid-latitude troposphere. *Quarterly Journal of the Royal Meteorological Society*, *122*(533), 1197–1210.

Jansen, M. F., & Held, I. M. (2014). Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Modelling*, *80*, 36–48.

Jansen, M. F., Held, I. M., Adcroft, A., & Hallberg, R. (2015). Energy budget-based backscatter in an eddy permitting primitive equation model. *Ocean Modelling*, *94*, 15–26.

Jovanovic, M. R., Schmid, P. J., & Nichols, J. W. (2014). Sparsity-promoting dynamic mode decomposition. *Physics of Fluids*, *26*(2), 024103. Retrieved from `https://doi.org/10.1063/1.4863670` doi: 10.1063/1.4863670

Juricke, S., Danilov, S., Kutsenko, A., & Oliver, M. (2019). Ocean kinetic energy backscatter parametrizations on unstructured grids: Impact on mesoscale turbulence in a channel. *Ocean Modelling*, *138*, 51 - 67. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1463500318303846` doi: https://doi.org/10.1016/j.ocemod.2019.03.009

Kutz, J., Brunton, S., Brunton, B., & Proctor, J. (2016). *Dynamic mode decomposition*. Philadelphia, PA: Society for Industrial and Applied Mathematics. Retrieved from `https://epubs.siam.org/doi/abs/10.1137/1.9781611974508` doi: 10.1137/1.9781611974508

Lasota, A., & Mackey, M. C. (1994). *Chaos, fractals and noise - stochastic aspects of dynamics*. Springer-Verlag.

Leith, C. (1996). Stochastic models of chaotic systems. *Physica D: Nonlinear Phenomena*, *98*, 481–491.

Majda, A., Franzke, C. L. E., & Crommelin, D. (2009). Normal forms for reduced stochastic climate models. *Proc. Natl. Acad. Sci. USA*, *106*, 3649–3653. doi: doi:10.1073/pnas.0900173106

Mémin, E. (2014). Fluid flow dynamics under location uncertainty. *Geophysical and*

Astrophysical Fluid Dynamics, *108*, 119–46.

Mezić, I. (2013). Analysis of fluid flows via spectral properties of the Koopman operator. *Annual Review of Fluid Mechanics*, *45*(1), 357-378. Retrieved from https://doi.org/10.1146/annurev-fluid-011212-140652 doi: 10.1146/annurev-fluid-011212-140652

Noack, B. R., Stankiewicz, W., Morzyaski, M., & Schmid, P. J. (2016). Recursive dynamic mode decomposition of transient and post-transient wake flows. *Journal of Fluid Mechanics*, *809*, 843-872. doi: 10.1017/jfm.2016.678

Palmer, T., & Williams, P. (Eds.). (2010). *Stochastic physics and climate modelling.* Cambridge: Cambridge University Press.

Pavliotis, G. A., & Stuart, A. (2008). *Multiscale methods: averaging and homogenization.* Springer Science & Business Media.

Penland, C. (1989). Random forcing and forecasting using principal oscillation pattern analysis. *Monthly Weather Review*, *117*(10), 2165-2185. Retrieved from https://doi.org/10.1175/1520-0493(1989)117<2165:RFAFUP>2.0.CO;2 doi: 10.1175/1520-0493(1989)117⟨2165:RFAFUP⟩2.0.CO;2

Penland, C., & Magorian, T. (1993). Prediction of nino 3 sea surface temperatures using linear inverse modeling. *J. Climate*, *6*(6), 1067–1076.

Porta Mana, P., & Zanna, L. (2014). Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, *79*, 1–20.

Resseguier, V., Mémin, E., & Chapron, B. (2017a). Geophysical flows under location uncertainty, part ii quasi-geostrophy and efficient ensemble spreading. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 177-208. Retrieved from https://doi.org/10.1080/03091929.2017.1312101 doi: 10.1080/03091929.2017.1312101

Resseguier, V., Mémin, E., & Chapron, B. (2017b). Geophysical flows under location uncertainty, part i random transport and general models. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 149-176. Retrieved from https://doi.org/10.1080/03091929.2017.1310210 doi: doi:10.1080/03091929.2017.1310210

Resseguier, V., Pan, W., & Fox-Kemper, B. (2019). Data-driven versus self-similar parameterizations for stochastic advection by Lie transport and location uncertainty. *Nonlinear Processes in Geophysics Discussions*, *2019*, 1–37. Retrieved from https://www.nonlin-processes-geophys-discuss.net/npg-2019-54/ doi: 10.5194/npg-2019-54

Riehl, H., Yeh, T., & La seur, N. E. (1950). A study of variations of the general circulation. *Journal of Meteorology*, *7*(3), 181–194.

Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., & Henningson, D. S. (2009). Spectral analysis of nonlinear flows. *J. Fluid Mech.*, *641*, 115–127. Retrieved from https://doi.org/10.1017/S0022112009992059 doi: 10.1017/S0022112009992059

Sakradzija, M., Seifert, A., & Heus, T. (2015). Fluctuations in a quasi-stationary shallow cumulus cloud ensemble. *Nonlin. Proc. Geophys.*, *22*(1), 65–85.

Salmon, R. (1988). Hamiltonian fluid mechanics. *Annual Review of Fluid Mechanics*, *20*(1), 225-256. Retrieved from https://doi.org/10.1146/annurev.fl.20.010188.001301 doi: 10.1146/annurev.fl.20.010188.001301

Sardeshmukh, P. D., & Sura, P. (2009). Reconciling non-gaussian climate statistics with linear dynamics. *J. Climate*, *22*(5), 1193–1207.

Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, *656*, 528. doi: 10.1017/S0022112010001217

Schmid, P. J., Li, L., Juniper, M. P., & Pust, O. (2011, Jun 01). Applications of the dynamic mode decomposition. *Theoretical and Computational Fluid Dynamics*, *25*(1), 249–259. Retrieved from https://doi.org/10.1007/s00162-010-0203-9 doi: 10.1007/s00162-010-0203-9

Schmid, P. J., Meyer, K. E., & Pust, O. (2009). Dynamic mode decomposition and proper orthogonal decomposition of flow in a lid-driven cylindrical cavity. In

8th international symposium on particle image veclocimetry – PIV09.

Shepherd, T. G. (1990). Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Advances in Geophysics*, *32*, 287–338.

Stensrud, D. J. (2007). *Parameterization schemes.* Cambridge University Press.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., & Kutz, J. N. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, *1*, 391. Retrieved from `http://aimsciences.org//article/id/1dfebc20-876d-4da7-8034-7cd3c7ae1161` doi: 10.3934/jcd.2014.1.391

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

von Storch, H. (1995). Spatial patterns: Eofs and cca. In H. von Storch & A. Navarra (Eds.), *Analysis of climate variability: Applications of statistical techniques* (pp. 227–257). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from `https://doi.org/10.1007/978-3-662-03167-4_13` doi: 10.1007/978-3-662-03167-4_13

von Storch, H., & Zwiers, F. W. (2003). *Statistical analysis in climate research.* Cambridge University Press.

Whitaker, J. S., & Sardeshmukh, P. D. (1998). A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.*, *55*(2), 237–258.

Williams, M. O., Kevrekidis, I. G., & Rowley, C. W. (2015). A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. *J. Nonlinear Sci.*, *25*(6), 1307–1346. Retrieved from `https://doi.org/10.1007/s00332-015-9258-5` doi: 10.1007/s00332-015-9258-5

Zhang, Y., & Held, I. M. (1999). A linear stochastic model of a gcms midlatitude storm tracks. *J. Atmos. Sci.*, *56*(19), 3416–3435.

Zurita-Gotor, P., Held, I. M., & Jansen, M. F. (2015). Kinetic energy-conserving hyperdiffusion can improve low resolution atmospheric models. *J. Adv. Mol. Earth Sys.*, *7*(3), 1117–1135.