Fluctuations of finite-time Lyapunov exponents in an intermediate-complexity atmospheric model: a multivariate and large-deviation perspective

Frank Kwasniok
Department of Mathematics, University of Exeter, Exeter, United Kingdom
Correspondence: Frank Kwasniok (F.Kwasniok@exeter.ac.uk)

Abstract. The stability properties as characterised by the fluctuations of finite-time Lyapunov exponents around their mean values are investigated in a three-level quasi-geostrophic atmospheric model with realistic mean state and variability. Firstly, the covariance structure of the fluctuation field is examined. In order to identify dominant patterns of collective excitation an empirical orthogonal function (EOF) analysis of the fluctuation field of all of the finite-time Lyapunov exponents is performed. The three leading modes are patterns where the most unstable Lyapunov exponents fluctuate in phase. These modes are virtually independent of the integration time of the finite-time Lyapunov exponents. Secondly, large-deviation rate functions are estimated from time series of finite-time Lyapunov exponents based on the probability density functions and using the Legendre transform method. Serial correlation in the time series is properly accounted for. A large-deviation principle can be established for all of the Lyapunov exponents. Convergence is rather slow for the most unstable exponent, becomes faster when going further down in the Lyapunov spectrum, is very fast for the near-neutral and weakly dissipative modes, and turns slow again for the strongly dissipative modes at the end of the Lyapunov spectrum. The curvature of the rate functions at the minimum is linked to the corresponding elements of the diffusion matrix. Also the joint large-deviation rate function for the first and the second Lyapunov exponent is estimated.

1 Introduction

The atmosphere is a high-dimensional nonlinear chaotic dynamical system; its time evolution is characterised by sensitivity to initial conditions (Lorenz, 1963; Kalnay, 2003). As a consequence predictability is limited; small errors in the initial states progressively grow under the time evolution until the forecast eventually becomes useless, that is, indistinguishable from the invariant measure or climatology of the system. Understanding the structure of this inherent instability is key to improve forecasts at all timescales.

Sensitivity to initial conditions and perturbation growth in nonlinear dynamical systems are often quantified using Lyapunov exponents (LEs) (e.g., Eckmann and Ruelle, 1985; Ott, 2002; Pikovsky and Politi, 2016). They describe the asymptotic growth or decay of infinitesimal perturbations. A system is chaotic if it has at least one positive Lyapunov exponent. However, the predictability properties may vary substantially across state space. Finite-time (or local) Lyapunov exponents (FTLEs) allow a characterisation of the stability of a particular initial state with respect to a predefined prediction horizon.

LEs have been calculated for various geophysical fluid systems, ranging from highly truncated atmospheric models (Legras and Ghil, 1985), to intermediate-complexity atmospheric models (Vannitsem and Nicolis, 1997; Schubert and Lucarini, 2015), to coupled atmosphere-ocean models (Vannitsem and Lucarini, 2016). A review has been published recently by Vannitsem (2017). Models tuned to realistic conditions turn out to possess quite a large number of positive LEs corresponding to a high-dimensional chaotic attractor.

The present paper investigates the fluctuations of FTLEs in an intermediate-complexity atmospheric model with realistic mean state and variability. It focuses on two aspects which have so far found little attention in the context of geophysical fluid systems. Firstly, the covariance structure of the fluctuation field of the FTLEs is studied by means of a principal component (PC) or empirical orthogonal function (EOF) analysis. Secondly, we are looking at the large-deviation behaviour of the FTLEs at long integration times.

The paper is organised as follows: In section 2 the atmospheric model is described. The methodology which consists of calculating LEs, the multivariate fluctuation analysis and the large-deviation theory is outlined in sections 3, 4 and 5. The results are presented and discussed in section 6. Some conclusions are drawn in section 7.

2 The atmospheric model

A quasi-geostrophic (QG) three-level model on the sphere, formulated in pressure coordinates, is used here as dynamical framework. The model is identical to that introduced by Kwasniok (2007) except for the horizontal resolution and the coefficient of hyperviscosity. A very similar model was introduced by Marshall and Molteni (1993). The dynamical equations are

$$\frac{\partial q_i}{\partial t} + J(\Psi_i, q_i) = D_i + S_i, \qquad i = 1,2,3 \tag{1}$$

where Ψ_i and q_i are the streamfunction and the potential vorticity at level i and J denotes the Jacobian operator on the sphere. All variables are nondimensional using the radius of the earth as the unit of length and the inverse of the angular velocity of the earth as the unit of time. The three pressure levels are located at 250, 500 and 750 hPa. Potential vorticity and the streamfunction are related by

$$q_1 = \nabla^2\Psi_1 - R_{1,2}^{-2}(\Psi_1 - \Psi_2) + f \tag{2}$$

$$q_2 = \nabla^2\Psi_2 + R_{1,2}^{-2}(\Psi_1 - \Psi_2) - R_{2,3}^{-2}(\Psi_2 - \Psi_3) + f \tag{3}$$

$$q_3 = \nabla^2\Psi_3 + R_{2,3}^{-2}(\Psi_2 - \Psi_3) + f + f_0 h \tag{4}$$

where ∇ is the horizontal gradient operator and f is the Coriolis parameter. The Rossby deformation radii $R_{1,2}$ and $R_{2,3}$ have dimensional values of 575 km and 375 km. The function $h = h(\lambda, \mu)$ represents a nondimensional topography which is related to the actual dimensional topography of the earth $h^* = h^*(\lambda, \mu)$ by $h = h^*/H$, where H is a scale height set to 8 km and f_0 is the Coriolis parameter at an average geographic latitude taken to be 45°N.

The dissipative terms are given as

$$D_1 = \tau_{\mathrm{N}}^{-1} R_{1,2}^{-2}(\Psi_1 - \Psi_2) - k_{\mathrm{H}} \nabla^8 \hat{q}_1 \quad (5)$$

$$D_2 = -\tau_{\mathrm{N}}^{-1} R_{1,2}^{-2}(\Psi_1 - \Psi_2) + \tau_{\mathrm{N}}^{-1} R_{2,3}^{-2}(\Psi_2 - \Psi_3) - k_{\mathrm{H}} \nabla^8 \hat{q}_2 \quad (6)$$

$$D_3 = -\tau_{\mathrm{N}}^{-1} R_{2,3}^{-2}(\Psi_2 - \Psi_3) - \tau_{\mathrm{E}}^{-1} \nabla^2 \Psi_3 - k_{\mathrm{H}} \nabla^8 \hat{q}_3 \quad (7)$$

They are Newtonian temperature relaxation with a radiative timescale of $\tau_{\mathrm{N}} = 25$ days, Ekman damping on the lowest level with a spindown timescale of $\tau_{\mathrm{E}} = 1.5$ days, and a strongly scale-selective horizontal diffusion of vorticity and temperature. The $\hat{q}_i$ is the time-dependent part of the potential vorticity at level i, that is, $\hat{q}_i = q_i - f - \delta_{i3} f_0 h$. The coefficient of horizontal diffusion $k_{\mathrm{H}} = \tau_{\mathrm{H}}^{-1}[n_{\mathrm{m}}(n_{\mathrm{m}}+1)]^{-4}$ is such that harmonics of total wavenumber $n_{\mathrm{m}} = 21$ are damped at a timescale of $\tau_{\mathrm{H}} = 1.5$ days. The terms $S_i = S_i(\lambda, \mu)$ are diabatic sources of potential vorticity which are independent of time but spatially varying.

The model is considered on the northern hemisphere. The boundary condition of no meridional flow, $v_i(\lambda, 0) = 0$, that is, vanishing streamfunction, $\Psi_i(\lambda, 0) = 0$, is applied at the equator on all three model levels. The horizontal discretization is spectral, triangularly truncated at total wavenumber $n_{\mathrm{m}} = 21$. The number of degrees of freedom is 231 for each level and $N = 693$ in total. The model is integrated in time using the third-order Adams–Bashforth scheme with a constant step size of 1 h.

The variables of the QG model are listed in Table 1; the model parameters are listed in Table 2 with their dimensional and nondimensional values.

In order to get a model behaviour close to that of the real atmosphere, the forcing terms S_i are determined from ECMWF reanalysis data by requiring that when computing potential vorticity tendencies for a large number of observed atmospheric fields, the average of these tendencies must be zero (Roads, 1987), in order for the ensemble of reanalysis data states to be representative of a statistically stable long-term behaviour of the QG model. The timescale of horizontal diffusion τ_{H} is determined such that the slope of the kinetic energy spectrum at the truncation level in the model matches that in the reanalysis data. See Kwasniok (2007) for details on the parameter tuning procedure. The QG model exhibits in a long-term integration a remarkably realistic mean state and variability pattern of streamfunction and potential vorticity (see Table 3).

3 Lyapunov exponents

We consider a nonlinear autonomous dynamical system with state vector $\mathbf{x} = (x_1, \ldots, x_N)^{\mathrm{T}}$ governed by the evolution equations

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}). \quad (8)$$

The linearised dynamics of a small perturbation $\delta\mathbf{x}$ are given as

$$\frac{d}{dt}\delta\mathbf{x} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\delta\mathbf{x}. \quad (9)$$

The propagation of the perturbation between time t_0 with initial state $\mathbf{x}_0 = \mathbf{x}(t_0)$ and time t ($t > t_0$) can be written as

$$\delta\mathbf{x}(t) = \mathbf{M}(\mathbf{x}_0, t)\delta\mathbf{x}(t_0) \quad (10)$$

Table 1. Variables and fields in the QG model and their nondimensionalization with the earth radius $a = 6.371 \times 10^6\,\mathrm{m}$ and the angular velocity of the earth $\Omega = 7.292 \times 10^{-5}\,\mathrm{s}^{-1}$.

Symbol	Description	Unit	Nondimensionalisation
t	time	s	Ω^{-1}
$\lambda \quad (0 \le \lambda < 2\pi)$	longitude (eastward)		
$\phi \quad (0 \le \phi \le \pi/2)$	latitude		
$\mu = \sin\phi \quad (0 \le \mu \le 1)$	sine of latitude		
Ψ_i	streamfunction at level i	$\mathrm{m}^2\,\mathrm{s}^{-1}$	$a^2\Omega$
q_i	potential vorticity at level i	s^{-1}	Ω
$\hat{q}_i$	time-dependent part of potential vorticity at level i	s^{-1}	Ω
$f = 2\mu$	Coriolis parameter	s^{-1}	Ω
h	topography of the earth	m	H
S_i	diabatic forcing at level i	s^{-2}	Ω^2
$u_i = -\sqrt{1-\mu^2}\,(\partial\Psi_i/\partial\mu)$	zonal velocity (eastward)	$\mathrm{m}\,\mathrm{s}^{-1}$	$a\Omega$
$v_i = (1/\sqrt{1-\mu^2})\,(\partial\Psi_i/\partial\lambda)$	meridional velocity (northward)	$\mathrm{m}\,\mathrm{s}^{-1}$	$a\Omega$

Table 2. Parameters in the QG model

Symbol	Description	Dimensional value	Nondimensional value
$R_{1,2}$	Rossby deformation radius between levels 1 and 2	575 km	9.025×10^{-2}
$R_{2,3}$	Rossby deformation radius between levels 2 and 3	375 km	5.886×10^{-2}
τ_N	timescale of temperature relaxation	25 days	50π
τ_E	timescale of Ekman damping	1.5 days	3π
τ_H	timescale of horizontal diffusion at wavenumber 21	1.5 days	3π
f_0	Coriolis parameter at 45°N	$1.031 \times 10^{-4}\,\mathrm{s}^{-1}$	$\sqrt{2}$
H	scale height	8 km	

where $\mathbf{M}$ is the resolvent matrix. If the system is ergodic then according to the theorem by Oseledets (1968) the limit

$$\mathbf{S} = \lim_{t\to\infty} (\mathbf{M}^\mathrm{T}\mathbf{M})^{\frac{1}{2(t-t_0)}} \tag{11}$$

exists and is the same for almost all initial conditions $\mathbf{x}_0$. The (global) LEs are defined as

$$\lambda_j = \log\omega_j, \qquad j = 1,\ldots,N, \tag{12}$$

where $\{\omega_j\}_{j=1}^N$ are the positive eigenvalues of the matrix $\mathbf{S}$. The set of all LEs, usually presented in non-increasing order, is called the *Lyapunov spectrum*. The LEs are independent of norm.

Table 3. Pattern correlation of various fields in the QG model with the corresponding fields in ECMWF reanalysis data

Level	$\langle \Psi_i \rangle$	$\sqrt{\langle \Psi_i'^2 \rangle}$	$\sqrt{\langle \hat{q}_i'^2 \rangle}$
250 hPa	0.99	0.99	0.97
500 hPa	0.99	0.99	0.98
750 hPa	0.96	0.97	0.94

In order to characterise perturbation growth or decay over a finite integration time τ the FTLEs $\Lambda_j^{(\tau)}(\mathbf{x}_0)$ are introduced. There are three different definitions of FTLEs. One can compute them by making reference to the backward, forward or covariant Lyapunov vectors; see, e.g., Kuptsov and Parlitz, 2012 for a review. In the limit of large integration time τ, which is the main focus of the present study, all of the three definitions become more and more equivalent (Kuptsov and Politi, 2011; Pazó et al., 2013). We here refer to the backward FTLEs as they are easiest to compute. They are calculated using the standard algorithm based on the Gram–Schmidt orthogonalisation (Shimada and Nagashima, 1979; Benettin et al., 1980). An ensemble of N linearly independent perturbations is initialised and integrated forward in time together with the nonlinear model trajectory. A transient period is discarded for the trajectory to settle on the attractor of the system and the perturbations to converge to the backward Lyapunov vectors. Then after every integration time interval $\Delta\tau$ the perturbations are re-orthonormalised using a QR-decomposition performed via the Gram–Schmidt procedure. The FTLEs are obtained as

$$\Lambda_j^{(\Delta\tau)}(\mathbf{x}_\alpha) = \Lambda_{j,\alpha}^{(\Delta\tau)} = \frac{1}{\Delta\tau} \log R_{jj}(t_\alpha, t_{\alpha+1}), \qquad \alpha = 0, \ldots, L-1, \tag{13}$$

where $R_{jj}(t_\alpha, t_{\alpha+1})$ are the diagonal elements of the triangular matrix $\mathbf{R}$ in the QR-decomposition resulting from the integration between times t_α and $t_{\alpha+1}$. We have $t_\alpha = t_0 + \alpha\Delta\tau$ and $\mathbf{x}_\alpha = \mathbf{x}(t_\alpha)$. The FTLEs $\Lambda_j^{(\tau)}$ for larger integration times $\tau = n\Delta\tau$ are obtained by averaging over n consecutive values of $\Lambda_j^{(\Delta\tau)}$:

$$\Lambda_{j,\alpha}^{(\tau)} = \frac{1}{n} \sum_{i=0}^{n-1} \Lambda_{j,\alpha+i}^{(\Delta\tau)}, \qquad \alpha = 0, \ldots, L-1 \tag{14}$$

For all integration times τ, we keep time series of FTLEs of the same length L, $\{\Lambda_{j,\alpha}^{(\tau)}\}_{\alpha=0}^{L-1}$, characterising the stability of the states $\{\mathbf{x}_\alpha\}_{\alpha=0}^{L-1}$ over the time horizon τ.

The FTLEs depend on the scalar product chosen in the Gram–Schmidt orthogonalisation procedure. We here use the total energy scalar product with its associated total energy norm (Ehrendorfer, 2000; Kwasniok, 2007). The dependence of the FTLEs on the norm becomes increasingly weaker in the limit of large integration time τ.

The FTLEs are related to the global LEs by

$$\lim_{\tau\to\infty} \Lambda_j^{(\tau)}(\mathbf{x}_0) = \lambda_j \tag{15}$$

for almost all initial states $\mathbf{x}_0$ and

$$\langle \Lambda_j^{(\tau)} \rangle = \lambda_j \tag{16}$$

for all τ where $\langle\cdot\rangle$ denotes an ensemble average over the attractor of the system which for ergodic systems can be estimated as a mean over a long time series.

4 Multivariate fluctuation analysis

The vector of global Lyapunov exponents is defined as $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^{\mathrm{T}}$ and the fluctuation field as $\boldsymbol{\Lambda}^{(\tau)} - \boldsymbol{\lambda} = \left(\Lambda_1^{(\tau)} - \lambda_1, \ldots, \Lambda_N^{(\tau)} - \lambda_N\right)^{\mathrm{T}}$. We study the correlations between the fluctuations of the FTLEs; to do this, preferred patterns of collective excitation are extracted. A canonical approach is a principal component (PC) or empirical orthogonal function (EOF) analysis based on the scaled covariance matrix $\mathbf{D}^{(\tau)}$ defined as

$$\mathbf{D}^{(\tau)} = \left\langle \left(\boldsymbol{\Lambda}^{(\tau)} - \boldsymbol{\lambda}\right)\left(\boldsymbol{\Lambda}^{(\tau)} - \boldsymbol{\lambda}\right)^{\mathrm{T}} \right\rangle \tau. \tag{17}$$

In the limit of large integration time τ we expect convergence to the diffusion matrix $\mathbf{D}$ (Kuptsov and Politi, 2011; Pikovsky and Politi, 2016):

$$\mathbf{D} = \lim_{\tau\to\infty} \mathbf{D}^{(\tau)} \tag{18}$$

The scaled covariance matrix is here estimated from the time series of the FTLEs as

$$\mathbf{D}^{(\tau)} = \frac{\tau}{L}\sum_{\alpha=0}^{L-1}\left(\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda}\right)\left(\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda}\right)^{\mathrm{T}} \tag{19}$$

The eigenvalues and eigenvectors of the symmetric, positive definite matrix $\mathbf{D}^{(\tau)}$ are calculated:

$$\mathbf{D}^{(\tau)}\mathbf{e}_j^{(\tau)} = \nu_j^{(\tau)}\mathbf{e}_j^{(\tau)} \tag{20}$$

The eigenvalues $\{\nu_j^{(\tau)}\}_{j=1}^N$ are arranged in non-increasing order. The eigenvectors form an orthonormal system:

$$\mathbf{e}_j^{(\tau)} \cdot \mathbf{e}_k^{(\tau)} = \delta_{jk} \tag{21}$$

The fluctuation field of the FTLEs is expanded as

$$\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda} = \sum_{j=1}^{N} y_{j,\alpha}^{(\tau)}\mathbf{e}_j^{(\tau)} \tag{22}$$

with $y_{j,\alpha}^{(\tau)} = \mathbf{e}_j^{(\tau)} \cdot (\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda})$. The principal components $\{y_j^{(\tau)}\}_{j=1}^N$ are uncorrelated and their variance is given by the corresponding eigenvalue:

$$\left\langle y_j^{(\tau)} y_k^{(\tau)} \right\rangle = \frac{1}{L}\sum_{\alpha=0}^{L-1} y_{j,\alpha}^{(\tau)} y_{k,\alpha}^{(\tau)} = \nu_j^{(\tau)}\delta_{jk} \tag{23}$$

The steepness or complexity of the eigenvalue spectrum is characterised by the fraction of variance explained by the principal component $y_j^{(\tau)}$ given as

$$r_j^{(\tau)} = \frac{\nu_j^{(\tau)}}{\sum_{k=1}^{N}\nu_k^{(\tau)}} \tag{24}$$

and the cumulative fraction of variance given as

$$c_j^{(\tau)} = \frac{\sum_{k=1}^{j} \nu_k^{(\tau)}}{\sum_{k=1}^{N} \nu_k^{(\tau)}} \tag{25}$$

As a possible further step, one may try to link the covariance structure of the FTLEs with investigations of the angles between the covariant Lyapunov vectors and the degree of entanglement and interaction of the various unstable and stable directions in tangent space (Yang et al., 2009). This is related to the hyperbolicity and the inertial manifold of the system.

5 Large-deviation theory for FTLEs

Large-deviation theory (Kifer, 1990; Touchette, 2009) is a powerful approach from statistical physics for estimating the probability of rare events with many applications. It has recently been applied to the behaviour of FTLEs at long integration times (Kuptsov and Politi, 2011; Laffargue et al., 2013; Johnson and Meneveau, 2015). Large-deviation theory is in the following briefly described in the form in which it is used in the present study.

5.1 Univariate theory

For a sequence of n identically distributed but not necessarily independent random variables, $\{X_i\}_{i=1}^{n}$, the sample mean

$$A_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{26}$$

is an unbiased estimator of and converges to the true mean, $\langle X \rangle$, as $n \to \infty$. By the Gärtner–Ellis theorem (Touchette, 2009), if the *scaled cumulant generating function* (SCGF)

$$\gamma(\theta) = \lim_{n\to\infty} \frac{1}{n} \log \left\langle e^{n\theta A_n} \right\rangle \tag{27}$$

exists and is differentiable everywhere then A_n follows a large-deviation principle,

$$p(A_n = z) \sim \exp[-nI(z)], \tag{28}$$

where the *large-deviation rate function* $I(z)$ is independent of n and given as the *Legendre–Fenchel transform* of the SCGF:

$$I(z) = \sup_{\theta\in\mathbb{R}} [\theta z - \gamma(\theta)] \tag{29}$$

The rate function $I(z)$ is non-negative and strictly convex. It has a unique zero and minimum at $z^* = \langle X \rangle$, that is, $I(\langle X \rangle) = 0$ and $I'(\langle X \rangle) = 0$. The curvature of the rate function at the minimum is given as (Touchette, 2009)

$$I''(\langle X \rangle) = \frac{1}{\lim\limits_{n\to\infty} n\langle (A_n - \langle A_n \rangle)^2 \rangle}. \tag{30}$$

The rate function $I(z)$ is calculated using the *block averaging method* (Rohwer et al., 2015). New random variables Y_i are introduced as block averages

$$Y_i = \frac{1}{b}\sum_{l=1}^{b} X_{(i-1)b+l}, \qquad i = 1,\ldots,n/b, \tag{31}$$

where the block size b is chosen large enough for the random variables Y_i to be independent. For simplicity, we assume that n is an integer multiple of b. The sample mean is then

$$A_n = \frac{b}{n}\sum_{i=1}^{n/b} Y_i \tag{32}$$

and we have

$$\begin{aligned}
\gamma(\theta) &= \lim_{n\to\infty}\frac{1}{n}\log\left\langle e^{b\theta\sum_{i=1}^{n/b}Y_i}\right\rangle \\
&= \lim_{n\to\infty}\frac{1}{n}\log\left\langle\prod_{i=1}^{n/b}e^{b\theta Y_i}\right\rangle \\
&= \lim_{n\to\infty}\frac{1}{n}\log\prod_{i=1}^{n/b}\left\langle e^{b\theta Y_i}\right\rangle \\
&= \frac{1}{b}\log\left\langle e^{b\theta Y}\right\rangle
\end{aligned} \tag{33}$$

and

$$\begin{aligned}
I(z) &= \sup_{\theta'\in\mathbb{R}}\left[\theta' z - \frac{1}{b}\log\langle e^{b\theta' Y}\rangle\right] \\
&= \frac{1}{b}\sup_{\theta\in\mathbb{R}}[\theta z - \log\langle e^{\theta Y}\rangle] \qquad (\theta = b\theta')
\end{aligned} \tag{34}$$

In view of eq.(14), FTLEs immediately lend themselves to large-deviation theory. For large integration time τ, one would expect the probability density of the FTLE $\Lambda_j^{(\tau)}$ to follow a large-deviation principle,

$$p\left(\Lambda_j^{(\tau)} = z\right) \sim \exp[-\tau I_j(z)] \tag{35}$$

with a large-deviation rate function $I_j(z)$ which is independent of τ and given as

$$I_j(z) = \frac{1}{\tau}\sup_{\theta\in\mathbb{R}}[\theta z - \log\langle e^{\theta\Lambda_j^{(\tau)}}\rangle]. \tag{36}$$

The rate function $I_j(z)$ has a unique zero and minimum at $z^* = \lambda_j$, that is, $I_j(\lambda_j) = 0$ and $I_j'(\lambda_j) = 0$. The curvature of the rate function at the minimum is linked to the diffusion matrix $\mathbf{D}$ as

$$I_j''(\lambda_j) = \frac{1}{\lim\limits_{\tau\to\infty}\left\langle\left(\Lambda_j^{(\tau)} - \lambda_j\right)^2\right\rangle\tau} = D_{j,j}^{-1} \tag{37}$$

A second-order Taylor expansion of the rate function in the vicinity of λ_j,

$$I_j(z) \approx \frac{1}{2} I_j''(\lambda_j)(z-\lambda_j)^2, \tag{38}$$

corresponds to a Gaussian probability density with mean λ_j and variance $D_{j,j}/\tau$, recovering the central limit theorem (CLT) as a limit case of large-deviation theory.

5.2 Estimating the rate function

There are two ways of estimating the rate functions $I_j(z)$ from data: via the probability density function (cf., eq.(35)) or via the Legendre transform (cf., eq.(36)).

5.2.1 Probability density function approach

By inverting eq.(35) we have

$$I_j(z) = -\lim_{\tau\to\infty} \frac{1}{\tau} \log p\left(\Lambda_j^{(\tau)} = z\right) \tag{39}$$

We take a maximum likelihood approach for estimating the rate function (Kwasniok, 2019a, 2019b). The probability density of $\Lambda_j^{(\tau)}$ is modelled as

$$p\left(\Lambda_j^{(\tau)} = z\right) = \frac{1}{Z_j^{(\tau)}} \exp\left[-U_j^{(\tau)}(z)\right] \tag{40}$$

with normalisation constant

$$Z_j^{(\tau)} = \int_{-\infty}^{\infty} \exp\left[-U_j^{(\tau)}(z)\right] dz \tag{41}$$

The potential function $U_j^{(\tau)}(z)$ is expanded into a polynomial basis in standardized variables:

$$U_j^{(\tau)}(z) = \sum_{i=1}^{M} \beta_i^{(\tau)} \left(\frac{z-\lambda_j}{\sigma_j^{(\tau)}}\right)^i \tag{42}$$

Here $\sigma_j^{(\tau)}$ is the standard deviation of $\Lambda_j^{(\tau)}$. The parameter M determines the complexity of the model. In order to have a normalisable probability density, we need M to be even and $\beta_M > 0$. The expansion coefficients $\{\beta_i\}_{i=1}^M$ are determined by maximising the likelihood function of the data $\{\Lambda_{j,\alpha}^{(\tau)}\}_{\alpha=0}^{L-1}$. This is a convex optimisation problem with a unique maximum which is numerically stable to solve. Model selection is performed with the Bayesian information criterion. For details of the algorithm, see Kwasniok (2019a, b).

The estimate of the rate function is given as

$$I_j(z) = \frac{1}{\tau}\left[U_j^{(\tau)}(z) - U_j^{(\tau)}(z^*)\right] \tag{43}$$

where z^* denotes the position of the minimum of the potential function $U_j^{(\tau)}(z)$. Note that, for finite τ, we do not necessarily have $z^* = \lambda_j$ as the mode of the probability density of $\Lambda_j^{(\tau)}$ may be different from its mean if the distribution is skewed; but we always have $z^* \to \lambda_j$ as $\tau \to \infty$. One would now estimate $I_j(z)$ from the probability density function of $\Lambda_j^{(\tau)}$ for various large values of τ and look for convergence.

The maximum likelihood method tends to provide very smooth and convex rate functions although convexity is not strictly guaranteed. It clearly improves on earlier work (e.g., Johnson and Meneveau, 2015) using histogram or kernel density estimates for the probability density and treating the normalisation constant only in the Gaussian approximation.

5.2.2 Legendre transform approach

Alternatively, the rate functions $I_j(z)$ can be determined by numerically implementing the Legendre transform of eq.(36) (Rohwer et al., 2015; Kwasniok, 2019b) with the moment generating function estimated by the sample mean over the time series:

$$\langle e^{\theta \Lambda_j^{(\tau)}} \rangle = \frac{1}{L} \sum_{\alpha=0}^{L-1} e^{\theta \Lambda_{j,\alpha}^{(\tau)}} \tag{44}$$

For each θ, this is a convex optimisation problem with a unique solution if any exists. Rate functions obtained via the Legendre transform method are guaranteed to be strictly convex with a unique zero and minimum at $z^* = \lambda_j$.

Rate function estimates from the Legendre transform method converge as soon as τ is large enough for successive values of $\Lambda_j^{(\tau)}$ over non-overlapping time intervals, $\Lambda_{j,\alpha}^{(\tau)}$ and $\Lambda_{j,\alpha+n}^{(\tau)}$, to be independent and thus for eq.(33) to hold. However, this gives no indication whether or not the probability density function actually is already in the large-deviation limit. Therefore we here consider both rate function estimates side by side.

5.3 Estimating the diffusion coefficients

The diffusion coefficients $D_{j,j}$ can be obtained from both rate function estimates as the inverse of the curvature at the minimum (cf., eq.(37). They can also be estimated directly from the time series of the FTLEs according to eq.(19). It can be shown that the estimates from the Legendre transform-based rate function and from the time series are always the same; any differences just stem from the error of the finite-difference approximation of the curvature as the Legendre transform is not available in closed form. For a Gaussian probability density model, that is, $M = 2$ in eq.(42), the diffusion coefficient estimates from the probability density-based rate function and from the time series are exactly the same; otherwise they are different.

5.4 Multivariate theory

The large-deviation analysis can be extended to a multivariate approach (Kuptsov and Politi, 2011; Johnson and Meneveau, 2015). Let now $\mathbf{\Lambda}^{(\tau)}$ denote the column vector of any K-dimensional subset of the N FTLEs and $\boldsymbol{\lambda}$ the corresponding vector of global LEs. We have $1 \leq K \leq N$ where $K = N$ corresponds to the full system and $K = 1$ recovers the univariate analysis. For

large integration time τ, the joint probability density function of the K FTLEs would then follow a large-deviation principle,

$$p\left(\mathbf{\Lambda}^{(\tau)}=\mathbf{z}\right)\sim\exp[-\tau I(\mathbf{z})], \tag{45}$$

where the *joint large-deviation rate function* $I(\mathbf{z})$ is independent of τ and given as the *multivariate Legendre–Fenchel transform*

$$I(\mathbf{z})=\frac{1}{\tau}\sup_{\boldsymbol{\theta}\in\mathbb{R}^K}\left[\boldsymbol{\theta}^{\mathrm{T}}\mathbf{z}-\log\langle e^{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{\Lambda}^{(\tau)}}\rangle\right] \tag{46}$$

The joint rate function $I(\mathbf{z})$ is non-negative and strictly convex. It has a unique zero and minimum at $\mathbf{z}^*=\boldsymbol{\lambda}$, that is, $I(\boldsymbol{\lambda})=0$ and $\partial I/\partial z_j=0$ at $\mathbf{z}=\boldsymbol{\lambda}$. The Hessian matrix of the joint rate function at the minimum is linked to the diffusion matrix $\mathbf{D}$ as

$$\left.\frac{\partial^2 I}{\partial z_j\partial z_k}\right|_{\mathbf{z}=\boldsymbol{\lambda}}=Q_{j,k}=\left(\mathbf{D}^{-1}\right)_{j,k}, \tag{47}$$

where here $\mathbf{D}$ denotes the $K\times K$ part of the diffusion matrix corresponding to the K retained FTLEs. A second-order Taylor expansion of the joint rate function in the vicinity of $\boldsymbol{\lambda}$,

$$I(\mathbf{z})\approx\frac{1}{2}(\mathbf{z}-\boldsymbol{\lambda})^{\mathrm{T}}\mathbf{Q}(\mathbf{z}-\boldsymbol{\lambda}), \tag{48}$$

corresponds to a multivariate Gaussian probability density with mean $\boldsymbol{\lambda}$ and covariance matrix $(\tau\mathbf{Q})^{-1}$, recovering the central limit theorem (CLT).

5.5 Estimating the joint rate function

There are again two ways of estimating the joint rate function from the time series of FTLEs: via the probability density function (cf., eq.(45)) or via the Legendre transform (cf., eq.(46)).

5.5.1 Probability density function approach

By inverting eq.(45) we get

$$I(\mathbf{z})=-\lim_{\tau\to\infty}\frac{1}{\tau}\log p\left(\mathbf{\Lambda}^{(\tau)}=\mathbf{z}\right) \tag{49}$$

The probability density of $\mathbf{\Lambda}^{(\tau)}$ is modelled as

$$p\left(\mathbf{\Lambda}^{(\tau)}=\mathbf{z}\right)=\frac{1}{Z^{(\tau)}}\exp\left[-U^{(\tau)}(\mathbf{z})\right] \tag{50}$$

with normalisation constant

$$Z^{(\tau)}=\int_{\mathbb{R}^K}\exp\left[-U^{(\tau)}(\mathbf{z})\right]d^K\mathbf{z} \tag{51}$$

The potential function $U^{(\tau)}(\mathbf{z})$ is expanded into suitable multinomial basis functions as

$$U^{(\tau)}(\mathbf{z})=\sum_{i=1}^{J}\beta_i^{(\tau)}\phi_i(\mathbf{z}) \tag{52}$$

subject to appropriate conditions to ensure a normalisable probability density. The expansion coefficients $\{\beta_i\}_{i=1}^{J}$ are determined from the time series of the FTLEs $\{\mathbf{\Lambda}_\alpha^{(\tau)}\}_{\alpha=0}^{L-1}$ via maximum likelihood (Kwasniok, 2019a, b) which is a convex optimisation problem with unique solution.

The estimate of the joint rate function is

$$I(\mathbf{z}) = \frac{1}{\tau}\left[U^{(\tau)}(\mathbf{z}) - U^{(\tau)}(\mathbf{z}^*)\right] \tag{53}$$

where $\mathbf{z}^*$ denotes the position of the minimum of the potential function $U^{(\tau)}(\mathbf{z})$ which for finite τ is not necessarily equal to $\boldsymbol{\lambda}$.

5.5.2 Legendre transform approach

Alternatively, the joint rate function $I(\mathbf{z})$ can be determined via the multivariate Legendre transform of eq.(46) (Kwasniok, 2019b) with the moment generating function estimated as the sample mean over the time series:

$$\langle e^{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{\Lambda}^{(\tau)}}\rangle = \frac{1}{L}\sum_{\alpha=0}^{L-1} e^{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{\Lambda}_\alpha^{(\tau)}} \tag{54}$$

Again, this is a convex optimisation problem and rate functions obtained from the Legendre transform method are guaranteed to be strictly convex with a unique zero and minimum at $\mathbf{z}^* = \boldsymbol{\lambda}$.

5.6 Estimating the diffusion matrix

The diffusion matrix $\mathbf{D}$ (or the part of it corresponding to the K considered FTLEs) can be obtained from both joint rate function estimates as the inverse of the Hessian matrix at the minimum (cf., eq.(47)). It can also be estimated directly from the time series of the FTLEs as given in eq.(19). The estimates from the Legendre transform-based joint rate function and from the time series are always the same, apart from errors in the finite-difference approximation of the second derivatives. The diffusion matrix estimates from the probability density-based joint rate function and from the time series are the same if the model in eq.(52) is a multivariate Gaussian probability density; otherwise they are different.

The different methods for estimating the rate function and the diffusion matrix in the univariate and the multivariate case are summarised in Tables 4 and 5.

In high-dimensional systems it is usually too ambitious a task to determine $I(\mathbf{z})$ beyond the Gaussian approximation for the full system. We here restrict ourselves to the bivariate case $K = 2$.

Table 4. Methods for estimating the rate function

	Probability density	Legendre transform
Univariate	$I_j(z) = \frac{1}{\tau}\left[U_j^{(\tau)}(z) - U_j^{(\tau)}(z^*)\right]$	$I_j(z) = \frac{1}{\tau}\sup_{\theta\in\mathbb{R}}\left[\theta z - \log\left(\frac{1}{L}\sum_{\alpha=0}^{L-1} e^{\theta\Lambda_{j,\alpha}^{(\tau)}}\right)\right]$
Multivariate	$I(\mathbf{z}) = \frac{1}{\tau}\left[U^{(\tau)}(\mathbf{z}) - U^{(\tau)}(\mathbf{z}^*)\right]$	$I(\mathbf{z}) = \frac{1}{\tau}\sup_{\boldsymbol{\theta}\in\mathbb{R}^K}\left[\boldsymbol{\theta}^{\mathrm{T}}\mathbf{z} - \log\left(\frac{1}{L}\sum_{\alpha=0}^{L-1} e^{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Lambda}_\alpha^{(\tau)}}\right)\right]$

Table 5. Methods for estimating the diffusion matrix

	Probability density	Legendre transform	Time series
Univariate	$D_{j,j}^{-1} = I_j''(z^*)$	$D_{j,j}^{-1} = I_j''(\lambda_j)$	$D_{j,j} = \frac{\tau}{L}\sum_{\alpha=0}^{L-1}\left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j\right)^2$
Multivariate	$(\mathbf{D}^{-1})_{j,k} = \frac{\partial^2 I}{\partial z_j \partial z_k}\Big\vert_{\mathbf{z}=\mathbf{z}^*}$	$(\mathbf{D}^{-1})_{j,k} = \frac{\partial^2 I}{\partial z_j \partial z_k}\Big\vert_{\mathbf{z}=\boldsymbol{\lambda}}$	$D_{j,k} = \frac{\tau}{L}\sum_{\alpha=0}^{L-1}\left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j\right)\left(\Lambda_{k,\alpha}^{(\tau)} - \lambda_k\right)$

6 Results

6.1 Lyapunov exponents

Time series of FTLEs of the QG model of length $L = 25000$ with a basic integration time $\Delta\tau = 1$ day are generated as described in Section 3. The (global) LEs are calculated as

$$\lambda_j = \frac{1}{L}\sum_{\alpha=0}^{L-1}\Lambda_{j,\alpha}^{(\Delta\tau)}. \tag{55}$$

Figure 1 displays the Lyapunov spectrum of the QG model. There are 91 positive LEs. The largest LE is estimated as $\lambda_1 = 0.344$/day, corresponding to an e-folding time of perturbation growth of 2.9 days which appears to be realistic for the real atmosphere. The spectrum starts off quite steep and then flattens at the near-zero exponents. For example, there are 69 LEs between 0.05/day and -0.05/day. The spectrum becomes steeper again at the trailing very stable exponents. Overall, there is a continuous spectrum of timescales with no clear timescale separation. This is in accordance with previous results for QG models (Vannitsem and Nicolis, 1997; Schubert and Lucarini, 2015; Vannitsem and Lucarini, 2016) and is probably because QG equations are scale-filtered equations.

Figure 2 shows the standard deviation of the fluctuations of the FTLEs around their mean values calculated as

$$\sigma_j^{(\tau)} = \left[\frac{1}{L}\sum_{\alpha=0}^{L-1}\left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j\right)^2\right]^{1/2}. \tag{56}$$

The standard deviation monotonically decreases with increasing integration time τ for all exponents. The fluctuations are largest for the leading LEs and then quickly decrease. They increase again towards the end of the Lyapunov spectrum with a particularly sharp increase for the most stable exponents at the very end of the spectrum. This is in line with similar findings

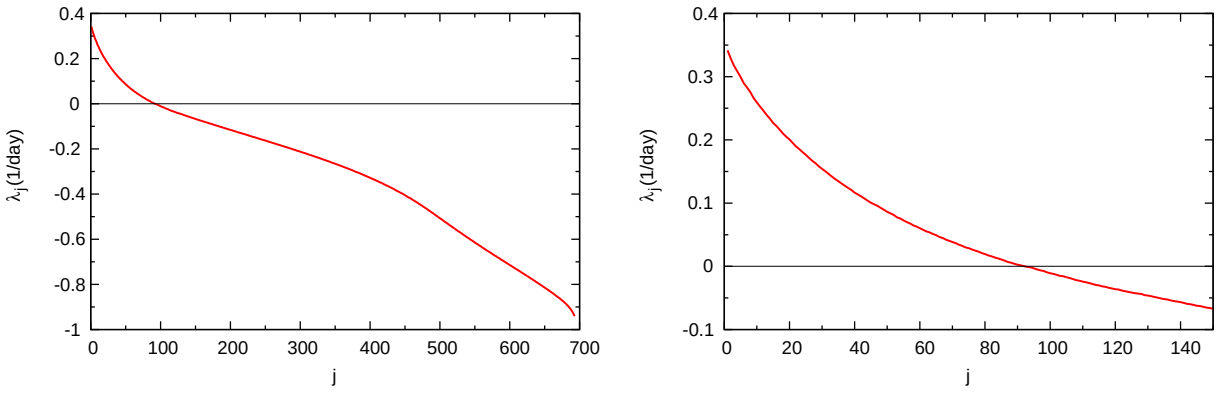

Figure 1. (a) Lyapunov spectrum of the QG model. (b) Close-up of (a).

in simple spatially extended systems (Kuptsov and Politi, 2011; Pazó et al., 2013) as well as in a QG atmosphere-ocean model (Vannitsem and Lucarini, 2016).

The scaled standard deviation $\sigma_j^{(\tau)}\tau^{1/2}$ shows clear convergence for all of the exponents at $\tau = 10 - 15$ days, that is, the scaled variance converges to the diagonal elements $D_{j,j}$ of the diffusion matrix $\mathbf{D}$. Convergence is reached at about $\tau = 10$ days for almost all of the exponents; it is particularly fast for the near-neutral and the weakly dissipative exponents where it is reached already at $\tau = 5 - 10$ days.

There is a kink-like feature at $j \approx 125$, separating regions with different slopes of the standard deviation. It is possible that this is linked to a distinction of the covariant Lyapunov vectors into interacting 'physical modes' and hyperbolically separated dissipative 'isolated modes', and thus to the dimension of the inertial manifold or the effective number of degrees of freedom of the system (Yang et al., 2009).

6.2 Multivariate fluctuation analysis

Figure 3 shows the explained variance and the cumulative explained variance of the principal components of the scaled Lyapunov fluctuations. There are three leading modes, then the eigenvalue spectrum sharply flattens off. The fraction of variance explained by the leading modes increases with increasing integration time τ. Going from $\tau = 1$ day to $\tau = 20$ days, the variance explained by the first principal component increases from just below 5% to more than 12%, and the variance explained by the second principal component increases from about 2% to more than 4%. However, due to the flatness of the bulk of the eigenvalue spectrum, even in the diffusion limit a substantial number of modes is necessary to explain large parts of the fluctuation variance. The eigenvalue spectrum is still not fully converged at $\tau = 20$ days. It is not completely clear what the reason for this is. There may be some indication that the off-diagonal elements of the diffusion matrix converge slightly more slowly than the diagonal elements. But there is probably also a finite sample size effect. With increasing τ, the time series of the FTLEs contain less and less uncorrelated information and fail to fully sample the high-dimensional covariance matrix which leads to an overestimation of the variance of the leading principal components.

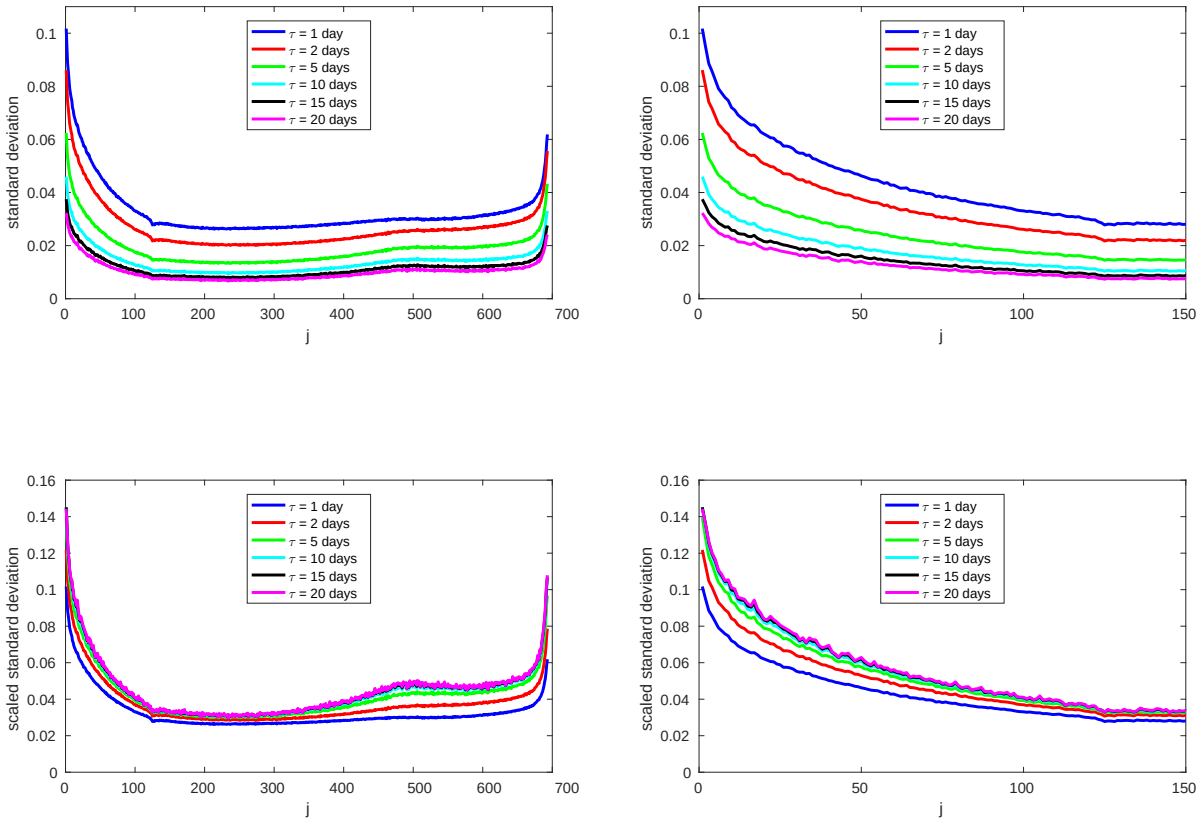

Figure 2. (a) Standard deviation $\sigma_j^{(\tau)}$ of the FTLEs. (b) Close-up of (a). (c) Scaled standard deviation $\sigma_j^{(\tau)}\tau^{1/2}$ of the FTLEs. (d) Close-up of (c).

In Figure 4 the three leading EOFs are displayed. The modes are largely independent of the integration time τ and have converged at about $\tau = 10$ days. The first EOF shows a pattern where all of the leading FTLEs fluctuate in phase. This incorporates all of the positive exponents and extends to the weakly dissipative ones. Then there is some negative correlation with the dissipative exponents in the second half of the Lyapunov spectrum. In the second EOF again the leading FTLEs fluctuate in phase; this here encompasses about the first 40 exponents. Then there is some negative correlation with the weakly dissipative exponents and substantial positive correlation with the strongly dissipative exponents at the end of the Lyapunov spectrum. The third EOF has the very stable exponents at the end of the spectrum fluctuating in phase and the most unstable exponents fluctuating in phase with each other, out of phase with the dissipative ones.

Complementary to the EOF analysis, Figure 5 shows the correlation of selected FTLEs with each of the other FTLEs for $\tau = 1$ day and $\tau = 15$ days. The pattern of the correlations is the same for both integration times but the amplitudes are very

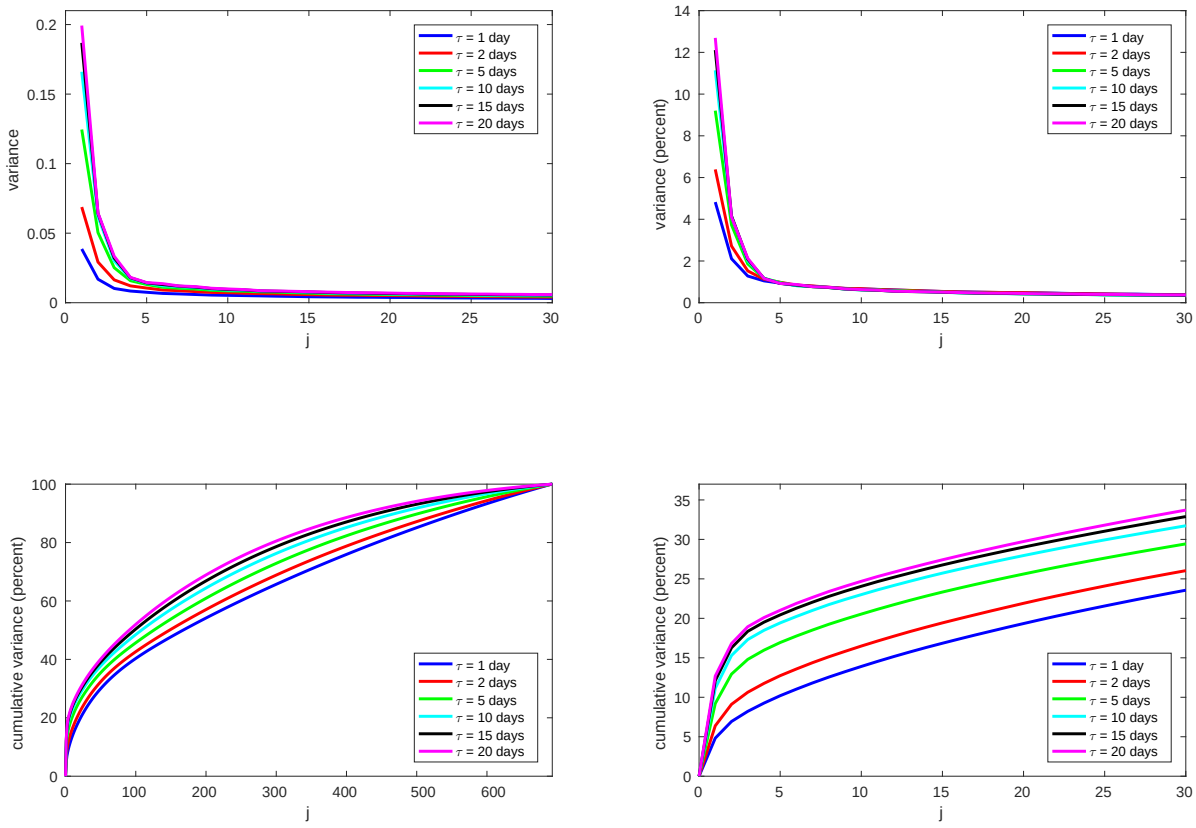

Figure 3. (a) Variance of the principal components of the finite-time Lyapunov fluctuations. (b) Fraction of variance. (c) Cumulative fraction of variance. (d) Close-up of (c).

low for $\tau = 1$ day and build up at larger integration times. This is in line with the results from the EOF analysis. The FTLEs have predominantly positive correlations with neighbouring exponents; these are strongest for the most unstable and the most stable exponents and weaker in between. There are also some relatively weak long-range correlations across the Lyapunov spectrum.

6.3 Large-deviation analysis

6.3.1 One-dimensional approach

We now investigate whether the fluctuations of the FTLEs obey a large deviation principle. As representative examples we look at the first and the fifth exponent as two strongly unstable modes, at the zero exponent, at a weakly dissipative exponent

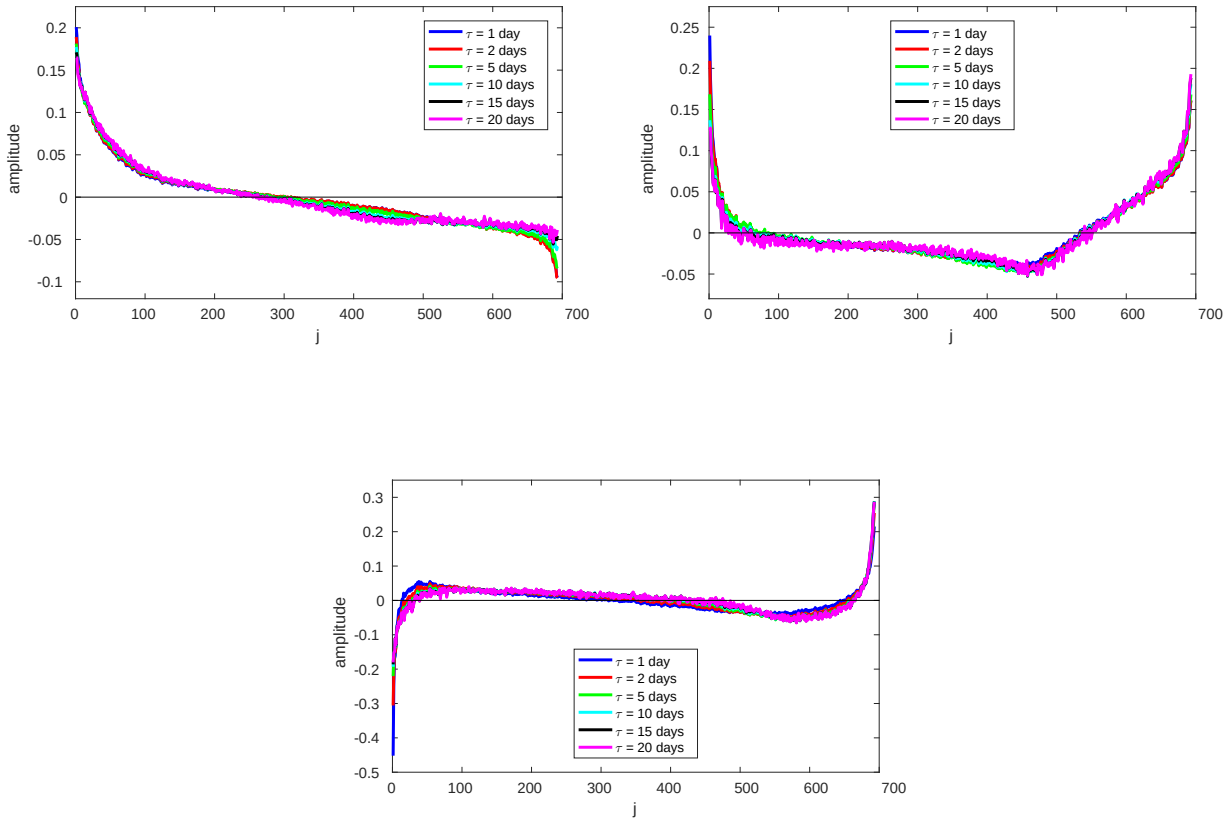

Figure 4. (a) First, (b) second and (c) third empirical orthogonal function of the finite-time Lyapunov fluctuations.

and at the smallest, most stable exponent. The large-deviation rate function is estimated as described in Section 5 from the probability density function and via the Legendre transform for various values of τ. The corresponding element $D_{j,j}$ of the diffusion matrix is calculated from the curvature of the two estimates of the rate function and directly from the time series of the FTLEs.

To model the probability density of the FTLEs two different choices for the potential function in eq.(42) are considered here: $M = 2$, that is, a Gaussian probability density and $M = 4$, a fourth-order polynomial. In view of the high degree of correlation in the time series of the FTLEs, particularly for large τ, model selection is here performed as follows. For $\tau = n\Delta\tau$, the time series of the FTLEs, $\{\Lambda_{j,\alpha}^{(\tau)}\}_{\alpha=0}^{L-1}$, are split into n disjoint subsets with non-overlapping integration time intervals, $\{\Lambda_{j,m}^{(\tau)}, \Lambda_{j,m+n}^{(\tau)}, \Lambda_{j,m+2n}^{(\tau)}, \ldots\}$, for $m = 0, \ldots, n-1$, the two models are fitted separately on the subsets and model selection is based on the average Bayesian information criterion over the subsets. Then the selected model is fitted on the whole time series.

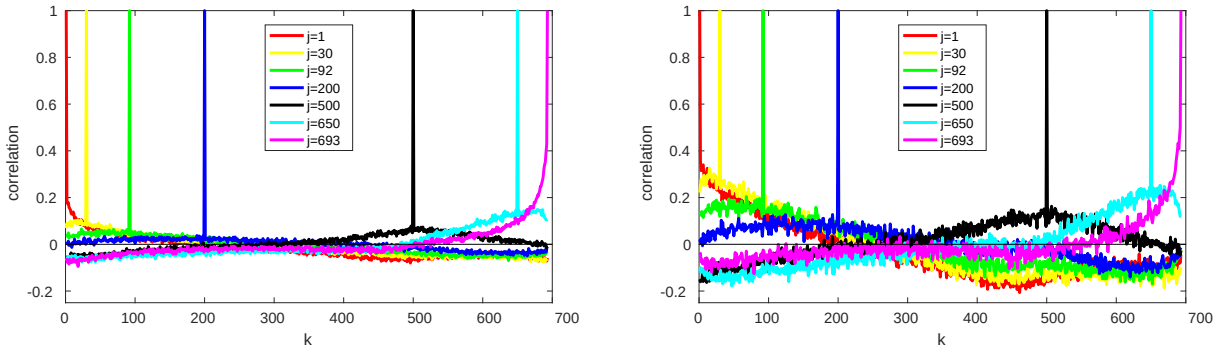

Figure 5. Correlation of the FTLEs $\Lambda_j^{(\tau)}$ and $\Lambda_k^{(\tau)}$ for (a) $\tau = 1$ day and (b) $\tau = 15$ days.

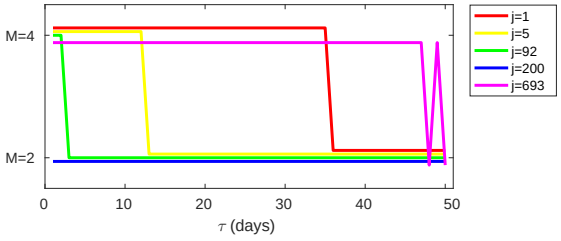

Figure 6. Order of the model for the probability density function of the FTLE $\Lambda_j^{(\tau)}$.

Figure 6 displays the order of the model for the probability density of the selected FTLEs as a function of the integration time τ. The leading unstable exponents exhibit strong non-Gaussianity. For the first exponent, it is detectable up to $\tau = 35$ days; for the fifth exponent, it is less pronounced and visible only up to $\tau = 12$ days. The zero exponent shows only very mild non-Gaussianity which is visible for $\tau = 1$ day and $\tau = 2$ days. The weakly dissipative exponent has Gaussian behaviour at all values of τ. The smallest, strongly dissipative exponent again displays marked deviations from Gaussianity; these are even more pronounced than those for the first exponent and detectable up to an integration time as large as $\tau = 49$ days. For the first and the last exponent, at small integration times τ it may be possible to even switch to the higher-order model $M = 6$ but this is not our concern here.

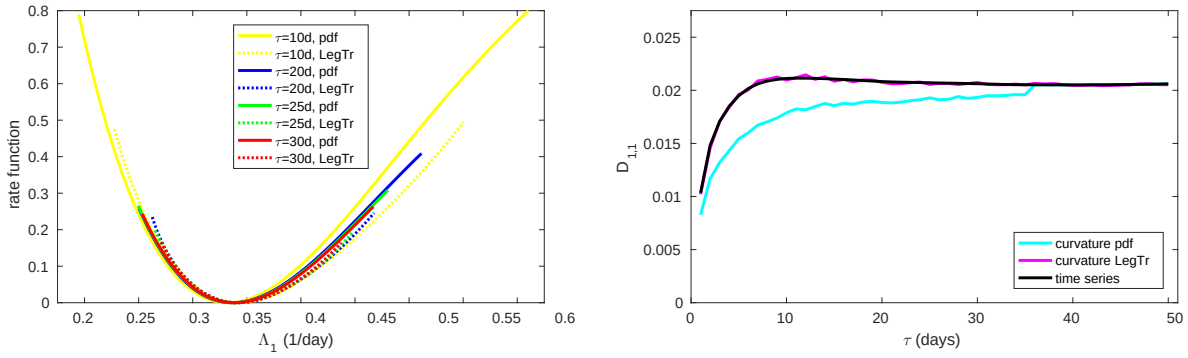

Figure 7. (a) Large-deviation rate function of the first FTLE. (b) Element $D_{1,1}$ of the diffusion matrix.

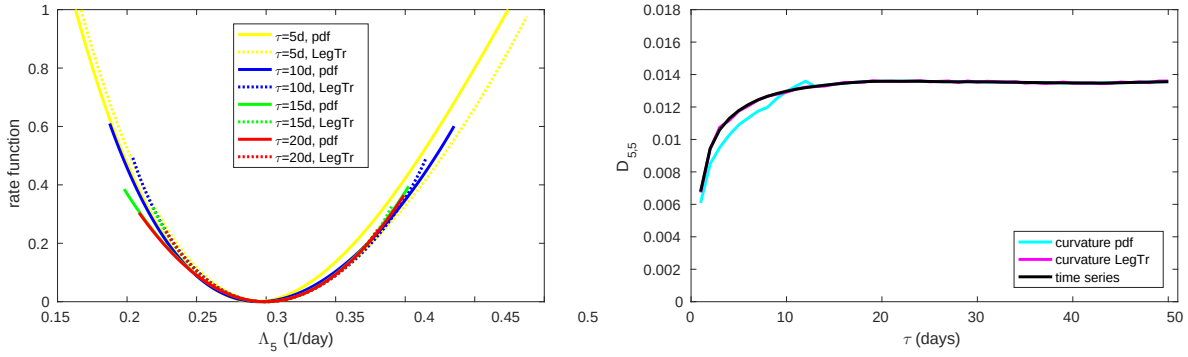

Figure 8. (a) Large-deviation rate function of the fifth FTLE. (b) Element $D_{5,5}$ of the diffusion matrix.

Figure 7 shows the results of the large-deviation analysis for the first FTLE. Convergence to a large-deviation principle is observed. At $\tau = 10$ days and even visible at $\tau = 20$ days the maximum of the probability density is still shifted away from the mean; nevertheless, some convergence among the probability density-based estimates of the rate function is reached at about $\tau = 20$ days. The Legendre transform-based estimates give a consistent picture already from $\tau = 10$ days. Good convergence is also observed for the corresponding element of the diffusion matrix.

For the fifth FTLE, a similar picture can be seen (Figure 8) but convergence is markedly faster than for the first FTLE. The probability density-based estimates are very consistent from $\tau = 10 - 15$ days; note that the model for the probability density

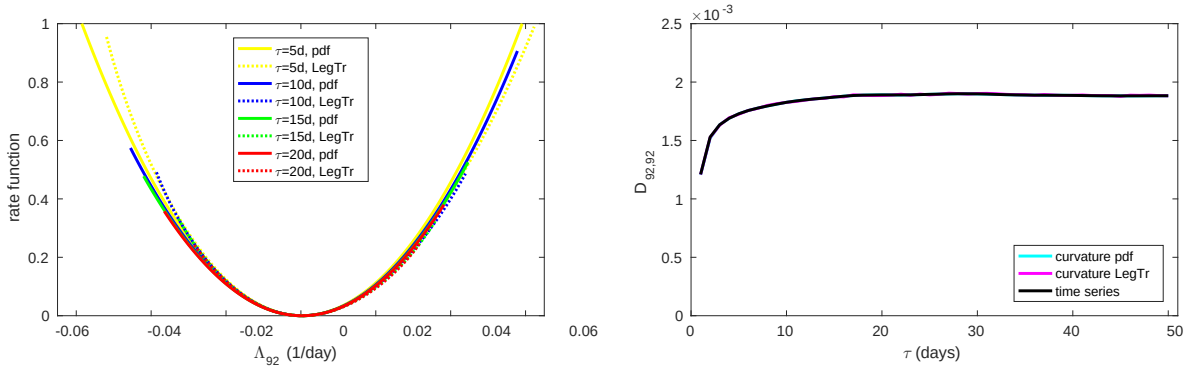

Figure 9. (a) Large-deviation rate function of the 92th FTLE. (b) Element $D_{92,92}$ of the diffusion matrix.

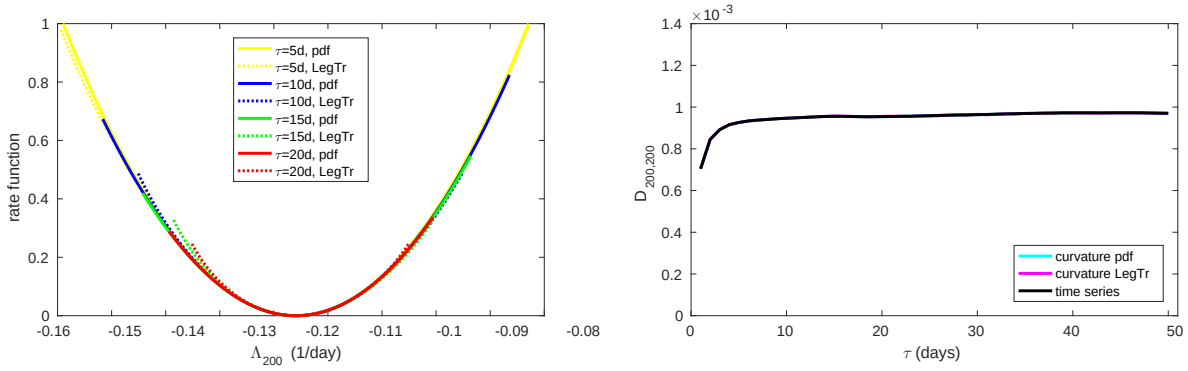

Figure 10. (a) Large-deviation rate function of the 200th FTLE. (b) Element $D_{200,200}$ of the diffusion matrix.

jumps from fourth-order to Gaussian for the higher values of τ. The Legendre transform gives close agreement for the rate function already from $\tau = 5$ days.

For the zero exponent (Figure 9), convergence is again markedly faster than for both positive exponents. A large-deviation principle can be established already from about $\tau = 10$ days and the two different estimates of the rate function are close together. The estimates of the diffusion coefficient all coincide.

For the fully Gaussian 200th FTLE (Figure 10), convergence is even faster. A large-deviation principle is valid from $\tau = 5$ days and all of the estimates of the rate function are in almost perfect agreement. The estimates of the diffusion coefficient show corresponding behaviour.

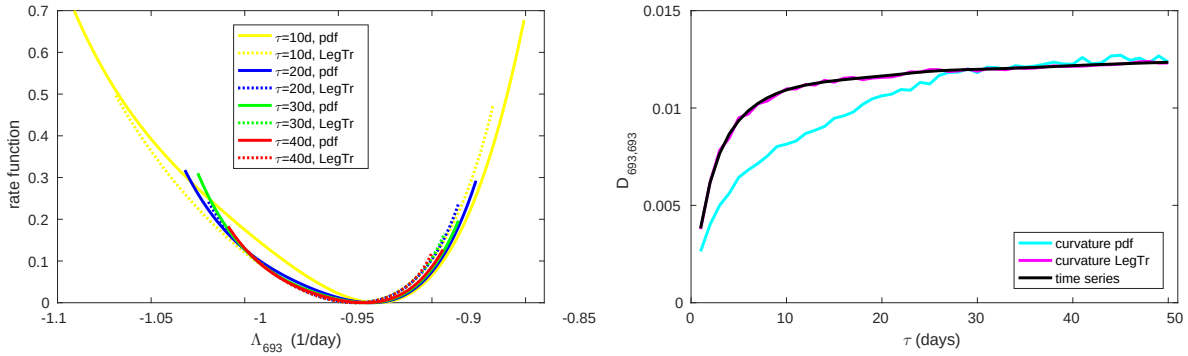

Figure 11. (a) Large-deviation rate function of the 693th FTLE. (b) Element $D_{693,693}$ of the diffusion matrix.

Table 6. Correlation length $l_j^{(\Delta\tau)}$ of the time series of the FTLE $\Lambda_j^{(\Delta\tau)}$ with $\Delta\tau = 1$ day

j	1	5	92	200	693
$l_j^{(\Delta\tau)}$	2.07	2.10	1.61	1.38	3.25

For the smallest, most dissipative exponent (Figure 11), the convergence to a large-deviation principle is very slow, even slower than for the first, most unstable exponent. A large-deviation principle is valid from about $\tau = 30$ days and the Legendre transform method gives reliable estimates of the rate function from $\tau = 10 - 20$ days. Also the convergence of the diffusion coefficient is markedly slow. The estimate from the non-Gaussian probability density is initially to low and converges at about $\tau = 25$ days.

The different speeds of convergence to a large-deviation principle for the different FTLEs can be understood from the degrees of serial correlation and non-Gaussianity of the FTLEs. Table 6 gives the correlation length of the selected FTLEs $\Lambda_j^{(\Delta\tau)}$ at the basic integration time $\Delta\tau = 1$ day defined as

$$l_j^{(\Delta\tau)} = 1 + 2\sum_{i=1}^{\infty} \rho_{j,i}^{(\Delta\tau)} \tag{57}$$

where $\rho_{j,i}^{(\Delta\tau)}$ is the autocorrelation at lag i of the time series of the FTLE $\Lambda_j^{(\Delta\tau)}$. Convergence to a large-deviation principle can certainly not be expected before the serial correlations have decayed. If the distribution is Gaussian convergence occurs immediately thereafter; otherwise it is delayed, generally the longer the larger the departure from Gaussianity is.

6.3.2 Two-dimensional approach

As an example of a multivariate large-deviation analysis, Figure 12 shows the joint large-deviation rate function of the first two FTLEs, Λ_1 and Λ_2. The estimates of the diffusion coefficients $D_{1,1}$, $D_{2,2}$ and $D_{1,2}$ are also shown. The potential function for the joint probability density is chosen as

$$U^{(\tau)}(z_1, z_2) = \sum_{i=1}^{M} \sum_{j=0}^{i} \beta_{i,j}^{(\tau)} \left(\frac{z_1 - \lambda_1}{\sigma_1^{(\tau)}} \right)^{i-j} \left(\frac{z_2 - \lambda_2}{\sigma_2^{(\tau)}} \right)^{j} \tag{58}$$

where the order of the model is fixed a priori at $M = 4$. The joint rate function displays markedly non-Gausian behaviour and some dependence between $\Lambda_1^{(\tau)}$ and $\Lambda_2^{(\tau)}$. Convergence to a large-deviation principle is mainly reached at $\tau = 15$ days as can be seen from the probability density-based estimates of the joint rate function. The estimates from the Legendre transform are in agreement and indicate the joint rate function already at $\tau = 10$ days. The elements of the diffusion matrix are overall well estimated with detailed convergence being somewhat slow in accordance with the univariate analysis for the first FTLE. The estimate of the off-diagonal element $D_{1,2}$ is particularly good.

7 Conclusions

The statistical properties of the fluctuations of FTLEs were investigated in a three-level quasi-geostrophic atmospheric model with realistic mean state and variability. The Lyapunov spectrum of the model has almost 100 positive LEs and displays no clear timescale separation.

A principal component analysis of the fluctuations of the FTLEs around their mean values was performed. The scaled covariance matrix of the fluctuations is converged to the limiting diffusion matrix at about $\tau = 15$ days. There are substantial correlations among the different FTLEs. The first three empirical orthogonal functions are patterns where the leading positive FTLEs fluctuate together in phase. These modes are largely independent of the integration time τ.

A large-deviation principle can be established for all of the FTLEs. The convergence to the large-deviation limit behaviour is slightly slow for the most unstable and the most stable FTLEs and very fast in between. Also a joint large-deviation rate function for the first and the second FTLE was successfully estimated. Good correspondance is found between the curvature of the rate functions at the minimum and the corresponding elements of the diffusion matrix.

Competing interests. The author declares that they have no conflict of interest.

References

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M.: Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: a method for computing all of them. Part 1: Theory, Meccanica, 15, 9–20, 1980.

Eckmann, J., and Ruelle, D.: Ergodic theory of chaos and strange attractors, Reviews of Modern Physics, 57, 617–656, 1985.

Ehrendorfer, M.: The total energy norm in a quasigeostrophic model, Journal of the Atmospheric Sciences, 57, 3443–3451, 2000.

Johnson, P. L. and Meneveau, C.: Large-deviation joint statistics of the finite-time Lyapunov spectrum in isotropic turbulence, Physics of Fluids, 27, 085110, 2015.

Kalnay, E.: Atmospheric Modeling, Data Assimilation, and Predictability, Cambridge University Press, 2003.

Kifer, Y.: Large deviations in dynamical systems and stochastic processes, Transactions of the American Mathematical Society, 321, 505–524, 1990.

Kuptsov, P. V. and Parlitz, U.: Theory and computation of covariant Lyapunov vectors, Journal of Nonlinear Science, 22, 727–762, 2012.

Kuptsov, P. V. and Politi, A.: Large-deviation approach to space-time chaos, Physical Review Letters, 107, 114101, 2011.

Kwasniok, F.: Reduced atmospheric models using dynamically motivated basis functions, Journal of the Atmospheric Sciences, 64, 3452–3474, 2007.

Kwasniok, F.: Maximum likelihood probability density estimation, to be published.

Kwasniok, F.: Data-driven estimation of large-deviation rate functions, to be published.

Laffargue, T., Lam, K.-D. N. T., Kurchan, J., and Tailleur, J.: Large deviations of Lyapunov exponents, Journal of Physics A: Mathematical and Theoretical, 46, 254002, 2013.

Legras, B., and Ghil, M.: Persistent Anomalies, Blocking, and Variations in Atmospheric Predictability, Journal of the Atmospheric Sciences, 42, 433–471, 1985.

Lorenz, E. N.: Deterministic Nonperiodic Flow, Journal of the Atmospheric Sciences, 20, 130–141, 1963.

Marshall, J. and Molteni, F.: Toward a dynamical understanding of planetary-scale flow regimes, Journal of the Atmospheric Sciences, 50, 1792–1818, 1993.

Oseledets, V. I.: A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems, Transactions of the Moscow Mathematical Society, 19, 179–210, 1968.

Ott, E.: Chaos in Dynamical Systems, Cambridge University Press, 2002.

Pazó, D., López, J. M., and Politi, A.: Universal scaling of Lyapunov-exponent fluctuations in space-time chaos, Physical Review E, 87, 062909, 2013.

Pikovsky, A. and Politi, A.: Lyapunov Exponents, Cambridge University Press, 2016.

Roads, J. O.: Predictability in the extended range, Journal of the Atmospheric Sciences, 44, 3495–3527.

Rohwer, C. M., Angeletti, F., and Touchette, H.: Convergence of large-deviation estimators, Physical Review E, 92, 052104, 2015.

Schubert, S., and Lucarini, V.: Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: Analysis of instabilities and feedbacks, Quarterly Journal of the Royal Meteorological Society, 141, 3040–3055, 2015.

Shimada, I. and Nagashima, T.: A Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems, Progress of Theoretical Physics, 61, 1605–1616, 1979.

Touchette, H.: The large deviation approach to statistical mechanics, Physics Reports, 478, 1–69, 2009.

Vannitsem, S.: Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics, Chaos, 27, 32101, 2017.

Vannitsem, S., and Lucarini, V.: Statistical and Dynamical Properties of Covariant Lyapunov Vectors in a Coupled Atmosphere-Ocean Model – Multiscale Effects, Geometric Degeneracy, and Error Dynamics, Journal of Physics A: Mathematical and Theoretical, 49, 224001, 2016.

Vannitsem, S., and Nicolis, C.: Lyapunov Vectors and Error Growth Patterns in a T21L3 Quasigeostrophic Model, Journal of the Atmospheric Sciences, 54, 347–361, 1997.

Yang, H.-L., Takeuchi, K. A., Ginelli, F., Chaté, H., and Radons, G.: Hyperbolicity and the Effective Dimension of Spatially Extended Dissipative Systems, Physical Review Letters, 102, 074102, 2009.

Responses to Reviewers' Comments

The author would like to thank both reviewers very much for their useful comments which helped improve the manuscript.

Reviewer 1

- Section 2

  Q:

  In the model, if I correctly understand, the "h" term is only a parameter (not a function). If so, I think the model is invariant under a zonal shift. This neutral transformation would imply a second vanishing Lyapunov exponent in the spectrum. Please, clarify this point.

  A:

  This is a misunderstanding; the model has no zonal symmetry. The topography $h = h(\lambda, \mu)$ is a function of space and is actually the real topography of the earth expanded into spherical harmonics. Also the diabatic source terms $S_i = S_i(\lambda, \mu)$ are functions of space, fitted from reanalysis data. So any symmetry is broken in the model and this is actually crucial for getting a realistic mean state and variability. This is clarified in the revised manuscript.

- Section 6.1

  Q:

  Concerning the last sentence in Sec. 6.1. I note that the equivalence between Lyapunov exponent fluctuations measured from Gram-Schmidt vectors and from covariant vectors, was detected already in Figs. 5 and 7 of Ref. [1].

  A:

  This is acknowledged in the revised manuscript.

  Q:

  In fact, the large fluctuations observed at the edges of the spectrum are not really surprising, at the light of the previous results on the diffusion coefficients in (Kuptsov and Politi, 2011) and [1].

  Q:

  This is referenced in the revised manuscript.

- Section 6.2

  Q:

  It is absolutely necessary to include one formula defining the fraction of explained variance, in order to ensure the self-consistency of the text.

  A:

  Done.

– Section 6.3

Q:

It is not said which is the total length of the time series used.

A:

The length of the time series is 25000 days; this information is now given in the revised manuscript.

Q:

The value of τ_r is "hidden" in Sec. 3.

A:

The value of τ_r is now again given in the results section.

Q:

The three methods used to measure $D_{j,j}$ are not fully clear to me. I think the author should make a list with the three methods specifying which formulas are used in each one. And which parameters are used. Now the explanation is hidden in the caption of Fig. 5, and is hardly understandable.

After reading it several times ...

A:

This part of the manuscript has been substantially revised. The different methods for estimating the rate function and the elements of the diffusion matrix are now explained more clearly and in more detail (see revised manuscript).

– Minor comments:

  – Q:

    Two lines after Eq. (13), I would write scalar product instead of norm, because this is what matters for Gram-Schmidt orthogonalisation. I suppose the energy norm is trivially related to the scalar product used.

    A:

    Corrected.

  – Q:

    Vectors should be always typed in bold face, also for Greek letters.

    A:

    Corrected.

  – Q:

    Orthogonality of the eigenvectors, Eq. (18), is better written after Eq. (16).

    A:

    Done.

  – Q:

    The equation in the text preceding Eq. (18) is apparently lacking of $-\lambda$.

A:

Corrected.

– Q:

When introducing Eq. (27), it would be important to cite at least (Touchette, 2009) again and to mention this is the Gärtner–Ellis theorem (if I'm not wrong).

A:

Done.

– Q:

The "log" symbol is missing in Eq. (35).

A:

Corrected.

– Q:

Figures 4–7 should be introduced in the text, one by one.

A:

Done.

– Q:

Last line of page 10. τ_c has not been defined.

A:

Corrected.

– Q:

Page 12, line 5. I don't appreciate smaller deviations of the rate function in this case than for the zero exponent.

A:

For $\tau = 1$ day and $\tau = 2$ days, there is some non-Gaussianity visible for the zero exponent but not for the 200th exponent as is shown in Figure 6 in the revised manuscript.

– Q:

In figures 5–8, I would use different colours for the lines in panel (b), at least for the coloured ones since they are not related to the same colours in panel (a).

A:

Done.

– Q:

In the conclusions, it is mentioned that the most unstable exponents exhibit slower convergence to the large-deviation limit. Let me to point out that this is fully consistent with [1].

A:

This is referenced in the revised manuscript.

– Q:

Labels (a), (b), etc. need to be included all the figures. This is critical in Fig. 9.

A:

Done.

Reviewer 2

– On section 1:

Q:

Line 8: based from what I understand: Could the covariance structure tell us something about how "close" or how "interactive" the various unstable and stable directions are? Could the covariance structure be related to the investigations of the inertial manifold using the angles between Lyapunov vectors. Maybe the work along the lines of Yang et al (2009) should be referenced here as a motivation.

A:

This is an interesting point but would clearly need the use of the covariant Lyapunov vectors. It is mentioned in the revised manuscript as a possible future research line and the work by Yang et al (2009) is referenced.

– On Section 2:

Q:

I think a short concise table listing all parameters of the model with their dimensional and a dimensional values would be beneficial to introduce the model setup.

A:

Tables listing all the variables and parameters of the model have been included in the revised manuscript.

– On Section 3:

Q:

It should be noted that the mean of the finite time Lyapunov exponents are in fact average growth rates of linear perturbations of the system. But the finite time LEs are not directly the growth rate of those perturbation. In fact one can define backward, forward and covariant LEs. There is a good review paper on this by Kuptsov and Parlitz which explains this distinction. I think using the FTLEs of the Gram Schmidt algorithm is alright, but it should be better clarified what type of FTLEs they actually are.

A:

Some comments on this have been added to the manuscript and the paper by Kuptsov and Parlitz is cited. In the limit of large integration time τ, which is the focus of the present paper, the three types of FTLEs are actually equivalent and the backward FTLEs are easiest to calculate.

- On Section 4:

  Q:

  I think this section should motivate better why one should use EOFs and what would be potentially alternatives to this approach.

  A:

  A couple of comments have been added here (see revised manuscript).

- On Section 6:

  - Section 6.1:

    Q:

    Since the model is zonally symmetric there should be two zero exponents. Can you verify this and could you include this in this discussion?

    A:

    This is a misunderstanding; the model has no zonal symmetry. The topography $h = h(\lambda, \mu)$ is a function of space and is actually the real topography of the earth expanded into spherical harmonics. Also the diabatic source terms $S_i = S_i(\lambda, \mu)$ are functions of space, fitted from reanalysis data. So any symmetry is broken in the model and this is actually crucial for getting a realistic mean state and variability. This is clarified in the revised manuscript.

    Q:

    Figure 2: This result should be referenced with the findings about fluctuations of the LE for covariant, backward and forward exponents in Vannitsem, Lucarini (2016). I think when you study collective excitations this is an interesting different viewpoint.

    A:

    Done.

  - Section 6.2:

    Q:

    I think it would be helpful to present the matrix D as well as a surface plot and also use the first EOF and second EOF to see what parts of the D matrix are actually reconstructed using the EOF method.

    A:

    I tried this but it turned out not to be really useful as the matrix D is strongly dominated by the diagonal and the leading EOFs do not explain that much variance, even in the diffusion limit. I have added a plot showing the correlation of selected FTLEs with all the other exponents (see Figure 5 in the revised manuscript).

- Section 7:

  Q:

  That no clear time scale separation is found is probably because QG equations are scale filtered quations. Similar results

were found before in QG models (Vannitsem 1997, Schubert 2015, Vannitsem 2016).

A:

I agree and a comment along this line has been added.

Q:

Can the collective excitations be traced to any anomalous behaviour in the nonlinear background state x? I think that would an interesting addition but of course not necessary in order to do this study.

A:

A separate study is actually underway by the author linking FTLEs to underlying weather regimes in the model state.

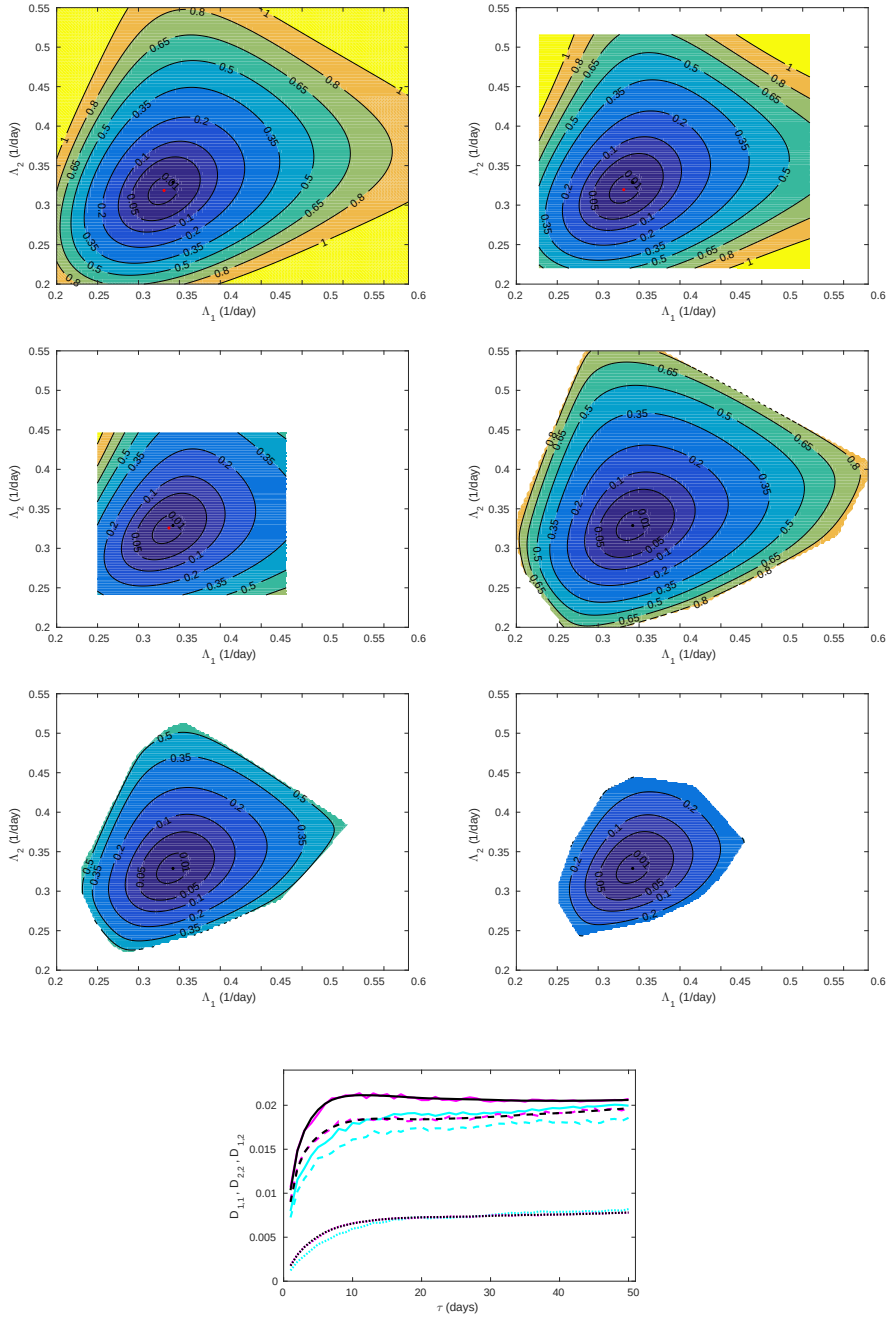

Figure 12. Joint large-deviation rate function of the first two FTLEs as estimated from the joint probability density for (a) $\tau = 10$ days, (b) $\tau = 15$ days and (c) $\tau = 25$ days; and with the Legendre transform for (d) $\tau = 10$ days, (e) $\tau = 15$ days and (f) $\tau = 25$ days. Black dots indicate the global LEs (λ_1, λ_2); red dots in the panels (a), (b) and (c) indicate the maximum of the joint probability density. (g) Elements $D_{1,1}$ (solid), $D_{2,2}$ (dashed) and $D_{1,2}$ (dotted) of the diffusion matrix as estimated from the curvature of the probability density-based rate function (cyan), from the curvature of the Legendre transform-based rate function (magenta) and from the time series of the FTLEs (black).