Nonlinear Processes
in Geophysics

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, UK

Correspondence: Frank Kwasniok (f.kwasniok@exeter.ac.uk)

Abstract. The stability properties as characterized 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 becomes 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 non-linear chaotic dynamical system; its time evolution is characterized 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, it 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 non-linear 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 characterization 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) and coupled atmosphere–ocean models (Vannitsem and Lucarini, 2016). A review has been published recently by Vannitsem (2017). Models tuned to realistic conditions were found 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 that have found little attention in the context of geophysical fluid systems thus far. 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 (Kuptsov and Politi, 2011). Secondly, we look at the large-deviation behaviour of the FTLEs (Kuptsov and Politi, 2011; Laffargue et al., 2013; Johnson and Meneveau, 2015). A large-deviation principle links the FTLEs at long integration times to the global LEs by providing a universal law for the probability density of fluctuations of the FTLEs around the mean value. It can be expected to hold for Axiom A dynamical systems and, invoking the chaotic hypothesis, also for certain types of non-Axiom A systems. In particular, a large-deviation law allows one to determine the probability of very stable or very unstable atmospheric states.

The paper is organized as follows: in Sect. 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 Sects. 3, 4 and 5; the results are presented and discussed in Sect. 6; and some conclusions are drawn in Sect. 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 as follows:

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

where Ψ_i and q_i are the stream function and the potential vorticity at level i, respectively, and J denotes the Jacobian operator on the sphere. All variables are non-dimensional 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 stream function 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 and 375 km, respectively. The function $h = h(\lambda, \mu)$ represents a non-dimensional 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 follows:

$$D_1 = \tau_N^{-1} R_{1,2}^{-2}(\Psi_1 - \Psi_2) - k_H \nabla^8 \hat{q}_1 \tag{5}$$

$$D_2 = -\tau_N^{-1} R_{1,2}^{-2}(\Psi_1 - \Psi_2) + \tau_N^{-1} R_{2,3}^{-2}(\Psi_2 - \Psi_3) - k_H \nabla^8 \hat{q}_2 \tag{6}$$

$$D_3 = -\tau_N^{-1} R_{2,3}^{-2}(\Psi_2 - \Psi_3) - \tau_E^{-1} \nabla^2 \Psi_3 - k_H \nabla^8 \hat{q}_3. \tag{7}$$

They are Newtonian temperature relaxation with a radiative timescale of $\tau_N = 25$ d, Ekman damping on the lowest level with a spin-down timescale of $\tau_E = 1.5$ d, 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 to say $\hat{q}_i = q_i - f - \delta_{i3} f_0 h$. The coefficient of horizontal diffusion $k_H = \tau_H^{-1}[n_m(n_m+1)]^{-4}$ is such that harmonics of total wave-number $n_m = 21$ are damped at a timescale of $\tau_H = 1.5$ d. 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 to say vanishing stream function, $\Psi_i(\lambda, 0) = 0$, is applied at the Equator on all three model levels. The horizontal discretization is spectral, triangularly truncated at total wave-number $n_m = 21$. 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 non-dimensional values.

In order to get a model behaviour close to that of the real atmosphere, the forcing terms S_i are determined from the European Centre for Medium-Range Weather Forecasts (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); this is done 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 a remarkably realistic mean state and variability pattern of stream function and potential vorticity in a long-term integration (see Table 3).

3 Lyapunov exponents

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

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

Symbol	Description	Unit	Non-dimensionalization
t	Time	s	Ω^{-1}
$\lambda \quad (0 \leq \lambda < 2\pi)$	Longitude (eastward)		
$\phi \quad (0 \leq \phi \leq \pi/2)$	Latitude		
$\mu = \sin\phi \quad (0 \leq \mu \leq 1)$	Sine of latitude		
Ψ_i	Stream function at level i	m^2 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)	m s^{-1}	$a\Omega$
$v_i = (1/\sqrt{1-\mu^2})\,(\partial\Psi_i/\partial\lambda)$	Meridional velocity (northward)	m s^{-1}	$a\Omega$

Table 2. Parameters in the QG model.

Symbol	Description	Dimensional value	Non-dimensional 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 d	50π
τ_E	Timescale of Ekman damping	1.5 d	3π
τ_H	Timescale of horizontal diffusion at wave-number 21	1.5 d	3π
f_0	Coriolis parameter at 45° N	1.031×10^{-4} s^{-1}	$\sqrt{2}$
H	Scale height	8 km	

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

tion equations

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}). \tag{8}$$

The linearized dynamics of a small perturbation $\delta\boldsymbol{x}$ are given as

$$\frac{d}{dt}\delta\boldsymbol{x} = \frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}}\delta\boldsymbol{x}. \tag{9}$$

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

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

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

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

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

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

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 the norm.

In order to characterize perturbation growth or decay over a finite integration time τ the FTLEs $\Lambda_j^{(\tau)}(\boldsymbol{x}_0)$ 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). Here, we refer to the backward FTLEs as they are easiest to compute. They are calculated using the standard algorithm based on the Gram–Schmidt orthogonalization (Shimada and Nagashima, 1979;

Benettin et al., 1980). An ensemble of N linearly independent perturbations is initialized and integrated forward in time along with the non-linear model trajectory. A transient period is discarded for the trajectory to settle on the attractor of the system and for the perturbations to converge to the backward Lyapunov vectors. Then, after every integration time interval $\Delta\tau$ the perturbations are re-orthonormalized using a QR-decomposition performed via the Gram–Schmidt procedure. The FTLEs are obtained as

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

where $R_{jj}(t_\alpha, t_{\alpha+1})$ are the diagonal elements of the upper 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 $\boldsymbol{x}_\alpha = \boldsymbol{x}(t_\alpha)$. The FTLEs $\Lambda_j^{(\tau)}$ for larger integration times $\tau = n\Delta\tau$ are obtained by averaging over n consecutive values of $\Lambda_j^{(\Delta\tau)}$:

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

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

The FTLEs depend on the scalar product chosen in the Gram–Schmidt orthogonalization procedure. Here, we 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

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

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

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

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

$$\mathbf{D}^{(\tau)} = \left\langle\left(\boldsymbol{\Lambda}^{(\tau)} - \boldsymbol{\lambda}\right)\left(\boldsymbol{\Lambda}^{(\tau)} - \boldsymbol{\lambda}\right)^{\mathrm{T}}\right\rangle\tau$$
$$= \frac{\tau}{L}\sum_{\alpha=0}^{L-1}\left(\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda}\right)\left(\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda}\right)^{\mathrm{T}}. \tag{17}$$

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

$$\mathbf{D}^{(\tau)}\boldsymbol{e}_j^{(\tau)} = \nu_j^{(\tau)}\boldsymbol{e}_j^{(\tau)}. \tag{19}$$

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

$$\boldsymbol{e}_j^{(\tau)} \cdot \boldsymbol{e}_k^{(\tau)} = \sum_{i=1}^N e_{j,i}^{(\tau)} e_{k,i}^{(\tau)} = \delta_{jk}. \tag{20}$$

The fluctuation field of the FTLEs is expanded as

$$\boldsymbol{\Lambda}_\alpha^{(\tau)} - \boldsymbol{\lambda} = \sum_{j=1}^N y_{j,\alpha}^{(\tau)}\boldsymbol{e}_j^{(\tau)} \tag{21}$$

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

$$\left\langle y_j^{(\tau)} y_k^{(\tau)}\right\rangle = \frac{1}{L}\sum_{\alpha=0}^{L-1} y_{j,\alpha}^{(\tau)} y_{k,\alpha}^{(\tau)} = \nu_j^{(\tau)}\delta_{jk}. \tag{22}$$

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

$$r_j^{(\tau)} = \frac{\nu_j^{(\tau)}}{\sum_{k=1}^N \nu_k^{(\tau)}} \tag{23}$$

and the cumulative fraction of variance given as

$$c_j^{(\tau)} = \frac{\sum_{k=1}^j \nu_k^{(\tau)}}{\sum_{k=1}^N \nu_k^{(\tau)}}. \tag{24}$$

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). In the following, large-deviation theory is 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

$$A_n = \frac{1}{n} \sum_{i=1}^n X_i \tag{25}$$

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

$$\gamma(\theta) = \lim_{n \to \infty} \frac{1}{n} \log \left\langle e^{n\theta A_n} \right\rangle \tag{26}$$

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

$$p(A_n = z) \sim \exp[-nI(z)], \tag{27}$$

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

$$I(z) = \sup_{\theta \in \mathbb{R}} [\theta z - \gamma(\theta)]. \tag{28}$$

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 to say $I(\langle X \rangle) = 0$ and $I'(\langle X \rangle) = 0$. The curvature of the rate function at the minimum is given as (Touchette, 2009)

$$I''(\langle X \rangle) = \frac{1}{\lim_{n \to \infty} n \langle (A_n - \langle A_n \rangle)^2 \rangle}. \tag{29}$$

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,

$$p\left(\Lambda_j^{(\tau)} = z\right) \sim \exp[-\tau I_j(z)], \tag{30}$$

where the large-deviation rate function $I_j(z)$ is independent of τ and given as

$$I_j(z) = \sup_{\theta \in \mathbb{R}} [\theta z - \gamma_j(\theta)] \tag{31}$$

with the SCGF

$$\gamma_j(\theta) = \lim_{\tau \to \infty} \frac{1}{\tau} \log \left\langle e^{\tau \theta \Lambda_j^{(\tau)}} \right\rangle. \tag{32}$$

Introducing $\theta' = \tau \theta$ and then dropping the prime again we get

$$I_j(z) = \lim_{\tau \to \infty} \frac{1}{\tau} \sup_{\theta \in \mathbb{R}} \left[\theta z - \log \left\langle e^{\theta \Lambda_j^{(\tau)}} \right\rangle \right]. \tag{33}$$

We expect convergence of the rate function $I_j(z)$ as soon as the integration time τ is large enough for consecutive values of $\Lambda_j^{(\tau)}$ taken over non-overlapping integration time intervals, $\Lambda_{j,\alpha}^{(\tau)}$ and $\Lambda_{j,\alpha+n}^{(\tau)}$, to be independent. This is actually an application of the block averaging method (Rohwer et al., 2015). Note, however, that convergence of the rate function at a particular value of τ here does not guarantee that the probability density function is already in the large-deviation limit at that value of τ.

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

$$I_j''(\lambda_j) = \frac{1}{\lim_{\tau \to \infty} \left\langle \left(\Lambda_j^{(\tau)} - \lambda_j \right)^2 \right\rangle \tau} = D_{j,j}^{-1}. \tag{34}$$

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

$$I_j(z) \approx \frac{1}{2} I_j''(\lambda_j)(z - \lambda_j)^2, \tag{35}$$

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. 30) or via the Legendre transform (cf., Eq. 33).

5.2.1 Probability density function approach

By inverting Eq. (30) we have

$$I_j(z) = -\lim_{\tau \to \infty} \frac{1}{\tau} \log p\left(\Lambda_j^{(\tau)} = z\right). \tag{36}$$

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

$$p\left(\Lambda_j^{(\tau)} = z\right) = \frac{1}{Z_j^{(\tau)}} \exp\left[-U_j^{(\tau)}(z) \right] \tag{37}$$

with normalization constant

$$Z_j^{(\tau)} = \int\limits_{-\infty}^{\infty} \exp\left[-U_j^{(\tau)}(z)\right] dz. \tag{38}$$

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

$$U_j^{(\tau)}(z) = \sum_{i=1}^{M} \beta_i^{(\tau)} \left(\frac{z - \lambda_j}{\sigma_j^{(\tau)}}\right)^i. \tag{39}$$

Here $\sigma_j^{(\tau)}$ is the standard deviation of the FTLE $\Lambda_j^{(\tau)}$:

$$\sigma_j^{(\tau)} = \left\langle \left(\Lambda_j^{(\tau)} - \lambda_j\right)^2 \right\rangle^{1/2} = \left[\frac{1}{L}\sum_{\alpha=0}^{L-1}\left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j\right)^2\right]^{1/2}. \tag{40}$$

The parameter M determines the complexity of the model. In order to have a normalizable probability density, we need M to be even and $\beta_M^{(\tau)} > 0$. The expansion coefficients $\{\beta_i^{(\tau)}\}_{i=1}^{M}$ are determined by maximizing the likelihood function of the data $\{\Lambda_{j,\alpha}^{(\tau)}\}_{\alpha=0}^{L-1}$. This is a convex optimization problem with a unique maximum which is numerically stable to solve. Model selection is performed with the Bayesian information criterion.

The estimate of the rate function is given as

$$I_j(z) = \lim_{\tau\to\infty} \frac{1}{\tau}\left[U_j^{(\tau)}(z) - U_j^{(\tau)}(z^*)\right], \tag{41}$$

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; however, 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 normalization 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. (33) (Rohwer et al., 2015) with the moment generating function estimated by the sample mean over the time series:

$$\left\langle e^{\theta \Lambda_j^{(\tau)}} \right\rangle = \frac{1}{L}\sum_{\alpha=0}^{L-1} e^{\theta \Lambda_{j,\alpha}^{(\tau)}}. \tag{42}$$

For each z, this is a convex optimization problem with a unique solution, if a solution 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 $\tau = n\Delta\tau$ is large enough for successive values of $\Lambda_j^{(\tau)}$ over non-overlapping integration time intervals, $\Lambda_{j,\alpha}^{(\tau)}$ and $\Lambda_{j,\alpha+n}^{(\tau)}$, to be independent. However, this gives no indication of whether or not the probability density function is actually already in the large-deviation limit. Therefore, here we 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. 34). They can also be estimated directly from the time series of the FTLEs according to Eq. (17). 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. (39), 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 $\mathbf{\Lambda}^{(\tau)}$ now denote the column vector of any K-dimensional subset of the N FTLEs and $\boldsymbol{\lambda}$ the corresponding vector of global LEs. We have $1 \le K \le N$, 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,

$$p\left(\mathbf{\Lambda}^{(\tau)} = z\right) \sim \exp[-\tau I(z)], \tag{43}$$

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

$$I(z) = \lim_{\tau\to\infty} \frac{1}{\tau} \sup_{\boldsymbol{\theta}\in\mathbb{R}^K}\left[\boldsymbol{\theta}^{\mathrm{T}} z - \log\left\langle e^{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{\Lambda}^{(\tau)}}\right\rangle\right]. \tag{44}$$

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

$$\left.\frac{\partial^2 I}{\partial z_j \partial z_k}\right|_{z=\boldsymbol{\lambda}} = Q_{j,k} = (\mathbf{D}^{-1})_{j,k}, \tag{45}$$

where $\mathbf{D}$ 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 $\boldsymbol{\lambda}$,

$$I(z) \approx \frac{1}{2}(z - \boldsymbol{\lambda})^{\mathrm{T}} \mathbf{Q}(z - \boldsymbol{\lambda}), \tag{46}$$

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. 43) or via the Legendre transform (cf., Eq. 44).

5.5.1 Probability density function approach

By inverting Eq. (43) we get

$$I(z) = -\lim_{\tau \to \infty} \frac{1}{\tau} \log p\left(\boldsymbol{\Lambda}^{(\tau)} = z\right). \tag{47}$$

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

$$p\left(\boldsymbol{\Lambda}^{(\tau)} = z\right) = \frac{1}{Z^{(\tau)}} \exp\left[-U^{(\tau)}(z)\right] \tag{48}$$

with normalization constant

$$Z^{(\tau)} = \int_{\mathbb{R}^K} \exp\left[-U^{(\tau)}(z)\right] d^K z. \tag{49}$$

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

$$U^{(\tau)}(z) = \sum_{i=1}^{J} \beta_i^{(\tau)} \phi_i(z) \tag{50}$$

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

The estimate of the joint rate function is

$$I(z) = \lim_{\tau \to \infty} \frac{1}{\tau} \left[U^{(\tau)}(z) - U^{(\tau)}(z^*) \right], \tag{51}$$

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

5.5.2 Legendre transform approach

Alternatively, the joint rate function $I(z)$ can be determined via the multivariate Legendre transform of Eq. (44) with the

moment generating function estimated as the sample mean over the time series:

$$\left\langle e^{\boldsymbol{\theta}^{\mathrm{T}} \boldsymbol{\Lambda}^{(\tau)}} \right\rangle = \frac{1}{L} \sum_{\alpha=0}^{L-1} e^{\boldsymbol{\theta}^{\mathrm{T}} \boldsymbol{\Lambda}_\alpha^{(\tau)}}. \tag{52}$$

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

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. 45). It can also be estimated directly from the time series of the FTLEs as given in Eq. (17). 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. (50) 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 summarized in Tables 4 and 5.

In high-dimensional systems it is usually too ambitious a task to determine $I(z)$ beyond the Gaussian approximation for the full system. Here, we restrict ourselves to the bivariate case $K = 2$.

6 Results

6.1 Lyapunov exponents

Time series of FTLEs of the QG model of length $L = 25\,000$ with a basic integration time $\Delta \tau = 1\,\mathrm{d}$ are generated as described in Sect. 3. The (global) LEs are calculated as

$$\lambda_j = \frac{1}{L} \sum_{\alpha=0}^{L-1} \Lambda_{j,\alpha}^{(\Delta \tau)}. \tag{53}$$

Figure 1 displays the Lyapunov spectrum of the QG model. There are 91 positive LEs. The largest LE is estimated as $\lambda_1 = 0.342\,\mathrm{d}^{-1}$, corresponding to an e-folding time of perturbation growth of $2.9\,\mathrm{d}$ 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 and $-0.05\,\mathrm{d}^{-1}$. 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

Table 4. Methods for estimating the rate function.

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

Table 5. Methods for estimating the diffusion matrix.

	Probability density	Legendre transform	Time series
Univariate	$D_{j,j}^{-1} = I_j''(z^*)$	$D_{j,j}^{-1} = I_j''(\lambda_j)$	$D_{j,j} = \lim\limits_{\tau \to \infty} \frac{\tau}{L} \sum_{\alpha=0}^{L-1} \left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j \right)^2$
Multivariate	$(\mathbf{D}^{-1})_{j,k} = \left. \frac{\partial^2 I}{\partial z_j \partial z_k} \right|_{z=z^*}$	$(\mathbf{D}^{-1})_{j,k} = \left. \frac{\partial^2 I}{\partial z_j \partial z_k} \right|_{z=\lambda}$	$D_{j,k} = \lim\limits_{\tau \to \infty} \frac{\tau}{L} \sum_{\alpha=0}^{L-1} \left(\Lambda_{j,\alpha}^{(\tau)} - \lambda_j \right) \left(\Lambda_{k,\alpha}^{(\tau)} - \lambda_k \right)$
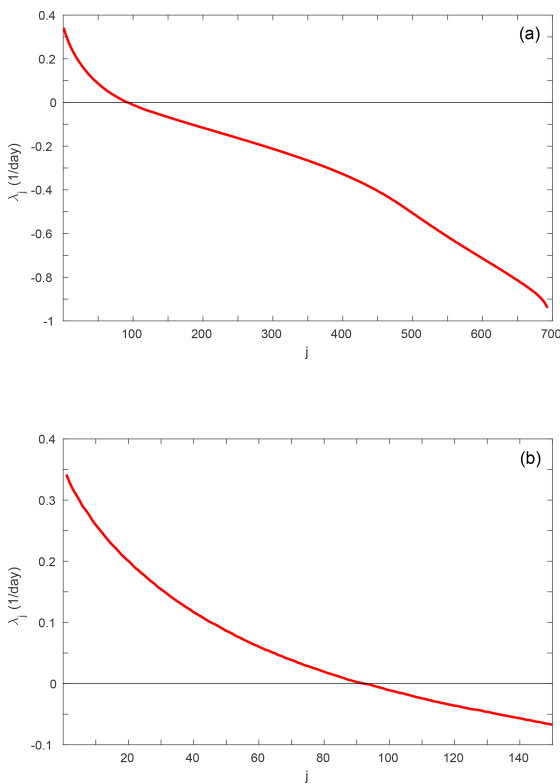
Figure 1. (a) Lyapunov spectrum of the QG model. **(b)** Close-up of **(a)**.

and Lucarini, 2015) and is probably because QG equations are scale-filtered equations.

Figure 2 shows the standard deviation $\sigma_j^{(\tau)}$ of the fluctuations of the FTLEs around their mean values (Eq. 40). 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 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 d, that is to say the scaled variance converges to the diagonal elements $D_{j,j}$ of the diffusion matrix $\mathbf{D}$. Convergence is reached at about $\tau = 10$ d for almost all of the exponents; it is particularly fast for the near-neutral and the weakly dissipative exponents where it is already reached at $\tau = 5$–10 d.

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 "isolated modes" (Yang et al., 2009). But this certainly needs further investigation.

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 $\tau = 1$ to $\tau = 20$ d, 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 $\tau = 20$ d. It is
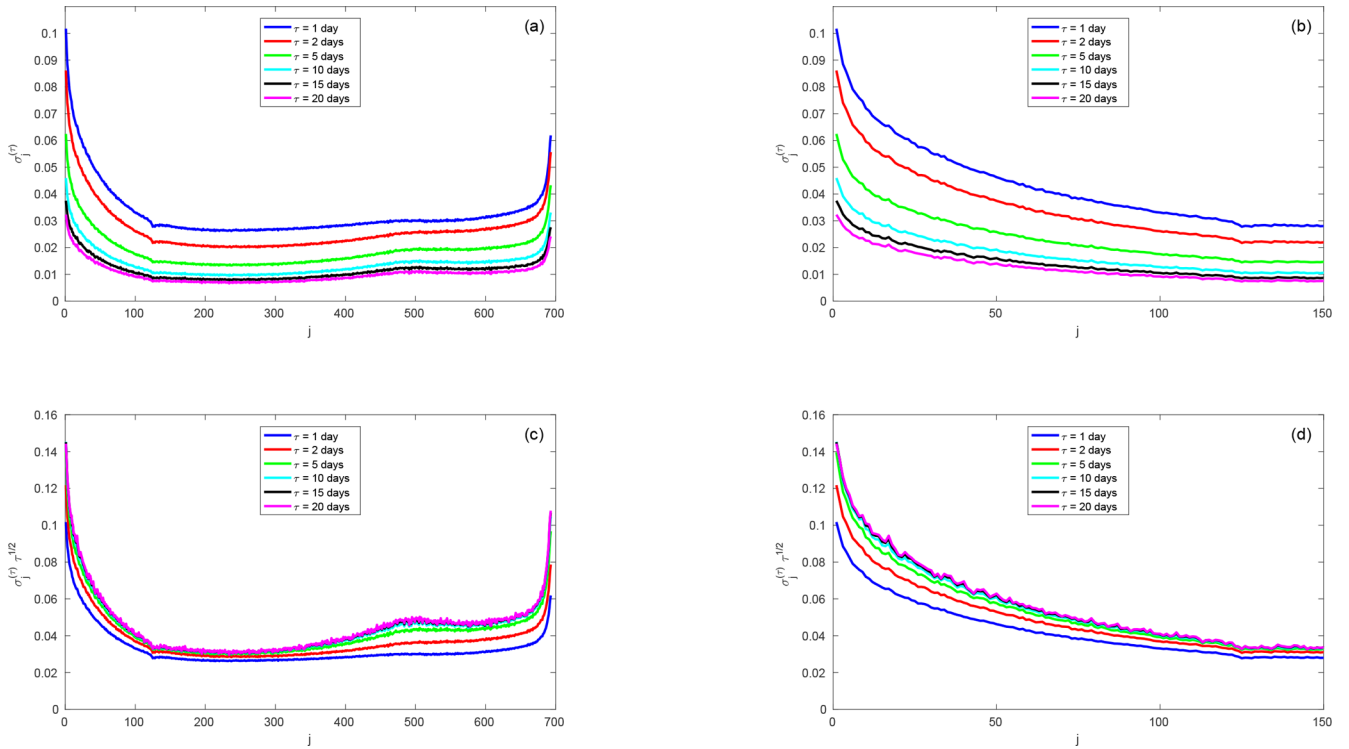
Figure 2. (a) Standard deviation $\sigma_j^{(\tau)}$ of the FTLEs. **(b)** Close-up of **(a)**. **(c)** Scaled standard deviation $\sigma_j^{(\tau)} \tau^{1/2}$ of the FTLEs. **(d)** Close-up of **(c)**.
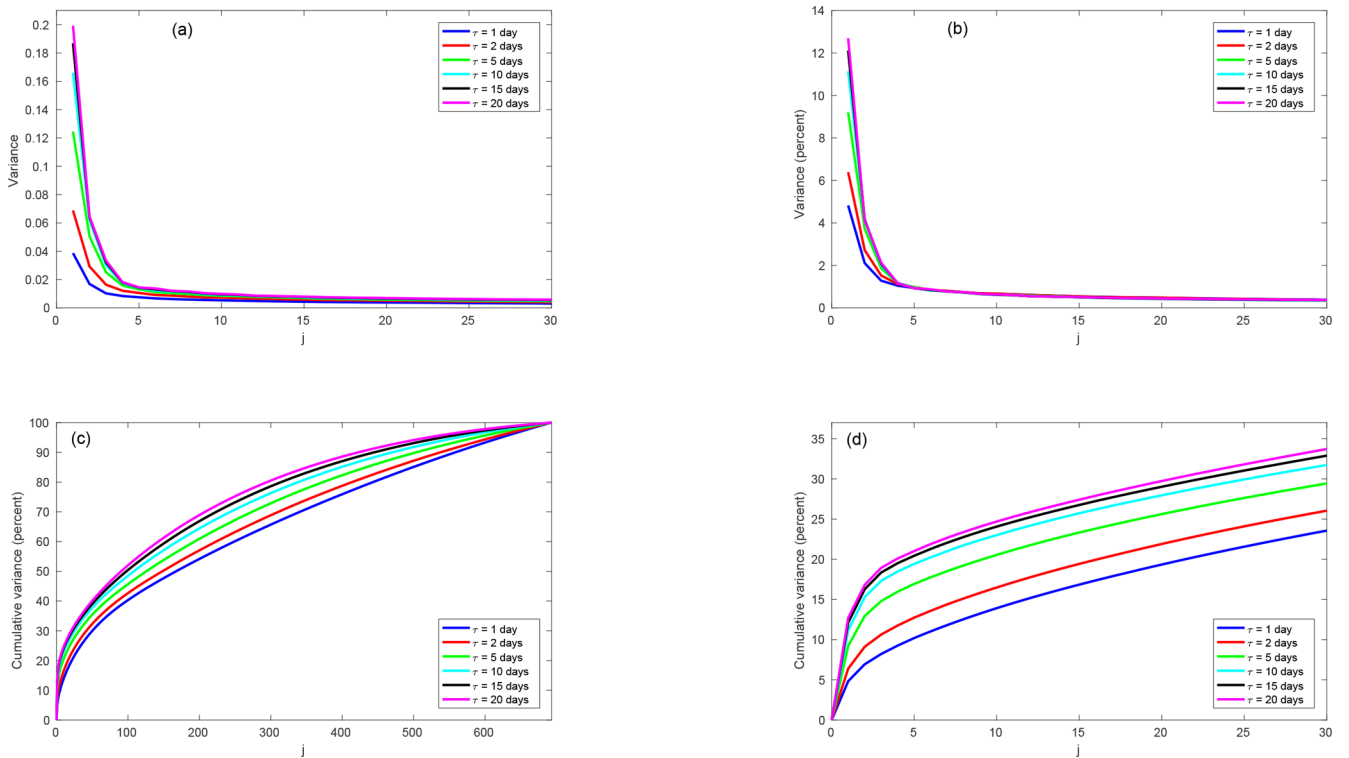

Figure 3. (a) Variance of the principal components of the finite-time Lyapunov fluctuations. **(b)** Fraction of variance. **(c)** Cumulative fraction of variance. **(d)** Close-up of **(c)**.
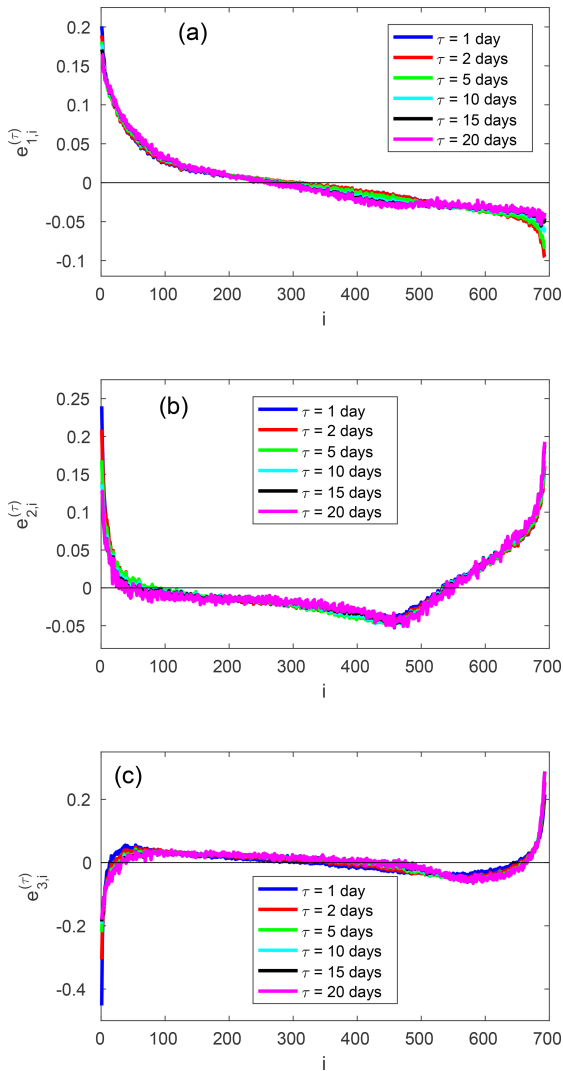
Figure 4. (a) First, (b) second and (c) third normalized empirical orthogonal function (EOF) of the finite-time Lyapunov fluctuations (cf., Eq. 20).

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

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.

In Fig. 4 the three leading EOFs are displayed. The modes are largely independent of the integration time τ and have converged at about $\tau = 10\,\mathrm{d}$. 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, the leading FTLEs again fluctuate in phase; here, this 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, Fig. 5 shows the correlation of selected FTLEs with each of the other FTLEs for $\tau = 1$ and $\tau = 15\,\mathrm{d}$. The pattern of the correlations is the same for both integration times but the amplitudes are very low for $\tau = 1\,\mathrm{d}$ 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.
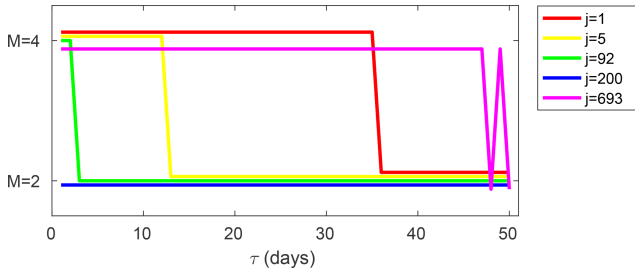
Figure 6. Order of the model for the probability density function of the FTLE $\Lambda_j^{(\tau)}$.

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 and at the smallest, most stable exponent. The large-deviation rate function is estimated as described in Sect. 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. (39) are considered here: $M = 2$, this is to say 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 performed here as follows. For $\tau = n\Delta\tau$, the time series of the FTLEs, $\{\Lambda_{j,\alpha}^{(\tau)}\}_{\alpha=0}^{L-1}$, are sub-sampled into n disjoint time series 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 length of the sub-sampled time series is the largest integer L' such that $m + (L'-1)n \leq L - 1$. The two probability density models are fitted separately on the n sub-sampled time series and model selection is based on the average Bayesian information criterion. Then the selected model is fitted on the whole time series.

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 $\tau = 35$ d; for the fifth exponent, it is less pronounced and visible only up to $\tau = 12$ d. The zero exponent shows only very mild non-Gaussianity which is visible for $\tau = 1$ and $\tau = 2$ d. 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
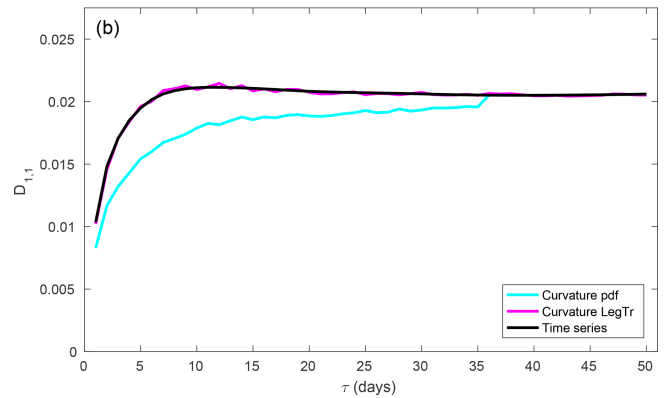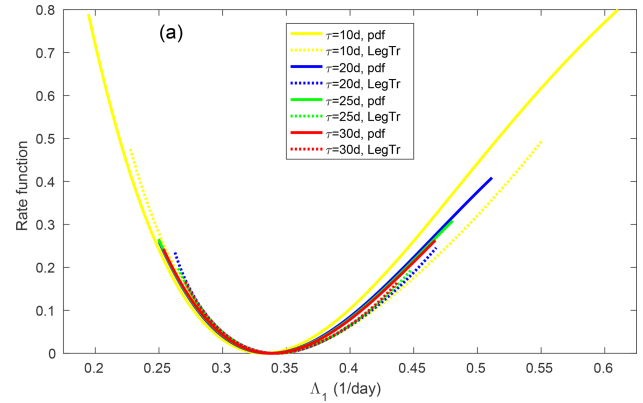

Figure 7. (a) Large-deviation rate function of the first FTLE. **(b)** Element $D_{1,1}$ of the diffusion matrix.

even more pronounced than those for the first exponent and detectable up to an integration time as large as $\tau = 49$ d. 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 shows the results of the large-deviation analysis for the first FTLE. Convergence to a large-deviation principle is observed. At $\tau = 10$ d and even visible at $\tau = 20$ d 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 $\tau = 20$ d. The Legendre transform-based estimates already give a consistent picture from $\tau = 10$ d. Good convergence is also observed for the corresponding element of the diffusion matrix.

For the fifth FTLE, a similar picture can be seen (Fig. 8) but convergence is markedly faster than for the first FTLE. The probability density-based estimates are very consistent from $\tau = 10$–15 d; note that the model for the probability density jumps from fourth-order to Gaussian for the higher values of τ. The Legendre transform already gives close agreement for the rate function from $\tau = 5$ d.
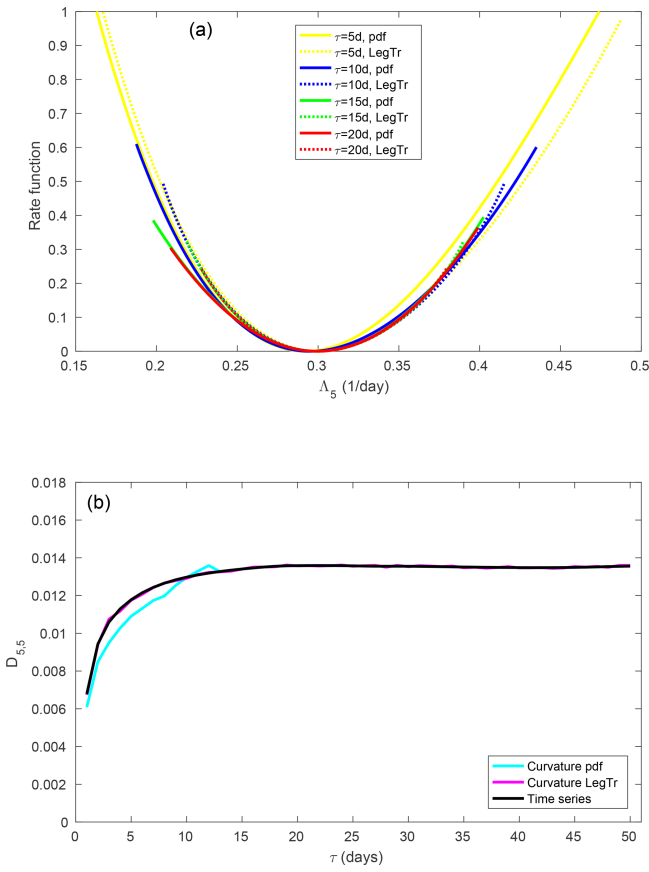
Figure 8. (a) Large-deviation rate function of the fifth FTLE. (b) Element $D_{5,5}$ of the diffusion matrix.

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

For the zero exponent (Fig. 9), convergence is again markedly faster than for both positive exponents. A large-deviation principle can already be established from about $\tau = 10$ d, 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 (Fig. 10), convergence is even faster. A large-deviation principle is valid from $\tau = 5$ d and all of the estimates of the rate function are in almost perfect agreement. The estimates of the diffusion coefficient show corresponding behaviour.

For the smallest, most dissipative exponent (Fig. 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 $\tau = 30$ d, and the Legendre transform method gives reliable estimates of the rate function from $\tau = 10$ to 20 d. Also, the convergence of the diffusion coefficient is markedly slow. The estimate from the non-Gaussian probability density is initially too low and converges at about $\tau = 25$ d.

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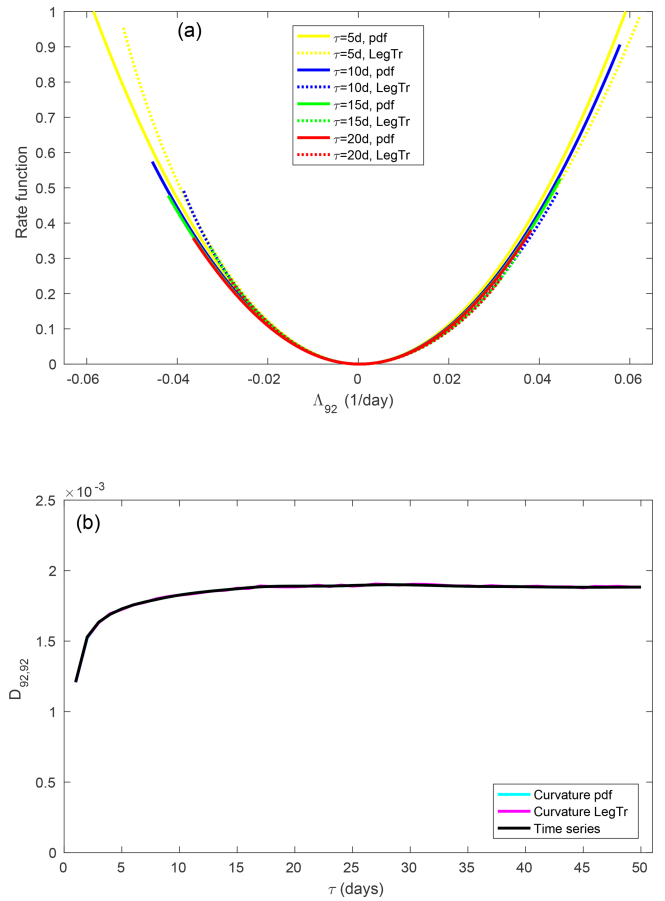
Table 6. Correlation length $l_j^{(\Delta\tau)}$ of the time series of the FTLE $\Lambda_j^{(\Delta\tau)}$ with $\Delta\tau = 1$ d.

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

FTLEs. The correlation length of the FTLE $\Lambda_j^{(\tau)}$ is defined as

$$l_j^{(\tau)} = 1 + 2\sum_{i=1}^{\infty} \rho_{j,i}^{(\tau)}, \qquad (54)$$

where $\rho_{j,i}^{(\tau)}$ is the autocorrelation of the FTLE $\Lambda_j^{(\tau)}$ at lag i. Note that a lag of 1 here refers to two consecutive but non-overlapping integration time intervals of length τ, that is to say $\rho_{j,i}^{(\tau)}$ is the autocorrelation at lag i of the sub-sampled time series of $\Lambda_j^{(\tau)}$ as introduced above for the model selection of the probability density model, $\{\Lambda_{j,m}^{(\tau)}, \Lambda_{j,m+n}^{(\tau)}, \Lambda_{j,m+2n}^{(\tau)}, \ldots\}$, for $m = 0, 1, \ldots, n-1$. There are n of these which can be used to generate n estimates of
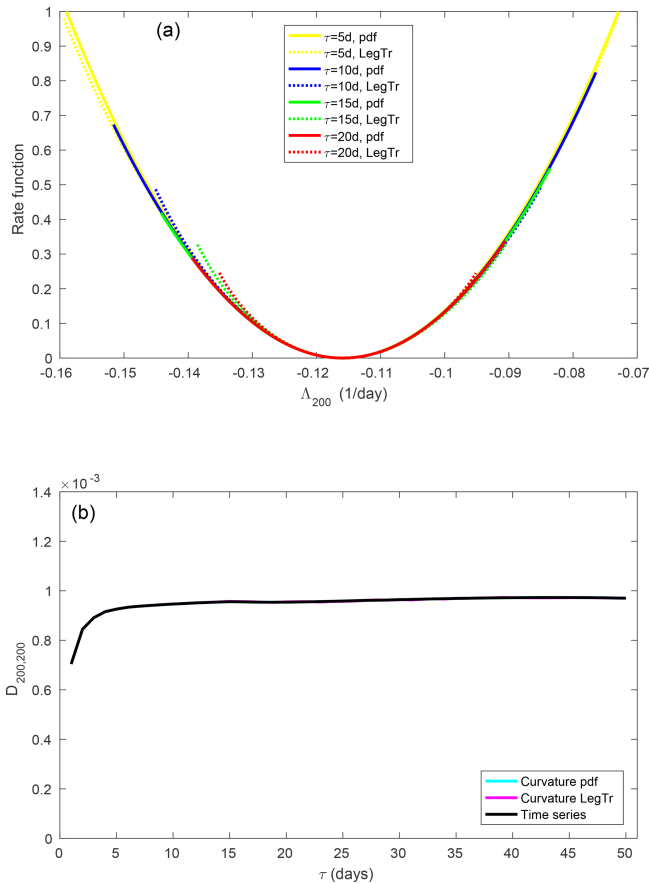
Figure 10. (a) Large-deviation rate function of the 200th FTLE. **(b)** Element $D_{200,200}$ of the diffusion matrix.
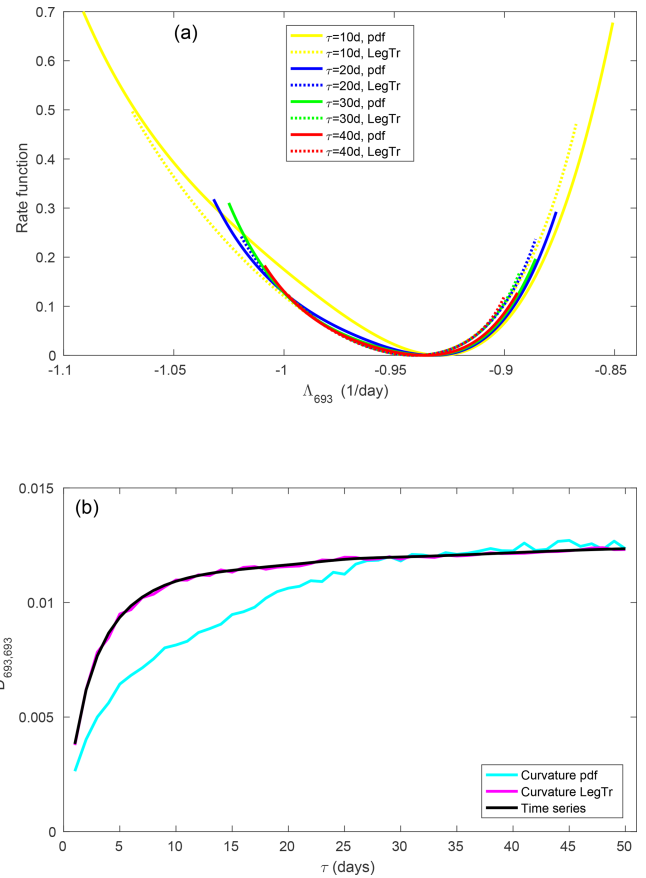

Figure 11. (a) Large-deviation rate function of the 693th FTLE. **(b)** Element $D_{693,693}$ of the diffusion matrix.

$l_j^{(\tau)}$ and then take the average. The definition of the correlation length of Eq. (54) occurs naturally in the formulation of the CLT for dependent random variables (e.g. Billingsley, 1995) under the assumption of a Markov process that is sufficiently mixing. Now consider two integration times $\tau_1 = n_1 \Delta\tau$ and $\tau_2 = n_2 \Delta\tau$ with $n_1 \leq n_2$ and $n' = n_2/n_1$ being an integer for simplicity; one could consider a continuous integration time τ in the limit $\Delta\tau \to 0$. The variances of $\Lambda_j^{(\tau_1)}$ and $\Lambda_j^{(\tau_2)}$ are linked as $\left[\sigma_j^{(\tau_2)}\right]^2 = \left[\sigma_j^{(\tau_1)}\right]^2 l_j^{(\tau_1)}/n'$, and the two estimates of the diffusion coefficient, as calculated from the time series or the Legendre transform, are linked as $D_{j,j}^{(\tau_2)} = D_{j,j}^{(\tau_1)} l_j^{(\tau_1)}$. This holds in the limit $n' \to \infty$, otherwise $l_j^{(\tau_1)}$ needs to be replaced with a counterpart that takes only a finite number of lags into account and also contains some correction terms. Convergence to a large-deviation principle is limited by serial correlation of the FTLEs. Convergence to the diffusion limit, that is to the Gaussian approximation of the large-deviation regime, can certainly not be expected before the serial correlations have decayed, that is to say when $l_j^{(\tau)} \approx 1$. If the distribution of the FTLEs is Gaussian or close to Gaussian the large-deviation limit is equivalent to the diffusion limit and convergence occurs immediately after correlation decay; otherwise it is further delayed, generally the longer the delay the larger the departure from Gaussianity.

Table 6 gives the correlation length of the selected FTLEs $\Lambda_j^{(\Delta\tau)}$ at the basic integration time $\Delta\tau = 1$ d. Note that $l_j^{(\Delta\tau)}$ does not allow one to directly calculate the value of τ at which convergence to a large-deviation principle occurs but it gives an impression of the timescales of temporal correlation and how they differ for the different FTLEs. Overall, temporal correlation is not very pronounced for all of the FTLEs, but the correlation length varies by a factor of 2.35 from the shortest to the longest. The rapid convergence to a large-deviation principle for the zero and weakly dissipative exponents is in line with their short correlation length and almost Gaussian distribution. For the first and the last exponent convergence is delayed beyond what is expected from the somewhat larger correlation length due to the strong non-Gaussianity .

6.3.2 Two-dimensional approach

As an example of a multivariate large-deviation analysis, Fig. 12 shows the joint large-deviation rate function of the
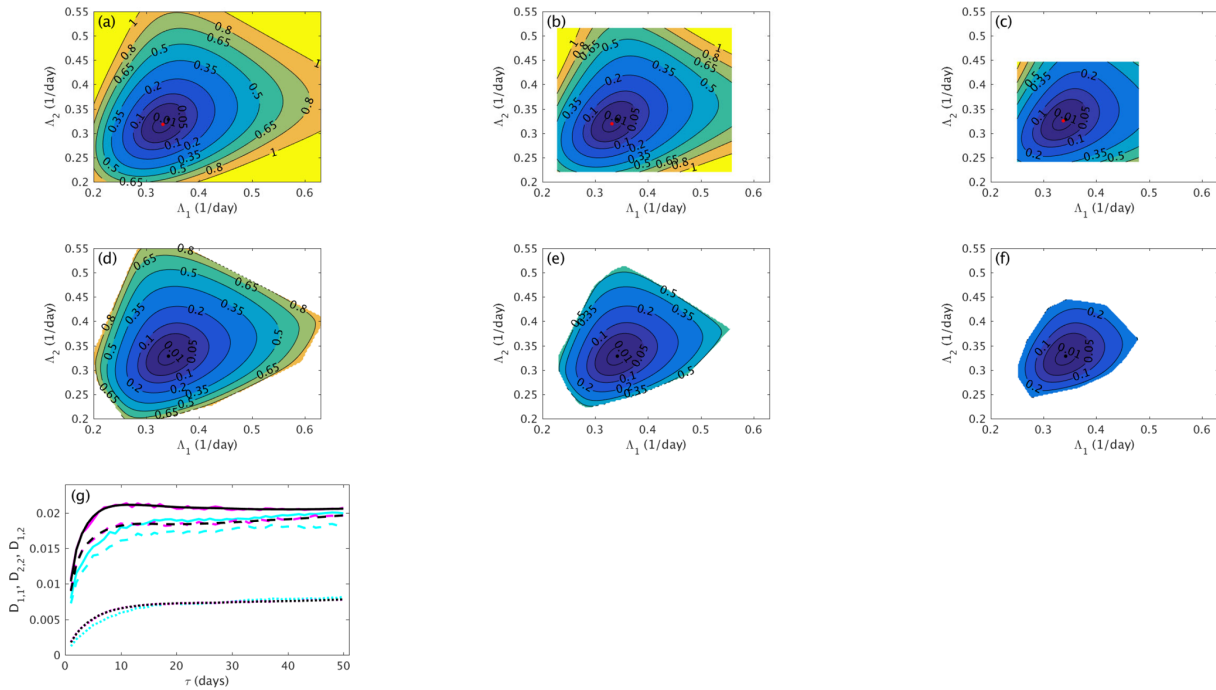
Figure 12. Joint large-deviation rate function of the first two FTLEs as estimated from the joint probability density for **(a)** $\tau = 10$ d, **(b)** $\tau = 15$ d and **(c)** $\tau = 25$ d; and with the Legendre transform for **(d)** $\tau = 10$ d, **(e)** $\tau = 15$ d and **(f)** $\tau = 25$ d. Black dots indicate the global LEs (λ_1, λ_2); and red dots in 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).

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

$$U^{(\tau)}(z_1, z_2) = \sum_{i=1}^{M} \sum_{j=0}^{i} \beta_{i,j}^{(\tau)} \left(\frac{z_1 - \lambda_1}{\sigma_1^{(\tau)}} \right)^{i-j} \left(\frac{z_2 - \lambda_2}{\sigma_2^{(\tau)}} \right)^{j},$$
(55)

where the order of the model is fixed a priori at $M = 4$. The joint rate function displays markedly non-Gaussian behaviour and some dependence between $\Lambda_1^{(\tau)}$ and $\Lambda_2^{(\tau)}$. Convergence to a large-deviation principle is mainly reached at $\tau = 15$ d, 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 already indicate the joint rate function at $\tau = 10$ d. The elements of the diffusion matrix are well estimated overall 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 $\tau = 15$ d. 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 is slightly slow for the most unstable and the most stable FTLEs and very fast in between. Convergence to the diffusion limit, that is to the Gaussian approximation of the large-deviation regime, is generally faster. Also a joint large-deviation rate function for the first and the second FTLE was successfully estimated beyond the Gaussian approximation. Good correspondence was found between the curvature of the rate functions at the minimum and the corresponding elements of the diffusion matrix.

Two different methods for estimating the large-deviation rate functions from the data were discussed: an approach via the probability density function and an approach using the Legendre transform. The Legendre transform method appears to be generally superior for finding the rate function as (i) convergence occurs at a smaller value of the integration time τ where more independent data are available and (ii) it yields diffusion coefficients fully consistent with their direct estimation from the data. Nevertheless, both methods should be considered side by side as the probability density approach allows one to monitor if/when the probability density function has actually reached the large-deviation regime.

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.

Billingsley, P.: Probability and Measure, 3rd edn., Wiley, New York, 1995.

Eckmann, J. and Ruelle, D.: Ergodic theory of chaos and strange attractors, Rev. Mod. Phys., 57, 617–656, 1985.

Ehrendorfer, M.: The total energy norm in a quasigeostrophic model, J. Atmos. Sci., 57, 3443–3451, 2000.

Johnson, P. L. and Meneveau, C.: Large-deviation joint statistics of the finite-time Lyapunov spectrum in isotropic turbulence, Phys. Fluids, 27, 085110, https://doi.org/10.1063/1.4928699, 2015.

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

Kifer, Y.: Large deviations in dynamical systems and stochastic processes, T. Am. Math. Soc., 321, 505–524, 1990.

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

Kuptsov, P. V. and Politi, A.: Large-deviation approach to space-time chaos, Phys. Rev.Lett., 107, 114101, https://doi.org/10.1103/PhysRevLett.107.114101, 2011.

Kwasniok, F.: Reduced atmospheric models using dynamically motivated basis functions, J. Atmos. Sci., 64, 3452–3474, 2007.

Laffargue, T., Lam, K.-D. N. T., Kurchan, J., and Tailleur, J.: Large deviations of Lyapunov exponents, J. Phys. A-Math. Theor., 46, 254002, https://doi.org/10.1088/1751-8113/46/25/254002, 2013.

Legras, B., and Ghil, M.: Persistent Anomalies, Blocking, and Variations in Atmospheric Predictability, J. Atmos. Sci., 42, 433–471, 1985.

Lorenz, E. N.: Deterministic Nonperiodic Flow, J. Atmos. Sci., 20, 130–141, 1963.

Marshall, J. and Molteni, F.: Toward a dynamical understanding of planetary-scale flow regimes, J. Atmos. Sci., 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, Cambridge, 2002.

Pazó, D., López, J. M., and Politi, A.: Universal scaling of Lyapunov-exponent fluctuations in space-time chaos, Phys. Rev. E, 87, 062909, https://doi.org/10.1103/PhysRevE.87.062909, 2013.

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

Roads, J. O.: Predictability in the extended range, J. Atmos. Sci., 44, 3495–3527, 1987.

Rohwer, C. M., Angeletti, F., and Touchette, H.: Convergence of large-deviation estimators, Phys. Rev. E, 92, 052104, https://doi.org/10.1103/PhysRevE.92.052104, 2015.

Schubert, S. and Lucarini, V.: Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: Analysis of instabilities and feedbacks, Q. J. Roy. Meteorol. Soc., 141, 3040–3055, 2015.

Shimada, I. and Nagashima, T.: A Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems, Prog. Theor. Phys., 61, 1605–1616, 1979.

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

Vannitsem, S.: Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics, Chaos, 27, 032101, https://doi.org/10.1063/1.4979042, 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, J. Phys. A-Math. Theor., 49, 224001, https://doi.org/10.1088/1751-8113/49/22/224001, 2016.

Vannitsem, S. and Nicolis, C.: Lyapunov Vectors and Error Growth Patterns in a T21L3 Quasigeostrophic Model, J. Atmos. Sci., 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, Phys. Rev. Lett., 102, 074102, https://doi.org/10.1103/PhysRevLett.102.074102, 2009.