Fluctuations and large deviations of Reynolds' stresses in zonal jet dynamics

Freddy Bouchet, J B Marston, T B Tangarife

Fluctuations and large deviations of Reynolds stresses in zonal jet dynamics

F. Bouchet,[1, a)] J. B. Marston,[2, b)] and T. Tangarife[1]

[1)] *Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France*

[2)] *Department of Physics, Box 1843, Brown University, Providence, RI 02912-1843 USA*

(Dated: 19 December 2017)

The Reynolds stress, or equivalently the average of the momentum flux, is key to understanding the statistical properties of turbulent flows. Both typical and rare fluctuations of the time averaged momentum flux are needed to fully characterize the slow flow evolution. The fluctuations are described by a large deviation rate function that may be calculated either from numerical simulation, or from theory. We show that, for parameter regimes in which a quasilinear approximation is accurate, the rate function can be found by solving a matrix Riccati equation. Using this tool we compute for the first time the large deviation rate function for the Reynolds stress of a turbulent flow. We study a barotropic flow on a rotating sphere, and show that the fluctuations are highly non-Gaussian. This work opens up new perspectives for the study of rare transitions between attractors in turbulent flows.

PACS numbers: 47.27.eb, 47.27.wg, 05.40.-a, 05.10.Gg

Keywords: two-dimensional turbulence, zonal jets, rare events, large deviation theory, quasi-linear dynamics, Riccati equation

[a)]Electronic mail: freddy.bouchet@ens-lyon.fr

[b)]Electronic mail: marston@brown.edu

TOMÁS

Regretfully, Tomás Tangarife suddenly and unexpectedly passed away a few months before completing the research reported in this paper. Most of the science discussed in this paper was developed in patient work by Tomás, and is part of his PhD thesis. F. Bouchet and J. B. Marston pay homage to the unique friendship and passion for science of Tomás, and would like to remember the intense and enriching collaboration that led to these scientific results. Tomás' quiet and constant character, his generosity, and his deep thoughts, were always a source of happiness and joy to his friends and colleagues.

I. INTRODUCTION

For a wide range of applications, in physics, engineering, and geophysics, the determination of the behavior of the average or typical behavior of a turbulent flow is a key issue. Since the work of Reynolds more than one century ago, the role of momentum fluxes and their divergence, or their averages called Reynolds stresses, have been recognized to play the key role. In order to be more specific, we now consider the very simple case of a two dimensional flow on a plane or in a channel, with an average flow that is parallel to the $\mathbf{e}_x$ direction, $U(y)\mathbf{e}_x$ (where x and y are Cartesian coordinates). We also assume that all averaged quantities do not depend on x. The spatially averaged equation of motion for the fluid then reads

$$\frac{\partial U}{\partial t} = -\frac{\partial}{\partial y}\mathbb{E}\left(< uv >\right) + D\left[U\right], \tag{1}$$

where $D[U]$ is the average dissipation operator, $\mathbb{E}\left(< uv >\right)$ is the Reynolds stress, and $\frac{\partial}{\partial y}\mathbb{E}\left(< uv >\right)$ is the momentum flux divergence along the $\mathbf{e}_x$ direction. The symbol $\mathbb{E}$ is either an ensemble or time average (for a time average $\partial U/\partial t = 0$), while $< . >$ denotes a spatial average. The spatial average is an average along the $\mathbf{e}_x$ direction. The spatial average can be avoided, but it is often useful to include for practical reasons. Because the Reynolds stress is the key quantity that determines the average flow behavior it has been extensively studied experimentally, numerically and theoretically, in a wide range of turbulent flows (see for instance classical turbulence textbooks[1,2].

Beyond the average value, fluctuations of the momentum flux $< uv >$, or its divergence $\frac{\partial}{\partial y}\left(< uv >\right)$, are very important quantities in a variety of dynamical circumstances. By

contrast with the average value, as far as we know, no work has been devoted so far to study such fluctuations, and we undertake this task as the main aim of the paper. An important example of when fluctuations play an important role is in the case of time scale separation between the typical time τ_U for the evolution of the parallel flow (or jet) and the time τ_e for the evolution of the turbulent fluctuations (or eddies): $\tau_e \ll \tau_U$. Such time scale separation is common when the parallel flow has a very large amplitude; classical examples include some regimes of two dimensional, geostrophic, or plasma turbulence. Then, following the classical results of stochastic averaging for systems with two timescales, a natural generalization of Reynolds average equation is

$$\frac{\partial U}{\partial t} = -\frac{\partial}{\partial y}\mathbb{E}_U\left(< uv >\right) + \frac{\partial}{\partial y}\zeta_U + D\left[U\right], \tag{2}$$

where now $\mathbb{E}_U$ means an average over a time window short compared to the typical time evolution of the parallel flow U, and we still call $\mathbb{E}_U\left(< uv >\right)$ the Reynolds stress that now depends on the state of U at time t, and $\zeta_U(y,t)$ characterizes the Gaussian typical fluctuations of the momentum flux $< uv >$. $\mathbb{E}_U\left(< uv >\right)$ and ζ_U represent two aspects of the action of the unresolved eddies on the mean flow, the average and typical fluctuations respectively. In such a situation of time scale separation, ζ_U is a white in time Gaussian field whose variance is related through a Kubo formula to the variance of the time average of the momentum flux

$$r_v = \frac{1}{T}\int \mathrm{dt}\ < uv >, \tag{3}$$

where the time average is over a time window of duration T, which is assumed to be short compared to the time scale for the evolution of U, but large compared with the evolution of the turbulent fluctuations: $\tau_e \ll T \ll \tau_U$. We call the fluctuation of (3) the Reynolds stress fluctuations (the fluctuation of the time averaged momentum fluxes, over finite but long times T).

In many instances, rarer and non Gaussian fluctuations are also important. Then (2) does not contain the relevant information and one wants to go beyond the study of the second moment of (3). In the asymptotic regime $\tau_e \ll T$, the probability distribution function of r_v takes a very simple form $P(r_v,T) \underset{T\to\infty}{\asymp} \exp\left(-TI_v[r_v]\right))$, where $\asymp$ is a logarithmic equivalence (the logarithms of the right and left hand sides of the equation are equivalent in the limit $T \to \infty$). This relation is called the large deviation principle. (For a review, see Ref. 3.) The large deviation rate function $I_v[r_v]$ then characterizes the fluctuations of

the time averaged Reynolds stress, both typical (the second variations of $I_v[r_v]$ then gives the statistics of ζ_U), and very rare. In many examples of turbulent flows, it has been observed that the dynamics has several "attractors" (see for instance[4] and references therein ; by "attractor" we mean here stationary solutions of the deterministic Reynolds equation $\frac{\partial U}{\partial t} = -\frac{\partial}{\partial y}\mathbb{E}_U\left(uv\right)$). Then rare fluctuations of the Reynolds stress characterized by the large deviation rate function I_v, are responsible for rare transitions between attractors. For all these reasons, it is very important to be able able to compute I_v and to be able to study its properties from a fluid mechanics point of view.

We develop theoretical and numerical tools to study Reynolds stress fluctuations, and compute the large deviation rate function I_v. First we sample empirically (from time series generated from numerical simulations) the large deviation rate function, using the method developed in 5. In addition to this empirical approach, we determine the Reynolds stress fluctuations and large deviation rate function directly for the case of the quasilinear approximation to the full non-linear dynamics. The quasilinear approximation amounts at neglecting the eddy-eddy interactions (fluctuation + fluctuation $\rightarrow$ fluctuation triads) while retaining interactions between the mean flow and the eddies, and may thus be expected to be accurate when the average flow is much larger than the fluctuations. Such a quasi-linear approximation, investigated at least as early as 1963 by Herring[6], is believed to be accurate for the 2D Navier-Stokes equation, barotropic flows, or quasigeostrophic models, on either a plane, a torus, or a sphere, for a range of parameters (discussed below). Two dimensional flows are a particularly favorable setting for the quasi-linear approximation because, as Kraichnan showed in his seminal 1967 paper[7], an inverse cascade of energy to the largest scales is expected, leading to the formation of coherent structures with non-trivial mean flows[8]. For unforced perfect flows, the large scale structures can be predicted through equilibrium statistical mechanics (see for instance[9]). For forced and dissipated flows eddies both sustain, and interact with, the large-scale flows, and both processes are captured by the quasi-linear approximation. By contrast, the scale-by-scale cascade of energy that plays a central role in Kraichnan's picture[7] relies on eddy + eddy $\rightarrow$ eddy processes that are neglected in the quasi-linear approximation[10,11].

The quasilinear approximation has been shown to be self-consistent[12] in the limit when a time scale separation exists between a typical large scale flow inertial time scale τ_i and

a flow spin up or spin down time scale τ_s: $\tau_i \ll \tau_s$ (then $\tau_U \simeq \tau_s$ and $\tau_e \simeq \tau_i$). This time scale separation condition may however not be necessary. Other factors may favor the validity of the quasilinear approximation, for instance the forcing of the flow through a large number of independent modes, through either a broad band spectrum, or a small scale forcing, keeping the total energy injection rate fixed. The energy transfer is then the same, but each eddy has reduced amplitude, lessening the interaction between eddies. The range of validity of the quasilinear approximation has not been fully understood yet. When the quasilinear approximation is valid, and when one further assumes that the force acts on small scales only, one can predict explicitly the averaged Reynolds stress[13–15] and sometimes the averaged velocity profile. The Gaussian fluctuations of the Reynolds stress may be parameterized phenomenologically[10,11]. The spatial structure of the Gaussian fluctuations has also been studied theoretically. It has been proven to have a singular part with white in space correlation function and a smooth part (see[16], section 1.4.3, or[17], see also[18]).

Within the context of the quasilinear approximation, we show that the Reynolds stress fluctuations and its large deviation rate function can be studied by solving a matrix Riccati equation. The equation can be easily implemented and solved by a generalization of the classical tools used to solve Lyapunov equation for the two points correlation functions. This mathematical result is the main reason why we study the Reynolds stress fluctuations for the quasilinear dynamics in this first study. Moreover we show that the matrix Riccati equation is a much more computationally efficient way to study rare fluctuations than through the traditional route of direct numerical solution. The calculation is illustrated for the case of barotropic flow on the sphere[11], for which the relevance of the quasilinear approximation, over certain parameter ranges, has been recognized for a some time now. For the case of a barotropic flow it is technically more convenient to discuss the dynamics in terms of the equation of motion for the vorticity, so we study the corresponding Reynolds stress that drives the vorticity.

Section II introduces the barotropic equation on the sphere and its quasilinear approximation. Section III discusses the fluctuations of the Reynolds stresses, without time average. Section IV is an introduction to averaging for stochastic processes. It explains pedagogically how an equation for the slow degrees of freedom, for instance the Reynolds equation (2), can be obtained. The relation between the statistics of the noise term, ζ_U, in equation (2), and the large deviation of the Reynolds stress (3) is explained. A short introduction

to the large deviation rate function is also provided. Finally, the matrix Riccati equation that permits the direct calculation of the large deviation rate function is derived both in a general framework, and in the case of the quasilinear approximation of the barotropic equation on the sphere. Section V uses the solution of the matrix Riccati equation in order to study numerically the zonal energy balance and the time scale separation in the inertial limit. Section VI discusses the computation of the large deviation rate function for the time averaged Reynolds stresses of the barotropic equation on the sphere. Section VII discusses the main conclusions and presents some perspectives.

II. BAROTROPIC EQUATION AND QUASI–LINEAR APPROXIMATION

Here we discuss the barotropic equation and its quasilinear approximation that is expected to be valid when a time scale separation exists between the typical time for the evolution of the zonal flow and that of the evolution of the eddies. We study the dynamics of zonal jets in the quasi-geostrophic one-layer barotropic model on a sphere of radius a, rotating at rate Ω,

$$\begin{cases} \dfrac{\partial \omega}{\partial t} + J(\psi,\omega) + \dfrac{2\Omega}{a^2}\dfrac{\partial \psi}{\partial \lambda} = -\kappa\omega - \nu_n\left(-\Delta\right)^n \omega + \sqrt{\sigma}\eta, \\ u = -\dfrac{1}{a}\dfrac{\partial \psi}{\partial \phi}, \quad v = \dfrac{1}{a\cos\phi}\dfrac{\partial \psi}{\partial \lambda}, \quad \omega = \Delta\psi \end{cases} \tag{4}$$

where ω is the relative vorticity, $\mathbf{v} = (u, v)$ is the horizontal velocity field, ψ is the stream function and $J(\psi,\omega) = \frac{1}{a^2\cos\phi}\left(\partial_\lambda\psi\cdot\partial_\phi\omega - \partial_\lambda\omega\cdot\partial_\phi\psi\right)$ is the Jacobian operator. The coordinates are denoted $(\lambda,\phi) \in [0, 2\pi]\times[-\pi/2, \pi/2]$, λ is the longitude and ϕ is the latitude. All fields ω, u, v and ψ can be decomposed onto the basis of spherical harmonics $Y_\ell^m(\phi,\lambda)$, for example

$$\psi\left(\phi,\lambda\right) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\psi_{m,\ell}\; Y_\ell^m(\phi,\lambda) \tag{5}$$

All fields ω, u, v and ψ are 2π-periodic in the zonal (λ) direction, so we can also define the Fourier coefficients in the zonal direction,

$$\psi_m(\phi) \equiv \frac{1}{2\pi}\int_0^{2\pi}\psi(\phi,\lambda)\;\mathrm{e}^{-im\lambda}\,\mathrm{d}\lambda = \sum_{\ell=|m|}^{\infty}\psi_{m,\ell}\;P_\ell^m(\sin\phi), \tag{6}$$

with the associated Legendre polynomials $P_\ell^m(\sin\phi)$.

In (4), κ is a linear friction coefficient, also known as Ekman drag or Rayleigh friction, that models the dissipation of energy at the large scales of the flow[19]. Hyper-viscosity $\nu_n\left(-\Delta\right)^n$

accounts for the dissipation of enstrophy at small scales and is used mainly for numerical reasons. Most of the dynamical quantities are independent of the value of ν_n, for small enough ν_n. η is a Gaussian noise with zero mean and correlations $\mathbb{E}\left[\eta\left(\lambda_1,\phi_1,t_1\right)\eta\left(\lambda_2,\phi_2,t_2\right)\right] = C\left(\lambda_1-\lambda_2,\phi_1,\phi_2\right)\delta\left(t_1-t_2\right)$, where C is a positive-definite function and $\mathbb{E}$ is the expectation over realizations of the noise η. C is assumed to be normalized such that σ is the average injection of energy per unit of time and per unit of mass by the stochastic force $\sqrt{\sigma}\eta$. There is no symmetry reason to enforce homogeneous forcing over a rotating sphere, which only has axial symmetry. Thus it is natural to consider forcing that varies with latitude. The barotropic equation is sometimes used to describe the vertically-averaged atmospheric dynamics. The stochastic forces then model the driving influence of the baroclinic instability on the barotropic flow. Baroclinic instabilities are typically strongest at mid-latitude.

A. Time scale separation between large scale and small scale dynamics

1. *Energy balance and non–dimensional equations*

The inertial barotropic model (eq. (4) with $\kappa = \nu_n = \sigma = 0$) conserves the energy $\mathcal{E}\left[\omega\right] = -\frac{1}{2}\int \omega\psi \,\mathrm{d}\mathbf{r}$ (we denote $\mathrm{d}\mathbf{r} = a^2\cos\phi\,\mathrm{d}\phi\mathrm{d}\lambda$), the moments of potential vorticity $\mathcal{C}_m\left[\omega\right] = \int(\omega+f)^m\,\mathrm{d}\mathbf{r}$ with the Coriolis parameter $f(\phi) = 2\Omega\sin\phi$, and the angular momentum $L[\omega] = \int \omega\cos\phi\,\mathrm{d}\mathbf{r}$. The average energy balance for the dissipated and stochastically forced barotropic equation is obtained applying the Ito formula[20] to (4). It reads

$$\frac{dE}{dt} = -2\kappa E - 2\nu_n Z_n + \sigma, \tag{7}$$

where $E = \mathbb{E}\left[\mathcal{E}\left[\omega\right]\right]$ is the total average energy and $Z_n = \mathbb{E}\left[-\frac{1}{2}\int\psi(-\Delta)^n\omega\,\mathrm{d}\mathbf{r}\right]$. The term $-2\nu_n Z_n$ in (7) represents the dissipation of energy at the small scales of the flow. In the regime we are interested in, most of the energy is concentrated in the large-scale zonal jet, so the main mechanism of energy dissipation is the linear friction (first term in the right-hand side of (7)). In this turbulent regime, energy dissipated by hyper-viscosity can be neglected. Then, in a statistically stationary state, $E_{stat} \simeq \frac{\sigma}{2\kappa}$, expressing the balance between stochastic forces and linear friction in (4).

The estimated total energy yields a typical jet velocity of $U \sim \sqrt{\frac{\sigma}{2\kappa}}$. The order of magnitude of the time scale of advection and stirring of turbulent eddies by this jet is $\tau_{eddy} \sim \frac{a}{U}$. We perform a non-dimensionalization of the stochastic barotropic equation (4)

using τ_{eddy} as unit time and a as unit length. The non-dimensionalization may be carried out by setting $a = 1$ and using the non-dimensionalized variables $t' = t/\tau_{eddy}$, $\omega' = \omega\tau_{eddy}$, $\psi' = \psi\tau_{eddy}$, $\Omega' = \Omega\tau_{eddy}$,

$$\alpha = \kappa\tau_{eddy} = \sqrt{\frac{2\kappa^3}{\sigma}}, \tag{8}$$

$\nu_n' = \nu_n\tau_{eddy}$, $\sigma' = \sigma\tau_{eddy}^3 = 2\alpha$, and a rescaled force $\eta' = \eta\sqrt{\tau_{eddy}}$ such that $\mathbb{E}\left[\eta'\left(\lambda_1, \phi_1, t_1'\right)\eta'\left(\lambda_2, \phi_2, t_2'\right)\right] = C\left(\lambda_1 - \lambda_2, \phi_1, \phi_2\right)\delta\left(t_1' - t_2'\right)$. In these new units, and dropping the primes for easiness in the notation, the stochastic barotropic equation (4) reads

$$\frac{\partial\omega}{\partial t} + J(\psi, \omega) + 2\Omega\frac{\partial\psi}{\partial\lambda} = -\alpha\omega - \nu_n\left(-\Delta\right)^n\omega + \sqrt{2\alpha}\eta. \tag{9}$$

In (9), α is an inverse Reynolds' number based on the linear friction and ν_n is an inverse Reynolds' number based on hyper-viscosity. The turbulent regime mentioned previously corresponds to $\nu_n \ll \alpha \ll 1$. In such regime and in the units of (9), the total average energy in a statistically stationary state is $E_{stat} = 1$.

We are interested in the dynamics of zonal jets in the regime of small forces and dissipation, defined as $\alpha \ll 1$. In the next section we show that the dynamics corresponds to a regime in which the zonal jet evolves much more slowly than the surrounding turbulent eddies.

2. Decomposition into zonal and non–zonal components

In order to decompose (9) into a zonally averaged flow and perturbations around it, we define the zonal projection of a field

$$\langle\psi\rangle\left(\phi\right) \equiv \psi_0(\phi) = \frac{1}{2\pi}\int_0^{2\pi}\psi(\lambda, \phi)\,\mathrm{d}\lambda.$$

The zonal jet velocity profile is then defined by $U(\phi) \equiv \langle u\rangle\left(\phi\right)$. In most situations of interest, the stochastic force in (9) does not act direcly on the zonal flow: $\langle\eta\rangle = 0$. Then the perturbations of the zonal jet is proportional to the amplitude of the stochastic force in (9). We thus decompose the velocity field as $\mathbf{v} = U\mathbf{e}_x + \sqrt{\alpha}\delta\mathbf{v}$ and the relative vorticity field as $\omega = \omega_z + \sqrt{\alpha}\delta\omega$ with $\omega_z \equiv \langle\omega\rangle$, and where α is the non-dimensional parameter defined in (8). We call the perturbation velocity $\delta\mathbf{v}$ and vorticity $\delta\omega$ the eddy velocity and eddy vorticity, respectively.

With the decomposition of the vorticity field, the barotropic equation (9) reads

$$\begin{cases} \dfrac{\partial \omega_z}{\partial t} = \alpha R - \alpha \omega_z - \nu_n \left(-\Delta\right)^n \omega_z \\ \dfrac{\partial \delta\omega}{\partial t} = -L_U\left[\delta\omega\right] - \sqrt{\alpha} NL\left[\delta\omega\right] + \sqrt{2}\eta, \end{cases} \tag{10}$$

with

$$R(\phi) \equiv -\left\langle J\left(\delta\psi, \delta\omega\right)\right\rangle \tag{11}$$

the zonally averaged advection term, where the linear operator L_U reads

$$L_U\left[\delta\omega\right] = \frac{1}{\cos\phi}\left(U(\phi)\partial_\lambda\delta\omega + \gamma(\phi)\partial_\lambda\delta\psi\right) + \alpha\delta\omega + \nu_n\left(-\Delta\right)^n\delta\omega, \tag{12}$$

with $\gamma\left(\phi\right) = \partial_\phi\omega_z(\phi) + 2\Omega\cos\phi$, and where

$$NL\left[\delta\omega\right] = J(\delta\psi, \delta\omega) - \left\langle J(\delta\psi, \delta\omega)\right\rangle$$

is the non-linear eddy-eddy interaction term.

Using $\omega_z\left(\phi\right) = -\frac{1}{\cos\phi}\partial_\phi\left(U\left(\phi\right)\cos\phi\right)$ and the first equation of (14), we get the evolution equation for the zonal flow velocity $U\left(\phi\right)$

$$\frac{\partial U}{\partial t} = \alpha f - \alpha U - \nu_n\left(-\Delta\right)^n U\,, \tag{13}$$

where $f\left(\phi\right)$ is such that $R\left(\phi\right) = -\frac{1}{\cos\phi}\partial_\phi\left(f\left(\phi\right)\cos\phi\right)$. f is minus the divergence of the Reynolds' stress.

3. Quasi-linear and linear dynamics

In this section we discuss the quasilinear approximation to the barotropic equation and the associated linear dynamics.

In the limit of small forces and dissipation $\alpha \ll 1$, the perturbation flow is expected to be of small amplitude. Then the non-linear term $NL[\delta\omega]$ in (10) is negligible compared to the linear term $L_U\left[\delta\omega\right]$. Neglecting these non-linear eddy-eddy interaction terms, we then obtain the so-called quasi-linear approximation of the barotropic equation[21],

$$\begin{cases} \dfrac{\partial \omega_z}{\partial t} = \alpha R - \alpha \omega_z - \nu_n \left(-\Delta\right)^n \omega_z \\ \dfrac{\partial \delta\omega}{\partial t} = -L_U\left[\delta\omega\right] + \sqrt{2}\eta. \end{cases} \tag{14}$$

The approximation leading to the quasi-linear dynamics (14) amounts at suppressing some of the triad interactions. As a consequence, the inertial quasi-linear dynamics has the same quadratic invariants as the initial barotropic equations. The average energy balance for the quasi-linear barotropic dynamics (14) is thus the same as the one for the full barotropic dynamics (10).

In many situations of interest, for example Jovian jets, the turbulent eddies $\delta\omega$ evolve much faster than the zonal jet velocity profile U[22]. In (10) and (14), the natural time scale of evolution of the zonal jet is of order $1/\alpha$, while the typical time scale of evolution of the perturbation vorticity $\delta\omega$ is of order 1. In the regime $\alpha \ll 1$, we thus expect to observe a separation of time scales between the evolution of ω_z and $\delta\omega$, consistent with the definition of α as the ratio of the inertial time scale τ_{eddy} and of the dissipative time scale $1/\kappa$, see (8).

In the regime $\alpha \ll 1$, it is natural to consider the linear dynamics of $\delta\omega$ with U held fixed,

$$\frac{\partial \delta\omega}{\partial t} = -L_U\left[\delta\omega\right] + \sqrt{2}\eta\,. \tag{15}$$

The relevance of (15) as an effective description of turbulent eddy dynamics is further discussed later. In particular, we show in section V B that the correlation time of Reynolds' stresses resulting from the linear dynamics (15) —the most relevant time scale related to the dynamics of eddies and their action on the evolution of the zonal jet— is of the order or smaller than τ_{eddy}, holding even as α decreases. It means that the time scale separation hypothesis that leads us to consider the linear dynamics (15) is self-consistent in the limit of weak forces and dissipation $\alpha \ll 1$.

4. Reynolds averaging for the vorticity equation

In the introduction we discussed Reynolds averaging and Reynolds stresses for the simplest possible case: a two dimensional flow that does not break the symmetry along the direction $\mathbf{e}_x$. We now adapt the discussion to two dimensional flows on a sphere. As it is much more convenient to work directly with the vorticity equation, we discuss Reynolds averaging for the vorticity equation only.

Our aim is to write the counterpart of Eq. (2) and (3), for the vorticity equation. In the cases when there is a time scale separation between the evolution of the slow zonal and the

fast non zonal part of the flow, averaging either Eq. (10) or Eq. (14) leads to an effective equation for the low frequency evolution of the zonal vorticity

$$\frac{\partial \omega_z}{\partial t} = \alpha \mathbb{E}\left(R\right) - \alpha \omega_z - \nu_n \left(-\Delta\right)^n \omega_z + \xi_{\omega_z}, \tag{16}$$

where $\mathbb{E}\left(R\right)$ is the average of the vorticity flux R (11), and the white in time Gaussian noise ξ_ω describes the typical fluctuations. We consider time averages of the vorticity flux

$$r = \frac{1}{T} \int \mathrm{dt}\, R(u). \tag{17}$$

The average of r is the term $\mathbb{E}\left(R\right)$ appearing in the Reynolds averaged equation (16). We call this term the vorticity Reynolds stress; however it does not have the same physical dimension as the usual stress. When the time average is over a time window of duration T which is assumed to be short compared to the time scale for the evolution of U, but large compared with the evolution of the turbulent fluctuations: $\tau_e \ll T \ll \tau_U$, we call the fluctuations of (17) the vorticity Reynolds stress fluctuations (the fluctuation of the time averaged vorticity fluxes, over finite but long times T). In the asymptotic regime $\tau_e \ll T$, the probability distribution function of r takes the simple large deviation form $P(r,T) \underset{T\to\infty}{\asymp} \exp\left(-TI[r]\right))$. The variance of ξ_ω is given by a Kubo formula, and is simply related to the second variations of I.

We note that there exists a simple relation between the Reynolds stress large deviations rate function I_v, that describes the averages of the actual momentum fluxes that appear in the velocity equation, and the vorticity Reynolds stress large deviation rate function I. In the following we study the vorticity Reynolds stress only. For simplicity, as there is no ambiguity, we call these quantities Reynolds stresses and Reynolds stress large deviation rate functions, omitting the word vorticity.

B. Numerical implementation

Direct numerical simulations (DNS) of the barotropic equation (10), the quasi-linear barotropic equation (14) and the linear equation (15) are performed using a purely spectral code with a fourth-order-accurate Runge-Kutta algorithm and an adaptive time step[23]. The spectral cutoffs defined by $\ell \leq L$, $|m| \leq \min\left\{\ell, M\right\}$ in the spherical harmonics decomposition of the fields are taken to be $L = 80$ and $M = 20$. In all the simulations, the rotation rate of the sphere is $\Omega = 3.7$ in the units defined previously.

The stochastic noise is implemented using the method proposed in Ref. 24, with a non-zero renewal time scale τ_r larger than the time step of integration. For τ_r much smaller than the typical eddy turnover time scale, the noise can be considered as white in time.

Whenever one considers the linear dynamics (15), modes with different values of m decouple, thanks to the zonal symmetry. Then the statistics of the contribution of the Reynolds stress coming from different values of m are independent. The statistics for the total Reynolds stress can be computed from the statistics of the contribution of each independent value of m. It is then natural and simpler to study the contribution from each different value of m independently. For this reason we consider in this study a force that acts on one mode only. However, as explained in the previous section the validity of the quasi-linear approximation is favored by the use of a broad band spectrum forcing, or a forcing acting on a large number of small scale modes, or both. Forcing only one mode is the most unfavorable case from the point of view of the accuracy of the quasilinear approximation. Larger time scale separation may be required in this case to ensure the accuracy of the quasilinear approximation. However whenever the quasilinear approximation is accurate, the statistics of the Reynolds stress arising from the studied mode are accurately described by the methods reported here.

The force only acts on the mode $|m| = 10$, $\ell = 10$, which is concentrated around the equator (see figure 1). With such a forcing spectrum and setting $\alpha = 0.073$, the integration of the quasi-linear barotropic equation (14) leads to a stationary state characterized by a strong zonal jet with velocity $U(\phi)$, represented in Figure 1. We spectrally truncate the jet to its first 25 spherical harmonics to fix the mean flow in the simulation of the linear barotropic equation (15). We use hyper-viscosity of order 4 with coefficient ν_4 such that the damping rate of the smallest mode is 4. To assess that hyper-viscosity is negligible in the large scale statistics, simulations of the linear equation with $\nu_4 = 4$ and $\nu_4 = 2$ are compared in sections III, V and VI.

III. EQUAL-TIME STATISTICS OF VORTICITY FLUXES

The aim of this section is to illustrate that fluctuations of equal-time vorticity flux R (11) may be strongly non Gaussian. We prove that vorticity flux fluctuations have exponential tails with a distribution close to that of Gaussian product statistics[25]. While equal-time

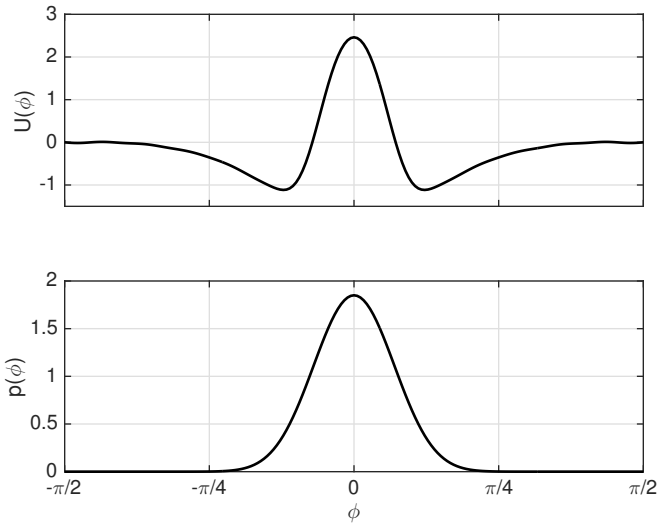


FIG. 1. Top pannel: the zonal flow velocity profile $U(\phi)$ used in numerical simulations of the linearized barotropic equation (15). Bottom panel: zonally averaged energy injection rate by the stochastic force η in (9), (14) and (15).

fluctuations of the vorticity flux are important for high frequency jet variability, Reynolds stresses (time average of the vorticity fluxes) are more important for the long term evolution of the jet. Beginning in section IV, we study Reynolds stresses, and their large deviations.

The evolution of the mean flow $\omega_z(\phi, t)$ is given by the advection term $R(\phi, t) = -\langle J(\delta\psi, \delta\omega)\rangle$, through (10) or (14). In most previous statistical approaches to zonal jet dynamics, only the averaged advection term, the Reynolds stress, was considered. This is for instance the case in S3T[26] and CE2[21,27,28] approaches. Such restriction gives a good approximation of the relaxation of zonal jets towards the attractors of the dynamics, that is expected to be quantitatively accurate in the inertial limit $\alpha \to 0$[12]. However, replacing the advection term R by its average does not describe fluctuations of the vorticity fluxes, that may lead to fluctuations of zonal jets. Understanding the statistics of vorticity fluxes beyond their average value is thus a very interesting perspective. In this section, we study the whole distribution function of vorticity fluxes, as computed from direct numerical simulations.

The zonally averaged advection term is a function of latitude ϕ, it can be decomposed in spherical harmonics according to (5). We denote $R_\ell(t) \equiv R_{0,\ell}(t)$ the ℓ-th component in

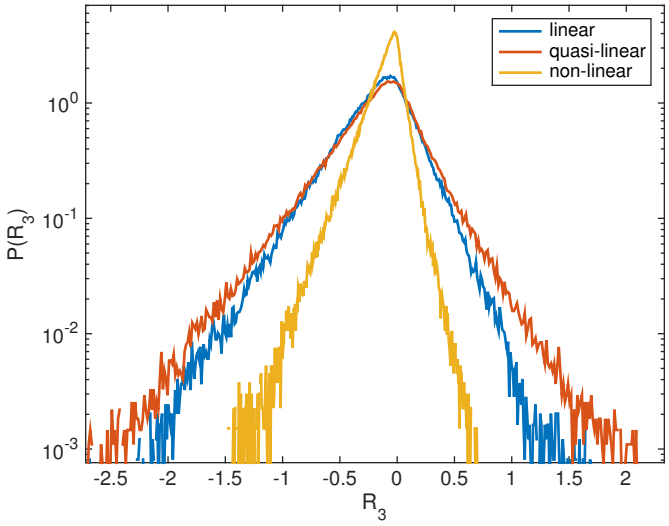


FIG. 2. Probability density functions of R_3, the third component in the spherical harmonics decomposition of the zonally averaged advection term (vorticity flux) $R(\phi)$, from direct numerical simulations of the linear barotropic equation (15) (blue), the quasi-linear barotropic equation (14) (orange), and the non-linear barotropic equation (10) (yellow). Exponential tails are observed in all of the different cases. The common parameters are $\alpha = 0.073$, $\Omega = 3.7$, total integration time $5,450$, and the forcing is concentrated in wavenumbers $|m| = 10$, $\ell = 10$.

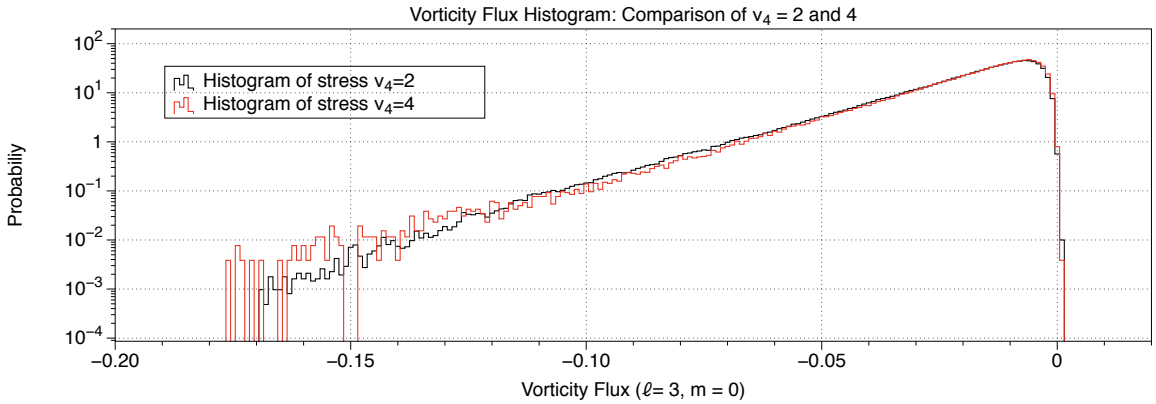


FIG. 3. Probability density functions of R_3, the third component in the spherical harmonics decomposition of the zonally averaged advection term (vorticity flux) R, from direct numerical simulations of the quasi-linear barotropic equation (14) with hyper-viscosity such that the smallest scale has a hyperviscous damping rate of 4 (red curve) and 2 (black curve). The two probability density functions are nearly identical, showing that hyper-viscosity can be considered to be negligible as far as the zonal jet statistics are concerned.

the spherical harmonics decomposition of $R(\phi, t)$. All R_l for odd values of l larger than one have non-zero amplitudes (the amplitude of the $l = 1$ mode is zero because total angular momentum about the polar axis remains zero). In the following, for simplicity, we focus our analysis on R_3 only, that has the largest contribution. The probability density functions of R_3, computed either from direct numerical simulations of the barotropic equation (10), or the quasi-linear barotropic equation (14) or the linear equation (15), with the forcing spectrum specified in section II B and with $\alpha = 0.073$, are shown in Figure 2. Figure 3 shows that the probability distribution of R_3 is not affected by the choice of small scale dissipation.

In the linear dynamics (15), the eddy vorticity evolves according to the linearized barotropic equation close to the fixed base flow $U(\phi)$ shown in Figure 1. In the quasi-linear dynamics (14), the zonal mean flow has the same average velocity profile $U(\phi)$, but this zonal flow is allowed to fluctuate. This difference in the dynamics of the zonal flow between linear and quasi-linear equations explains the slight difference observed in the corresponding advection term histograms (respectively blue curve and orange curve in Figure 2), namely, the probability density function is more spread (the vorticity fluxes fluctuate more) in the quasi-linear dynamics than in the linear dynamics.

In contrast, the probability density function of R_3 computed from the non-linear integration (yellow curve in Figure 2) is very different from the other ones for two reasons: the average zonal flow is different from the fixed zonal flow used in the linear dynamics, and the dynamics of $\delta\omega$ is also different from the quasi-linear dynamics because of the non-linear eddy-eddy interaction terms in (10) (this is expected, as forcing a single mode is the most unfavorable case from the point of view of the validity of the quasilinear approximation, as explained in section II).

In all three cases, the probability distribution functions in Figure 2) show large fluctuations and heavy tails. For instance it is clear that typical fluctuations of the vorticity flux have much larger amplitude than the value of their average (the variance is much larger than the average). While essential for understanding the high frequency and small variability of the jets, on the slow time scale, the jet evolution is described by time averaged vorticity fluxes (the Reynolds stress).

In all of the simulations, the distribution of the vorticity flux shows exponential tails.

This can be easily understood for the case of the linear equation (15). Indeed, in this case the statistics of the eddy vorticity are exactly Gaussian ($\delta\omega$ is an Ornstein-Uhlenbeck process[20]). Then, the statistics of $R(\phi)$ can be calculated explicitly, as we explain now.

Using (6) we can write the vorticity flux as

$$R(\phi) = -\frac{1}{\cos\phi}\sum_m im\left(\psi_m \cdot \partial_\phi \omega_{-m} + \partial_\phi \psi_m \cdot \omega_{-m}\right), \tag{18}$$

where $\omega_m(\phi)$ is the m-th Fourier coefficient of $\delta\omega$, and $\psi_m(\phi)$ is the associated stream function. The Ornstein-Uhlenbeck process $\omega_m(\phi)$ is a Gaussian random variable at each latitude ϕ. The sum of Gaussian random variables is a Gaussian random variable, so $\psi_m(\phi)$, $\partial_\phi\psi_m(\phi)$ and $\partial_\phi\omega_m(\phi)$ are also Gaussian random variables at each latitude ϕ. All these Gaussian random variables have zero mean, and in general they are correlated in a non-trivial way.

The vorticity flux (18) is thus of the form $R = \xi_1\xi_2 + \ldots + \xi_{M-1}\xi_M$ where $\xi_1, \ldots, \xi_M$ are M real-valued[29] correlated Gaussian variables with zero mean. We denote by ξ the column vector with components $\xi_1, \ldots, \xi_M$. By definition, the probability distribution function of ξ is

$$P_\xi(\xi) = \frac{1}{Z}\exp\left(-\frac{1}{2}\xi^T G^{-1}\xi\right),$$

where ξ^T denotes the transpose vector of ξ, G is the covariance matrix of ξ, and Z is a normalisation constant. The probability density function of R, denoted P_R, is given by

$$\begin{aligned} P_R(R) &= \int \mathrm{d}\xi\, P_\xi(\xi)\,\delta\left(R - \xi_1\xi_2 - \ldots - \xi_{M-1}\xi_M\right) \\ &= \int \mathrm{d}\xi_2 \ldots \mathrm{d}\xi_m \frac{1}{|\xi_2|} P_\xi\left(\frac{R - \xi_3\xi_4 - \ldots - \xi_{M-1}\xi_M}{\xi_2}, \xi_2, \ldots \xi_M\right). \end{aligned}$$

Using the change of variable $\zeta_m = \xi_m/\sqrt{|R|}$ for $m = 2, \ldots, M$, the first argument of P_ξ becomes $\sqrt{|R|}\frac{\frac{R}{|R|} - \zeta_3\zeta_4 - \ldots - \zeta_{M-1}\zeta_M}{\zeta_2}$, so we obtain:

$$P_R(R) = \frac{1}{Z}\int \mathrm{d}\zeta_2 \ldots \mathrm{d}\zeta_M \frac{|R|^{\frac{M-2}{2}}}{|\zeta_2|}\exp\left(-|R|\, Q_\pm(\zeta_2, \ldots, \zeta_M)\right),$$

where $Q_\pm$ is a function of $(\zeta_2, \ldots, \zeta_M)$, that depends only on the sign of R, according to $R = \pm|R|$. The tails of the distribution P_R correspond to the limits $R \to \pm\infty$. In both limits, $|R| \to \infty$ so we can perform a saddle-point approximation in the above integral, and get

$$\ln\left(P_R(R)\right) \underset{R\to\pm\infty}{\sim} -|R|\,\mu_\pm, \tag{19}$$

where the rates of decay are defined by

$$\mu_{\pm} = \min_{\zeta_2,\ldots,\zeta_M} \left\{ Q_{\pm}\left(\zeta_2,\ldots,\zeta_M\right) \right\}. \tag{20}$$

The exponential tails of the distribution P_R are direct consequences of the fact that the eddy vorticity $\delta\omega$ evolving according to the linear equation (15) is a Gaussian process and of the fact that R is quadratic in $\delta\omega$. This simple argument explains the exponential tails observed in probability density functions of the zonally averaged advection term in simulations of the linear dynamics (15) (blue curve in Figure 2), where the vorticity field is exactly an Ornstein-Uhlenbeck process.

In the quasi-linear and non-linear dynamics, the zonal flow and eddies evolve at the same time. As a consequence, the dynamics of the eddy vorticity is not linear, and its statistics are not Gaussian. However, we observe that the probability density functions of eddy vorticity are nearly Gaussian (skewness -0.0147 and kurtosis 3.8079 in the quasi-linear case, skewness -0.0037 and kurtosis 3.3964 in the non-linear case, compared to skewness 0.0172 and kurtosis 3.0028 in the linear case). The previous argument can thus also be applied empirically to explain the exponential tails observed in the curves corresponding to quasi-linear and non-linear simulations in Figure 2.

The same analysis has been performed on direct numerical simulations of the deterministic 2-layer quasi-geostrophic baroclinic model[19], see Figure 4. In this case, the eddy vorticity statistics are highly non-Gaussian, while statistics of the vorticity flux have exponential tails similar to those found in the one-layer case. The observation indicates that the previous explicit calculation might not be the most general explanation of the exponential distribution of vorticity fluxes.

IV. AVERAGING AND LARGE DEVIATIONS IN SYSTEMS WITH TIME SCALE SEPARATION

As explained in section II, we are interested in the regime where zonal jets evolve much slower than the surrounding turbulent eddies. In this section, we present some theoretical tools (stochastic averaging, large deviation principle) that can be applied to study the

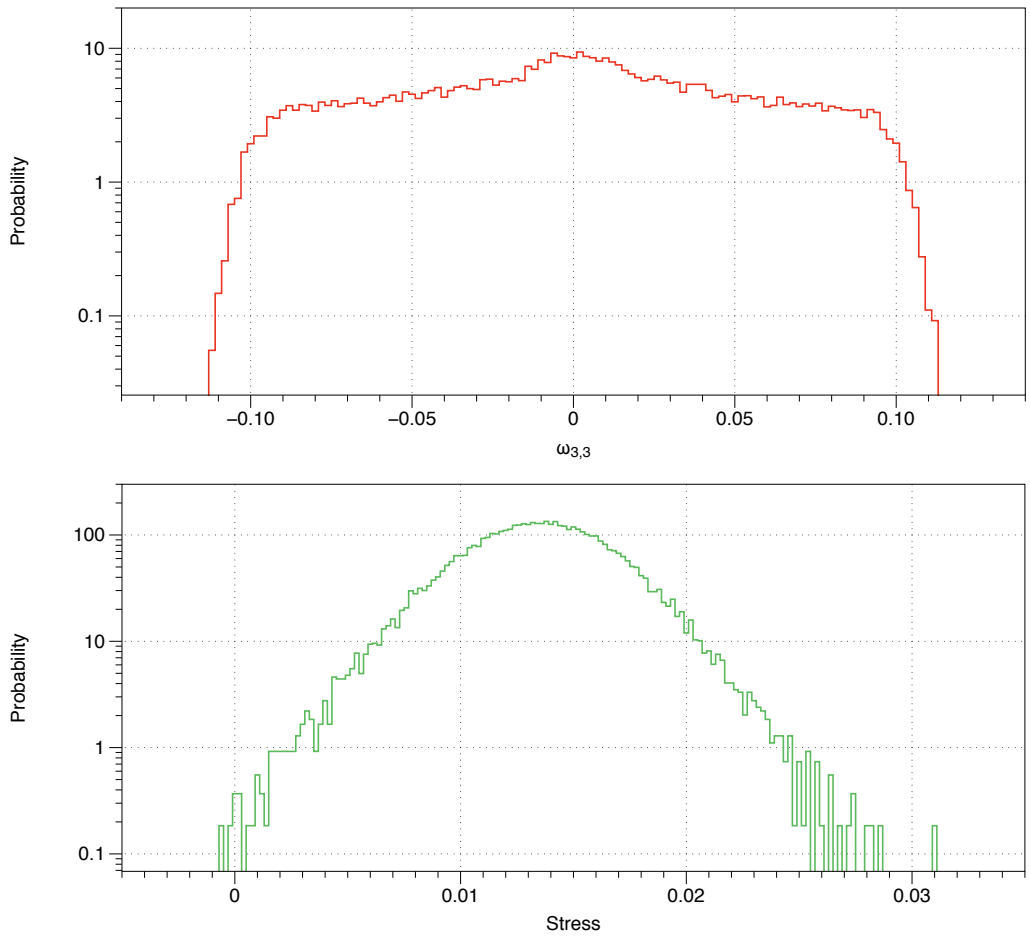


FIG. 4. Probability density functions of the vorticity component $\omega_{3,3}$ (top panel) and zonally averaged advection term (vorticity flux) R_3 (bottom pannel) from a direct numerical simulation of the deterministic 2-layer quasi-geostrophic baroclinic equation. The eddy vorticity is clearly non-Gaussian, and yet the advection term distribution has exponential tails as in the one-layer cases (Figure 2). This observation calls for a more general study of vorticity flux statistics close to a zonal jet.

effective dynamics and statistics of slow dynamical variables coupled to fast stochastic processes. Most of these tools are classical ones[20,30,31], except for the explicit results presented in section IV C 2[32]. Application of these general tools to the quasi-linear barotropic model is considered in sections V and VI.

Consider the stochastic dynamical system

$$\begin{cases} \dfrac{dx}{dt} = \alpha f\left(x,y\right) \\ \dfrac{dy}{dt} = b\left(x,y\right) + \eta \end{cases} \tag{21}$$

where $0 < \alpha \ll 1$, and where η is a Gaussian random column vector with zero mean and correlations $\mathbb{E}\left[\eta\left(t_1\right)\eta^T\left(t_2\right)\right] = C\delta\left(t_1 - t_2\right)$ with the correlation matrix C. In the case we are interested in, the random vector y is actually the eddy vorticity field, and x is the zonal jet vorticity or velocity. We use vectorial notations $x = \left(x_\ell\right)_{1\leq\ell\leq L}$ for simplicity in this section, the formal generalization to the field case is straightforward, see sections V and VI.

In (21), the variable x typically evolves on a time scale of order $1/\alpha$, while y evolves on a time scale of order 1. In the time scale separation regime between zonal jets and eddies defined by $\alpha \ll 1$, the quasi-linear barotropic equation (14) is a particular case of the system (21). Note however that in that case, dissipation terms of order α are present in $b(x,y)$. The general results presented in this section usually do not take into account such terms[20,30,31]. As a consequence, in sections V and VI we make sure that our results do not depend on the dissipative terms in the limit $\alpha \to 0$.

The goal of stochastic averaging is to give an effective description of the dynamics of x over time scales of order $1/\alpha$, where the effect of the fast process y is averaged out. The effective dynamics describes the attractors of x, the relaxation dynamics towards these attractors and the small fluctuations around these attractors, in the regime $\alpha \ll 1$. For quasi-geostrophic zonal jets dynamics, stochastic averaging leads to a kinetic description of zonal jets[12], related to statistical closures of the dynamics (S3T[26] and CE2[21,28,33]).

The effective dynamics obtained through stochastic averaging or statistical closures is not able to describe arbitrarily large fluctuations of the slow process x. Such rare events are of major importance in the long-term dynamics of x. For instance in the case where the system (21) has several attractors, transitions between the attractors are governed by large fluctuations of the system. The description of such transitions (transition probability, typical transition path) cannot be done through a stochastic averaging procedure.

Large deviation theory is a natural framework to describe large fluctuations of x in the regime $\alpha \to 0$. The large deviation principle[30] gives the asymptotic form of the probability density of paths $\left\{x(t)\right\}_{0\leq t\leq T}$ when $\alpha \ll 1$, with the effect of the fast process y averaged

out. Information about the typical effective dynamics of x as obtained through stochastic averaging is captured, but the principle allows us to go further to describe arbitrarily rare events. In cases of multistability of x, the Large Deviation Principle yields the asymptotic expression of the transition probability from one attractor to another, the average relative residence time in each attractor, and the typical transition path $\{x(t)\}_{0\leq t\leq T}$ that links two attractors in a given time $T \gtrsim 1/\alpha$, among other relevant statistical quantities. Implementing the large deviation principle in practice for systems like (21) and for the quasilinear dynamics is one of the goals of this work.

In the effective descriptions of x provided by stochastic averaging and the Large Deviation Principle, the dynamics of y is approximated by its stationary dynamics with x held fixed, the so-called virtual fast process. The mathematics is described in section IV A. The effective dynamics of x over time scales $t \gg 1$ provided by stochastic averaging is then presented in section IV B. The Large Deviation Principle for (21) is stated in section IV C, and in section IV D 2 we give a method to estimate the quantities involved in the Large Deviation Principle from simulations of the virtual fast process.

A. The virtual fast process

In slow-fast systems like (21), the time scale separation implies that at leading order, the statistics of y are very close to the stationary statistics of the virtual fast process $\tilde{y}(u)$

$$\frac{d\tilde{y}}{du} = b\left(x, \tilde{y}(u)\right) + \eta(u), \tag{22}$$

where x is held fixed[20,30]. The time scale separation hypothesis is relevant only when the fast process described by (22) is stable (for instance has an invariant measure and is ergodic). The stationary process (22) depends parametrically on x, and the expectation over the invariant measure of (22) is thus denoted $\mathbb{E}_x$. The statistics of $\tilde{y}$ then change when x evolves adiabatically on longer timescales.

For quasilinear barotropic dynamics (14), the virtual fast process is the linearized barotropic equation close to the fixed stable zonal flow U (15) (the necessity for U to be stable for the quasilinear hypothesis to be correct was emphasized in reference[12].)

The process (22) is relevant only if a time scale separation effectively exists between the evolutions of x and y. In practice, the time scale separation hypothesis in (21) can be considered to be self-consistent if the typical time scale of evolution of the virtual fast process (22) is of order one, while the slow variable evolves on a time scale of order $1/\alpha$. From the point of view of the interaction with the dynamics of x, the most relevant time scales related to the evolution of $\tilde{y}(u)$ are the correlation times of processes $f_\ell\left(x,\tilde{y}(u)\right)$ and $f_{\ell'}\left(x,\tilde{y}(u)\right)$, defined as[34,35]

$$\tau_{\ell,\ell'}=\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_0^t\frac{\mathbb{E}_x\left[\left[f_\ell\left(x,\tilde{y}\left(u_1\right)\right)f_{\ell'}\left(x,\tilde{y}\left(u_2\right)\right)\right]\right]}{2\mathbb{E}_x\left[\left[f_\ell\left(x,\tilde{y}\right)f_{\ell'}\left(x,\tilde{y}\right)\right]\right]}\mathrm{d}u_1\mathrm{d}u_2 \tag{23}$$

where $\mathbb{E}_x\left[\left[X_1\left(u_1\right)X_2\left(u_2\right)\right]\right]\equiv\mathbb{E}_x\left[X_1\left(u_1\right)X_2\left(u_2\right)\right]-\mathbb{E}_x\left[X_1\left(u_1\right)\right]\mathbb{E}_x\left[X_2\left(u_2\right)\right]$ is the covariance of X_1 at time u_1 and X_2 at time u_2. If $\ell=\ell'$, $\tau_{\ell,\ell}$ is called the auto-correlation time of the process $f_\ell\left(x,\tilde{y}(u)\right)$. In all these expressions, x is fixed and $\mathbb{E}_x$ is the average over realizations of the fast process (22) in its statistically stationary state. The correlation times $\{\tau_{\ell,\ell'}\}$ give an estimate of the time scales of evolution of the terms that force the slow process x in (21).

In the regime $\alpha\ll 1$, we can consider a time Δt much larger than the auto-correlation times $\tau_{\ell,\ell'}$ but much smaller than the typical time for the evolution of x itself: $\tau_{\ell,\ell'}\ll\Delta t\ll 1/\alpha$. Over such time scale, (21) can be integrated to give

$$x(t+\Delta t)=x(t)+\alpha\int_t^{t+\Delta t}f\left(x(u),y(u)\right)\mathrm{d}u\simeq x(t)+\alpha\int_t^{t+\Delta t}f\left(x(t),\tilde{y}(u)\right)\mathrm{d}u, \tag{24}$$

where in obtaining the last equality we have used the fact that over time Δt the process x has almost not evolved. The relation (24) is used in the following to derive equations for the average behaviour, typical fluctuations and large fluctuations of x, in the time scale separation limit $\alpha\ll 1$.

B. Average evolution and energy balance for the slow process

We now describe the typical dynamics of x over time scales Δt such that $\tau_{\ell,\ell'}\ll\Delta t\ll 1/\alpha$, recovering classical results from stochastic averaging[20]. Because the time Δt in (24) is much larger than the typical correlation time of the components of $f\left(x,\tilde{y}(u)\right)$, by the Law of Large Numbers we can replace the time average by a statistical average: $\frac{1}{\Delta t}\int_t^{t+\Delta t}f\left(x,\tilde{y}(u)\right)\mathrm{d}u\simeq F(x)$ where $F(x)\equiv\mathbb{E}_x\left[f\left(x,\tilde{y}(u)\right)\right]$ is the average force acting on x, computed in the statistically stationary state of the virtual fast process (22). Then, the average evolution of x at

leading order in $\alpha\Delta t \ll 1$ is

$$\frac{\Delta x}{\Delta t} \equiv \frac{x(t+\Delta t) - x(t)}{\Delta t} \simeq \alpha F(x(t)). \tag{25}$$

In the case of zonal jet dynamics in barotropic models, x is the zonally averaged vorticity (or velocity) and $F(x)$ is the average advection term R. The effective dynamics (25) is then very close to S3T-CE2 types of closures[11,21,26–28] or to kinetic theory[12]. This point is further discussed in section V.

The effective dynamics (25) is not enough to describe the effective energy balance related to the slow process x. Indeed, replacing the time averaged force in (24) by its statistical average amounts to neglecting fluctuations in the process $f(x, \tilde{y}(u))$. The fluctuations are however relevant in the evolution of quadratic forms of x. In particular, if we define the energy of the slow degrees of freedom as $E = \frac{1}{2}x \cdot x = \sum_\ell E_\ell$ with $E_\ell = \frac{1}{2}x_\ell^2$, an equation for E_ℓ can be derived using (24),

$$\begin{aligned} E_\ell(t+\Delta t) \simeq E_\ell(t) + \alpha x_\ell(t) \int_t^{t+\Delta t} f_\ell\left(x(t), \tilde{y}(u)\right) \mathrm{d}u \\ + \frac{\alpha^2}{2} \int_t^{t+\Delta t} \int_t^{t+\Delta t} f_\ell\left(x(t), \tilde{y}\left(u_1\right)\right) f_\ell\left(x(t), \tilde{y}\left(u_2\right)\right) \mathrm{d}u_1 \mathrm{d}u_2. \end{aligned} \tag{26}$$

Define

$$Z_{\ell,\ell'}(x) \equiv \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_0^{\Delta t} \int_0^{\Delta t} \mathbb{E}_x \left[\left[f_\ell\left(x, \tilde{y}\left(u_1\right)\right) f_{\ell'}\left(x, \tilde{y}\left(u_2\right)\right)\right]\right] \mathrm{d}u_1 \mathrm{d}u_2 \,, \tag{27}$$

then using again that Δt is much larger than the correlation time of $f\left(x, \tilde{y}(u)\right)$ we get

$$\frac{\Delta E_\ell}{\Delta t} \simeq \alpha x_\ell F_\ell(x) + \frac{\alpha^2}{2} Z_{\ell,\ell}(x). \tag{28}$$

This relation is the energy balance for the slow evolution of x: $p_{mean,\ell} = \alpha x_\ell F_\ell(x)$ is the average energy injection rate by the mean force $F(x)$, and $p_{fluct,\ell} = \frac{\alpha^2}{2} Z_{\ell,\ell}(x)$ is the average energy injection rate by the typical fluctuations of the force f, as quantified by $Z(x)$. Neglecting the term $p_{fluct,\ell}$ in (28), we recover the energy balance we would have obtained by computing the evolution of E_ℓ from (25). This observation confirms the fact that fluctuations of f, which are not taken into account in (25), are relevant in the effective dynamics of x.

C. Large Deviation Principle for the slow process

1. *Large deviation rate function for the action of the fast variable on the slow variable*

Equations (25) and (28) give the evolution of x and $x \cdot x$ at leading order in $\alpha \ll 1$. Such effective evolution equations can also be found in a more formal way using stochastic averaging[20,30]. The effective equations only describe the low-order statistics of the slow process: The average evolution and typical fluctuations (variance or energy). In contrast, the Large Deviation Principle gives access to the statistics of both typical and rare events, also in the limit $\alpha \ll 1$. For system (21), the Large Deviation Principle was first proved by Freidlin (see Ref. 30 and references therein). It states that the probability density of a path of the slow process x, denoted $\mathcal{P}[x]$, satisfies[30]

$$\ln \mathcal{P}\left[x\right] \underset{\alpha \to 0}{\sim} -\frac{1}{\alpha} \int \mathcal{L}\left(x(t), \dot{x}(t)\right) \mathrm{d}t \tag{29}$$

with $\mathcal{L}\left(x, \dot{x}\right) \equiv \min_\theta \left\{\dot{x} \cdot \theta - H\left(x, \theta\right)\right\}$ and where $H\left(x, \theta\right)$ is the scaled cumulant generating function

$$H\left(x, \theta\right) \equiv \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \ln \mathbb{E}_x \left[\exp\left(\theta \cdot \int_0^{\Delta t} f\left(x, \tilde{y}(u)\right) \mathrm{d}u\right)\right], \tag{30}$$

where we recall that $\mathbb{E}_x$ is an average over realisations of the virtual fast process (22) in its statistically stationary state. Quantities H and $\mathcal{L}$ are classical definitions from Large Deviation Theory[30]. The knowledge of the function $H(x, \theta)$ is equivalent to the knowledge of $\mathcal{L}\left(x, \dot{x}\right)$, which gives the probability of any path of the slow process x through (29). Computing $H\left(x, \theta\right)$ is thus a very efficient way to study the effective statistics of $x(t)$, even when extremely rare events that are not described in the effective equations (25) and (28) play an important role.

Because the Large Deviation Principle (29) describes both rare events and typical events, information about the effective dynamics (25, 28) is encoded in the definition of the scaled cumulant generating function. Indeed, a Taylor expansion in powers of θ in (30) gives

$$H\left(x, \theta\right) = \sum_\ell \theta_\ell F_\ell(x) + \frac{1}{2} \sum_{\ell, \ell'} \theta_\ell \theta_{\ell'} Z_{\ell, \ell'}(x) + O\left(\theta^3\right), \tag{31}$$

with $F(x) \equiv \mathbb{E}_x\left[f\left(x, \tilde{y}(u)\right)\right]$ and Z given by (27). The terms appearing in the leading order evolution of x (25) and of the energy (28) are thus contained in the scaled cumulant generating function, through (31).

Higher-order terms in (31) involve cubic and higher-order cumulants of large time averages of the process $f(x, \tilde{y}(u))$. If this process is a Gaussian process, then its statistics are only given by its first and second order cumulants[20]. As a consequence, for such process $H(x, \theta)$ is quadratic in θ and (31) is exact (corrections of order θ^3 are exactly zero).

In practice, the scaled cumulant generating function (30) involves the virtual fast process (22). This stochastic process depends only parametrically on x, which means that we do not have to study the coupled system (21) in order to compute $H(x, \theta)$. This result is consistent with the time scale separation property of (21). In quasi-linear systems such as the quasi-linear barotropic dynamics, the virtual fast process is an Ornstein-Uhlenbeck process, which is particularly simple to study. This specific class of systems is considered next in section IV C 2.

2. *Quasi-linear systems with action of the fast process on the slow one through a quadratic force: the matrix Riccati equation*

We are particularly interested in the more specific class of systems defined by

$$\begin{cases} \dfrac{dx}{dt} = \alpha y^T \mathcal{M} y + \alpha g(x) \\ \dfrac{dy}{dt} = -L_x[y] + \eta \end{cases} \tag{32}$$

where $\mathcal{M}$ is a symmetric matrix, and L_x is a linear operator acting on y that depends parametrically on x. The system (32) is a particular case of (21) with $f(x, y) = y^T \mathcal{M} y + g(x)$ and $b(x, y) = -L_x[y]$.

When x is the zonal flow vorticity profile and y is the eddy vorticity, the quasi-linear barotropic dynamics (14) is an example of such a system, where the quadratic form $y^T \mathcal{M} y$ defines the zonally averaged advection term R and $g(x)$ contains the dissipative terms acting on the large-scale zonal flow x, and where L_x is the linearized barotropic operator close to the zonal flow x (see also section VI).

We now describe the effective dynamics and large deviations of x in the system (32), in the limit $\alpha \to 0$. In this limit, the statistics of y are very close to the statistics of the virtual

fast process (22), which in this case reads

$$\frac{d\tilde{y}}{dt} = -L_x\left[\tilde{y}\right] + \eta, \tag{33}$$

where x is frozen. Equation (33) describes an Ornstein-Uhlenbeck process, which stationary distribution is Gaussian[20]. Then, the stationary statistics of (33) are fully determined by the mean and covariance of $\tilde{y}$. The mean is zero, and the covariance $G_{ij} = \mathbb{E}\left[\tilde{y}_i\tilde{y}_j\right]$ is given by the Lyapunov equation

$$\frac{dG}{dt} + L_x G + G L_x^T = C. \tag{34}$$

The Lyapunov equation (34) converges to a unique stationary solution whenever (33) has an invariant measure. We recall that such an invariant measure is required for the time scale separation hypothesis to be relevant. The effective dynamics of x over times $\Delta t \ll 1/\alpha$ is given by (25). In the case of (32), it reads

$$\frac{\Delta x}{\Delta t} \simeq \alpha\left[\mathcal{M}\cdot G_\infty(x) + g(x)\right] \tag{35}$$

with $\mathcal{M}\cdot G_\infty(x) = \sum_{i,j}\mathcal{M}_{ij}\left(G_\infty\right)_{ij}(x)$ where G_∞ is the stationary solution of the Lyapunov equation (34). Simulating the effective slow dynamics (35) can be done by integrating the Lyapunov equation (34), using standard methods[36]. It provides an effective description of the attractors of x, and of the relaxation dynamics towards the attractors. Examples of such numerical simulations of (35) in the case of zonal jet dynamics in the barotropic model can be found for instance in Refs. 11, 21, 26, 28, and 33.

In order to describe large fluctuations of x in (32), we need to use the Large Deviation Principle (29). In practice, we compute the scaled cumulant generating function (30). As proven in Ref. 32, for the system (32), the scaled cumulant generating function is given by

$$H\left(x,\theta\right) = \theta\cdot g(x) + \mathrm{tr}\left(CN_\infty\left(x,\theta\right)\right) \tag{36}$$

where C is the covariance matrix of the noise η in (32) and $N_\infty\left(x,\theta\right)$ is a symmetric matrix, stationary solution of

$$\frac{dN}{dt} + NL_x + L_x^T N = 2NCN + \theta\mathcal{M}. \tag{37}$$

Equation (37) is a particular case of a matrix Ricatti equation, and in the following we refer to (37) as the Ricatti equation. θ is the parameter of the cumulant generating function

(30) that defines H. Whenever θ is in the range of parameter for which the limit in (30) exists, called the admissible θ range, Eq. (37) has a stationary solution. For the case in this section, with a linear dynamics with a quadratic observable, the admissible θ range is easily studied through the analysis of the positivity of a quadratic form. One can conclude that the admissible θ range is an interval containing 0. All the information regarding the large deviation rate function is contained in the values of H for θ in this range.

The Ricatti equation (37) is similar to the Lyapunov equation (34), and it can be solved using similar methods[37]. Moreover, the numerical implementation of (36, 37) can be easily checked using the relation with the Lyapunov equation (34). Namely, (31) implies that

$$\left.\frac{dH}{d\theta}\right|_{\theta=0} = \mathcal{M} \cdot G_\infty(x) + g(x).$$

The first term in the right-hand side is computed from the Lyapunov equation (34), while the left-hand side is computed from the Ricatti equation (37) together with (36).

In section VI, we present a numerical resolution of (37) for the case of the quasi-linear barotropic equation on the sphere, and then compute directly the scaled cumulant generating function using (36). We show that (37) can be very easily solved for a given value of θ. This means that the result (36) permits the study of arbitrarily rare events in zonal jet dynamics extremely easily, through the Large Deviation Principle (29). Such result is in clear contrast with approaches through direct numerical simulations, which require that the total time of integration increases as the probability of the event of interest decreases. This limitation of direct numerical simulations in the study of rare events statistics is made more precise in next section.

D. Estimation of the large deviation function from time series analysis

In this section we present a way to compute the scaled cumulant generating function (30) from a time series of the virtual fast process (22), for instance one obtained from a direct numerical simulation. Many of the technical aspects of this empirical approach follow Ref. 5.

Consider a time series $\{\tilde{y}(u)\}_{0\leq u\leq T}$ of the virtual fast process (22), with a given total time window $u \in [0, T]$. Because the quantities of interest like $H(x, \theta)$ involve expectations in the stationary state of the virtual fast process, we assume that the time series $\{\tilde{y}(u)\}_{0\leq u\leq T}$ corresponds to this stationary state. We use the continuous time series notation for simplicity, the

generalization of the following formulas to the case of discrete time series is straightforward. We also denote for simplicity $R(u) \equiv f\left(\tilde{y}(u)\right)$ the quantity for which the scale cumulant generating function $H\left(\theta\right) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E} \exp\left(\theta \int_0^t R(u)\,\mathrm{d}u\right)$ should be estimated.

The basic idea to estimate the scaled cumulant generating function (30) is to divide the full time series $\{\tilde{y}(u)\}_{0\leq u\leq T}$ into blocks of length Δt, to compute the integrals $\int_{t_0}^{t_0+\Delta t} R(u)\,\mathrm{d}u$ over those blocks, and then to average the quantity $\exp\left(\theta \cdot \int_{t_0}^{t_0+\Delta t} R(u)\,\mathrm{d}u\right)$. For a small block length Δt, the large-time regime defined by the limit $\Delta t \to \infty$ in the theoretical expression of H (30) is not attained. On the other hand, too large values of Δt —typically of the order of the total time T— lead to a low number of blocks, and thus to a very poor estimation of the empirical mean. Estimating H thus requires finding an intermediate regime for Δt. More precisely, we expect this regime to be attained for Δt equal to a few times the correlation time of the process $R(u)$, defined by[34,35]

$$\tau \equiv \lim_{\Delta t\to\infty} \frac{\int_0^{\Delta t}\int_0^{\Delta t} \mathbb{E}_z\left[\left[\, R(u_1)R(u_2)\,\right]\right]\,\mathrm{d}u_1\mathrm{d}u_2}{2\Delta t\,\mathbb{E}_z\left[\left[\, R^2\,\right]\right]} = \frac{\int_0^{\infty} \mathbb{E}_z\left[\left[\, R(u)R(0)\,\right]\right]\,\mathrm{d}u}{\mathbb{E}_z\left[\left[\, R^2\,\right]\right]}\,, \tag{38}$$

where $\mathbb{E}_z[[R(u_1)R(u_2)]]$ is the covariance of R at time u_1 and at time u_2. The second equality is easily obtained assuming that the process $R(u)$ is stationary. Because of the infinite-time limit in (38), the estimation of τ suffers from the same finite sampling problem as the estimation of H.

Finding a block length Δt such that the estimation of H and τ is accurate is thus a tricky problem. In the following, we propose a method to find the optimal Δt and estimate the quantities we are interested in. The proposed method is close to the "data bunching" method used to estimate errors in Monte Carlo simulations[38].

1. Estimation of the correlation time

We first consider the problem of the estimation of τ in a simple solvable case, so the numerical results can be compared directly to explicit formulas. Consider the stochastic process $R = w^2$ where w is the one-dimensional Ornstein-Uhlenbeck process

$$\frac{dw}{dt} = -w + \eta, \tag{39}$$

where η is a Gaussian white noise with correlation $\mathbb{E}\left(\eta(t)\eta(t')\right) = \delta(t-t')$. A direct calculation gives the correlation time of R, $\tau = 1/2$. Using (36) and (37), the scaled cumulant

generating function can also be computed explicitly (see for instance Ref. 32). We obtain

$$H(\theta) = \frac{1}{2} - \frac{1}{2}\sqrt{1-2\theta}, \tag{40}$$

defined for $\theta \leq 1/2$.

For a time series $\{R(u)\}_{0\leq u\leq T}$, we denote $\bar{R}_T = \frac{1}{T}\int_0^T R(u)\,\mathrm{d}u$ and $\mathrm{var}_T(R) = \frac{1}{T}\int_0^T \left(R(u) - \bar{R}_T\right)^2 \mathrm{d}u$ respectively the empirical mean and variance of R over the full time series. We then estimate the correlation time τ defined in (38) using an average over blocks of length Δt,

$$\tau_{\Delta t} = \frac{1}{2\Delta t\,\mathrm{var}_T(R)}\mathbb{E}_{\frac{T}{\Delta t}}\left[\left(\int_{t_0}^{t_0+\Delta t}\left(R(u)-\bar{R}_T\right)\,\mathrm{d}u\right)^2\right], \tag{41}$$

where $\mathbb{E}_{\frac{T}{\Delta t}}\left[h_{t_0}\right]$ is the empirical average over realisations of the quantity h_{t_0} inside the brackets[39].

To find the optimal value of Δt, we plot $\tau_{\Delta t}$ as a function of Δt, see Figure 5. For small values of Δt, the large-time limit in (38) is not achieved, which explains the low values of $\tau_{\Delta t}$. For too large values of Δt, the empirical average $\mathbb{E}_{\frac{T}{\Delta t}}$ in (41) is not accurate due to the small value of $\frac{T}{\Delta t}$ (small number of blocks), which explains the increasing fluctuations in $\tau_{\Delta t}$ as Δt increases. The optimal value of Δt —denoted $\Delta t^\star$ in the following— is the one in between these artificial behaviours. It should satisfy $T \gg \Delta t^\star \gg \tau_{\Delta t^\star}$. Here, one can read $\Delta t^\star \simeq 10$ and $\tau_{\Delta t^\star} \simeq 0.5$, so this optimal $\Delta t^\star$ satifies the aforementioned condition. The estimated value $\tau_{\Delta t^\star}$ is in agreement with the theoretical value $\tau = 1/2$.

The error bars for $\tau_{\Delta t}$ are given by $\Delta\tau_{\Delta t} = \sqrt{\mathrm{var}\left(\tau_{\Delta t}\right)/N_{terms}}$, where $\mathrm{var}\left(\tau_{\Delta t}\right)$ is the empirical variance associated with the average $\mathbb{E}_{\frac{T}{\Delta t}}$ defined in (59), and N_{terms} is the number of terms in this sum (roughly $N_{terms} \simeq 2T/\Delta t$).

2. *Estimation of the scaled cumulant generating function*

The self-consistent estimation of the correlation time τ presented in the previous section gives the optimal value $\Delta t^\star$ of the block length. Then, the scaled cumulant generating function is computed for a given value of θ as

$$H_T\left(\theta\right) \equiv \frac{1}{\Delta t^\star}\ln\mathbb{E}_{\frac{T}{\Delta t^\star}}\left[\exp\left(\theta\int_{t_0}^{t_0+\Delta t^\star} R(u)\,\mathrm{d}u\right)\right], \tag{42}$$

where $\mathbb{E}_{\frac{T}{\Delta t}}$ is the empirical average over the blocks, as defined in (59). However, the knowledge of $H\left(x,\theta\right)$ for an arbitrarily large value of $|\theta|$ leads to the probability of an arbitrarily

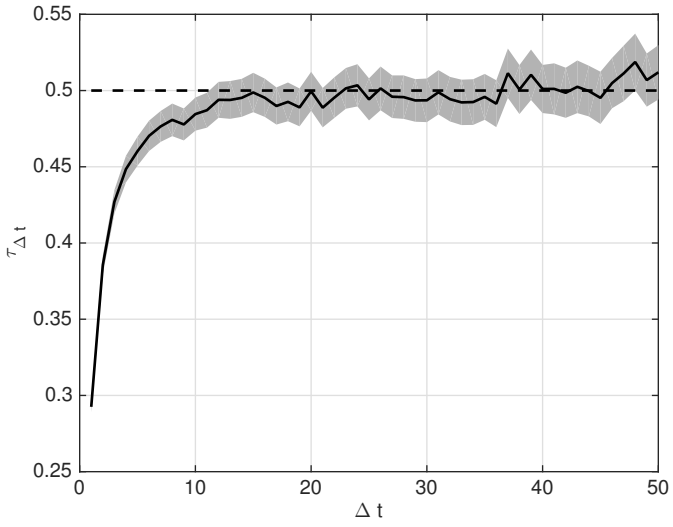


FIG. 5. Plot of the estimated correlation time $\tau_{\Delta t}$ (black line) and error bars (grey shading) as functions of Δt. For small values of Δt, the large-time limit in (38) is not achieved, which explains the low values of $\tau_{\Delta t}$. For too large values of Δt, the empirical average $\mathbb{E}_{\frac{T}{\Delta t}}$ in (41) is not accurate due to the small value of $\frac{T}{\Delta t}$, which explains the increasing fluctuations in $\tau_{\Delta t}$ as Δt increases. The optimal value $\Delta t^\star$ is the one in between these artificial behaviour. Here, one can read $\Delta t^\star \simeq 20$ and $\tau_{\Delta t^\star} \simeq 0.5$, in agreement with the exact value $\tau = 1/2$ (dashed line). The Ornstein-Uhlenbeck process (39) has been integrated over $T = 5.10^4$ using the method proposed in Ref. 40, with time step 10^{-3}.

rare event for the slow process x through the Large Deviation Principle (29). This is in contradiction with the fact that the available time series $\{R(u)\}_{0\leq u\leq T}$ is finite. In other words, the range of values of θ for which the scaled cumulant generating function $H_T(\theta)$ can be computed with accuracy depends on T.

Indeed, for large positive values of θ, the sum $\mathbb{E}_{\frac{T}{\Delta t^\star}}$ in (42) is dominated by the largest term $\exp(\theta I_{max})$ where $I_{max} = \max_{t_0}\left\{\int_{t_0}^{t_0+\Delta t} R(u)\,\mathrm{d}u\right\}$ is the largest value of $\int_{t_0}^{t_0+\Delta t} R(u)\,\mathrm{d}u$ over the finite sample $\{R(u)\}_{0\leq u\leq T}$. Then $H_T(\theta) \sim \frac{1}{\Delta t^\star} I_{max}\theta$ for $\theta \gg 1$. This phenomenon is known as linearization[5], and is clearly an artifact of the finite sample size. We denote θ_{max} the value of θ such that linearization occurs for $\theta > \theta_{max}$. Typically, we expect θ_{max} to be a positive increasing function of T. The same way, $H_T(\theta) \sim -\frac{1}{\Delta t^\star} I_{min}\theta$ for $\theta < 0$ and $|\theta| \gg 1$, with $I_{min} = \min_{t_0}\left\{\int_{t_0}^{t_0+\Delta t} R(u)\,\mathrm{d}u\right\}$. We thus define similarly θ_{min} such that linearization

occurs for $\theta < \theta_{min}$. Typically, we expect θ_{min} to be a negative decreasing function of T.

The convergence of estimators like (42) is studied in Ref. 5, in particular it is shown that error bars can be computed in the range $[\theta_{min}/2, \theta_{max}/2]$ for a given time series $\{R(u)\}_{0\leq u\leq T}$. An example of computation of $H_T(\theta)$ is shown in Figure 6 for the one-dimensional Ornstein-Uhlenbeck process, and compared to the explicit solution. The full error bars in Figure 6 are given by the error from the estimation of τ and the statistical error described in Ref. 5. The method shows excellent agreement with theory, and exposes non-Gaussian behavior.

In sections V and VI, we apply the tools (estimation of the correlation time and of the scaled cumulant generating function) to the study of Reynolds' stresses statistics in zonal jet dynamics.

V. ZONAL ENERGY BALANCE AND TIME SCALE SEPARATION IN THE INERTIAL LIMIT

In this section we discuss the effective evolution and effective energy balance for zonal flows in the inertial regime $\nu_n \ll \alpha \ll 1$, using the general results of section IV B and numerical simulations.

A. Effective dynamics and energy balance for the zonal flow

Using (13) and (25), the effective evolution of the zonal jet velocity profile $U(\phi, t)$ in the regime $\nu_n \ll \alpha \ll 1$ reads

$$\frac{\partial U}{\partial t} \simeq \alpha F[U] - \alpha U - \nu_n(-\Delta)^n U, \tag{43}$$

with $F[U] \equiv \mathbb{E}_U[f]$ where f is minus the Reynolds' stress divergence and $\mathbb{E}_U$ is the average in the statistically stationary state of the linear barotropic dynamics (15), with U held fixed.

Equation (43) describes the effective slow dynamics of zonal jets in the regime $\nu_n \ll \alpha \ll 1$, it is the analogous of the kinetic equation proposed in Ref. 12. In particular, the attractors of (43) are the same as the attractors of a second order closure of the barotropic dynamics[27,41].

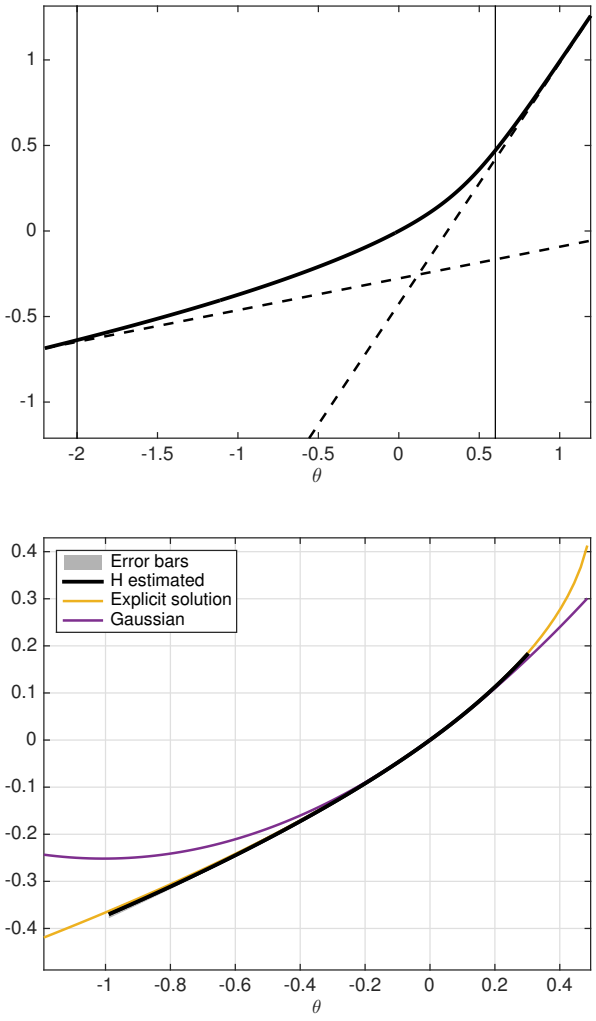


FIG. 6. Computation of the scaled cumulant generating function from (42) for the one-dimensional Ornstein-Uhlenbeck process (39). Upper panel: illustration of the linearization effect for large values of $|\theta|$. The solid curve is the estimated scaled cumulant generating function H_T, and the dashed lines are the expected linear tails, which are artifacts of the finite sample size[5]. The thin vertical lines show the range $\theta \in [\theta_{min}, \theta_{max}]$ for which we consider that linearization does not take place. Bottom pannel: the converged scaled cumulant generating function estimator H_T on $\theta \in [\theta_{min}/2, \theta_{max}/2]$ (thick black curve, with error bars in grey shading). The yellow curve is the exact scaled cumulant generating function (40), it fits the estimated one within statistical errors. The purple curve is the quadratic approximation, that corresponds to a Gaussian process $R(u)$ (see equation (31)). This quadratic approximation is computed using the exact mean, variance and correlation time of R. The Ornstein-Uhlenbeck process (39) has been integrated over $T = 5 \times 10^4$ using the method proposed in Ref. 40, with time step 10^{-3}.

As explained in a general setting in section IV B, equation (43) only takes into account the average Reynolds' stresses (through the term $F[U]$), as a consequence it does not describe accurately the effective zonal energy balance. Quantifying the influence of fluctuations of Reynolds' stresses on the zonal energy balance is one of the goals of this study. We now derive the effective zonal energy balance, and then describe the relative influence of average and fluctuations of Reynolds' stresses using numerical simulations.

First note that the hyperviscous terms in (13) essentially dissipate energy at the smallest scales of the flow. In the turbulent regime we are interested in, such small-scale dissipation is negligible in the global energy balance. For this reason, the viscous terms can be neglected in (43) and in the zonal energy balance. Note however that some hyper-viscosity is still present in the numerical simulations of the linear barotropic equation (15), in order to ensure numerical stability. For consistency, we make sure that the hyper-viscous terms do not influence the numerical results, see Figure 7.

The kinetic energy contained in zonal degrees of freedom reads $E_z = \int \mathrm{d}\phi\, E(\phi)$ with $E(\phi) = \pi \cos\phi\, U^2(\phi)$. Using (28) we get the equation for the effective evolution of $E(\phi)$,

$$\frac{1}{\alpha}\frac{dE}{dt} = p_{mean}(\phi) - 2E + \alpha p_{fluct}(\phi)\,, \tag{44}$$

with the instantaneous energy injection rates into the zonal mean flow, respectively by the average Reynolds' stresses $p_{mean}(\phi) \equiv 2\pi \cos\phi\, F[U](\phi)\, U(\phi)$ and by the fluctuations of Reynolds' stresses $\alpha p_{fluct}(\phi) \equiv \alpha\pi \cos\phi\, Z[U](\phi)$, where

$$Z[U](\phi) \equiv \lim_{\Delta t\to\infty} \frac{1}{\Delta t}\int_0^{\Delta t}\int_0^{\Delta t} \mathbb{E}_U\left[\left[f(\phi,u_1)\, f(\phi,u_2)\right]\right] \mathrm{d}u_1 \mathrm{d}u_2\,, \tag{45}$$

Integrating (44) over latitudes, we obtain the total zonal energy balance

$$\frac{1}{\alpha}\frac{dE_z}{dt} = P_{mean} - 2E_z + \alpha P_{fluct}, \tag{46}$$

with $P_{mean} \equiv \int \mathrm{d}\phi\, p_{mean}(\phi)$ and $\alpha P_{fluct} \equiv \int \mathrm{d}\phi\, \alpha p_{fluct}(\phi)$.

All the terms appearing in (44) and (46) can be easily estimated using data from a direct numerical simulation of the linearized barotropic equation (15). Indeed, $F[U](\phi)$ can be computed as the empirical average of $f(\phi)$ in the stationary state of (15), and $Z[U](\phi)$ can be computed using the method described in section IV D 1 to estimate correlation times[42].

Functions $F[U]$ and $Z[U](\phi)$ may be computed directly from the scaled cumulant generating function H, using (31). Computing H using the Ricatti equation (36, 37) and then

using (31), we have a very easy way to compute the terms appearing in the effective slow dynamics (43) or in the zonal energy balance equations (44) and (46), without having to simulate directly the fast process (15).

We now describe the results obtained by solving numerically the linearized barotropic equation (15), where we have used for the mean flow U the flow obtained from a quasilinear simulation as described in the end of section II B, and represented in Figure 1. The energy injection rates P_{mean} and αP_{fluct}, computed using both of the methods explained above, with different values of the non-dimensional damping rate α are represented in Figure 7. The first term P_{mean} (solid curve) is roughly of the order of magnitude of the dissipation term in (46) (recall we use units such that $E_z \simeq 1$). The second term αP_{fluct} is about an order of magnitude smaller than P_{mean}. In this case, the energy balance (46) implies that the zonal velocity is actually slowly decelerating.

Here, neglecting αP_{fluct} in (46) leads to an error in the zonal energy budget of about 5–10%. This confirms the fact that fluctuations of Reynolds' stresses are only negligible in a first approximation, but that they should be taken into account in order to obtain a quantitative description of zonal jet evolution. However, we emphasize that here only one mode is stochastically forced (see section II B for details). When several modes are forced independently, the Reynolds' stress divergence $f(\phi)$ is computed as the sum of independent contributions due to each mode. If the number K of forced modes becomes large, then the Central Limit Theorem implies that the typical fluctuations of $f(\phi)$ (and thus αP_{fluct}) roughly scale as $1/K$. In Figure 7, $K = 1$ so we are basically considering the case where fluctuations of Reynolds' stresses are the most important in the zonal energy balance. In other words, this is the worst case test for CE2 types of closures. In most previous studies of second order closures like CE2, a large number of modes is forced[28,43], so in these cases $p_{fluct}(\phi)$ and αP_{fluct} are most likely to be negligible in the zonal energy balance.

We also observe that P_{mean} increases up to a finite value as $\alpha \ll 1$, while αP_{fluct} is nearly constant over the range of values of α considered. We further comment the behavior in the following.

The spatial distribution of the energy injection rates $p_{mean}(\phi)$ and $p_{fluct}(\phi)$ are represented in Figures 8 and 9(a), 9(b). Both $p_{mean}(\phi)$ and $p_{fluct}(\phi)$ are concentrated in the jet region

$\phi \in [-\pi/4, \pi/4]$, which is also the region where the stochastic forces act (see Figure 1).

In Figure 9(a), we observe that p_{mean} is always positive. This means that the turbulent perturbations are everywhere injecting energy into the zonal degrees of freedom, i.e. the average Reynolds' stresses are intensifying the zonal flow $U(\phi)$ at each latitude. This effect is predominant at the jet maximum and around the jet minima (around $\phi = \pm\pi/8$). We also observe that p_{mean} (and thus $F[U]$) converges to a finite value as α decreases. A similar result has been obtained for the two dimensional Navier-Stokes equation under the assumption that the base flow has no mode, using theoretical arguments[12]. Those assumptions are not satisfied here, thus indicating that the finite limit of $F[U]$ as α vanishes is a more general result. This result is extremely important, indeed it implies that the effective dynamics (43) is actually well-posed in the limit $\alpha \to 0$.

By definition, $p_{fluct}(\phi)$ is necessarily positive. In Figure 9(b), we see that $p_{fluct}(\phi)$ keeps increasing as α decreases in the region away from the jet maximum (roughly for $|\phi| \in [\pi/16, \pi/4]$). This is in contrast with the behaviour of $p_{mean}(\phi)$ (fig. 9(a)). We note that such a behaviour for $p_{fluct}(\phi)$ has been obtained recently for the two-dimensional Navier-Stokes equation under the assumption that the base flow has no mode[18]. However, the range of values of α considered here is not wide enough to check precisely those theoretical results.

We also observe in Figure 9(b) that $p_{fluct}(\phi)$ is relatively small in the region of jet maximum $\phi \simeq 0$. This means that Reynolds' stresses tend to fluctuate less in this area. In the context of the deterministic two-dimensional Euler equation linearized around a background shear flow, it is known that extrema of the background flow lead to a decay of the perturbation vorticity (depletion of the vorticity at the stationary streamline[44]). In a stochastic context, this implies that the perturbation vorticity $\delta\omega$ is expected to fluctuate less in the region of jet extrema, in qualitative agreement with what is observed in Figure 9(b).

B. Empirical validation of the time scale separation hypothesis

In this paper we assumed a large separation in time scales: the eddies $\delta\omega$ evolves much faster than the zonal flow U, permitting the quasilinear approximation. It has been shown in Ref. 12 and 17 that for the linearized dynamics close to a zonal jet U, the autocorrelation function of both the eddy velocity and the Reynolds stresses are finite in the limit $\alpha \to 0$, even if the dissipation vanishes in this limit. An effective dissipation then takes place,

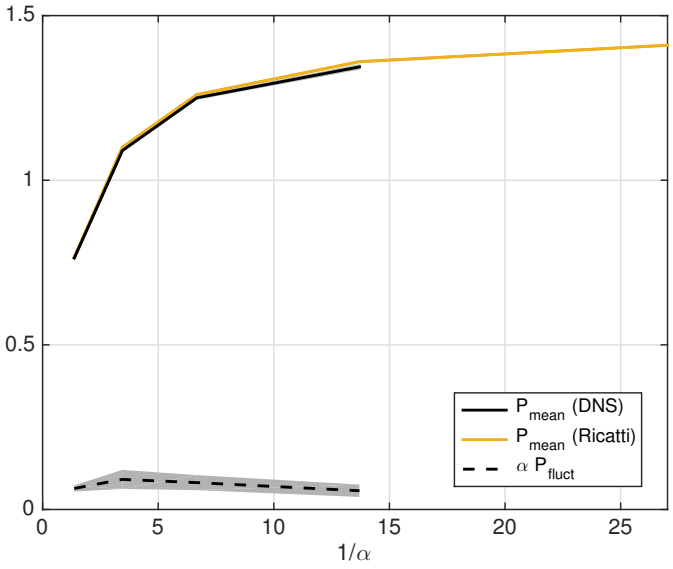


FIG. 7. Total energy injection rate into the zonal flow by the mean Reynolds' stresses P_{mean} (first term in the r.h.s of (46), in solid line) and by the fluctuations of Reynolds' stresses αP_{fluct} (last term in the r.h.s of (46), in dashed line with statistical error bars in grey shading) as a function of $1/\alpha$. The quantities are estimated from direct numerical simulations (DNS) of the linearized barotropic equation (52) with parameters given in section II B, and P_{mean} is also computed directly using the Ricatti equation (37) (yellow curve).

thanks to the Orr mechanism (see Refs. 12 and 17). This result then ensures that time scale separation assumption is valid for small enough α (the eddies $\delta\omega$ evolve on a time scale of order one, and the zonal flow U evolves on a time scale of order $1/\alpha$).

The consistency of this assumption for any value of α can also be tested numerically. For this purpose, we compute the maximum correlation time of the Reynolds' stress divergence $f(\phi)$, defined as[45]

$$\tau^{\alpha}_{max} \equiv \max_{\phi} \lim_{t\to\infty} \frac{1}{t} \int_0^t \int_0^t \frac{\mathbb{E}^{\alpha}_U \left[\left[f\left(\phi, s_1\right) f\left(\phi, s_2\right) \right] \right]}{2\mathbb{E}^{\alpha}_U \left[\left[f^2\left(\phi\right) \right] \right]} \mathrm{d}s_1 \mathrm{d}s_2. \tag{47}$$

We then check whether or not $\tau^{\alpha}_{max} \ll 1/\alpha$, where $1/\alpha$ is the dissipative time scale. The results are summarized in Figure 10. We observe the convergence of τ^{α}_{max} to a finite value as α decreases, as expected from the theoretical analysis[12,17], and this value is smaller than the inertial time scale (equal to one by definition of the time units). This means that the typical time scale of evolution of the Reynolds' stress divergence is much smaller than the

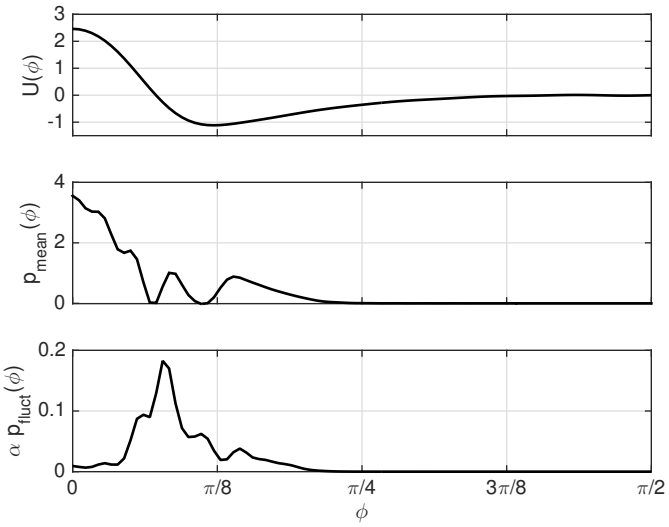
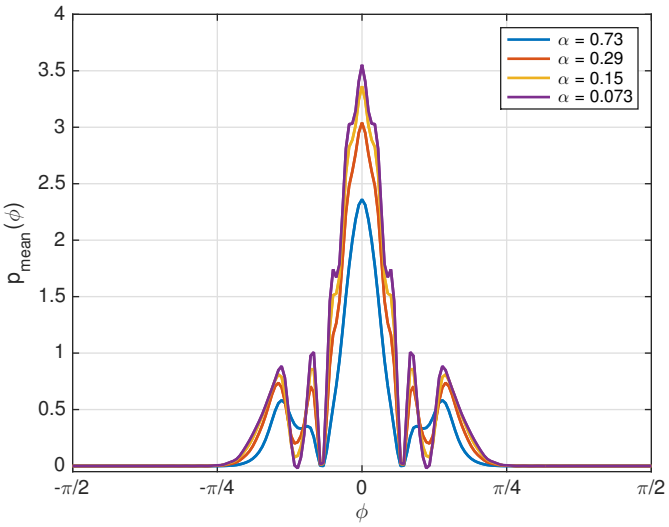


FIG. 8. From top to bottom: zonal velocity profile $U(\phi)$, energy injection rate by the average Reynolds' stresses $p_{mean}(\phi)$ and energy injection rate by the fluctuations of Reynolds' stresses $\alpha p_{fluct}(\phi)$, as functions of latitude ϕ restricted to the northern hemisphere. The values in the southern hemisphere are symmetric with respect to northern hemisphere, see Figures 1, 9(a) and 9(b). p_{mean} and p_{fluct} are estimated from numerical simulations of (52) with parameters given in section II B, and $\alpha = 0.073$. p_{mean} is always positive, meaning that the average Reynolds' stresses are intensifying the zonal flow $U(\phi)$ at each latitude. We see that fluctuations of Reynolds' stresses are lower at the jet extrema (p_{fluct} is relatively small), in particular close to the equator $\phi = 0$. This can be understood as a consequence of the depletion of vorticity at the stationary streamline[44]. Error bars are not shown here, see Figures 9(a) and 9(b).

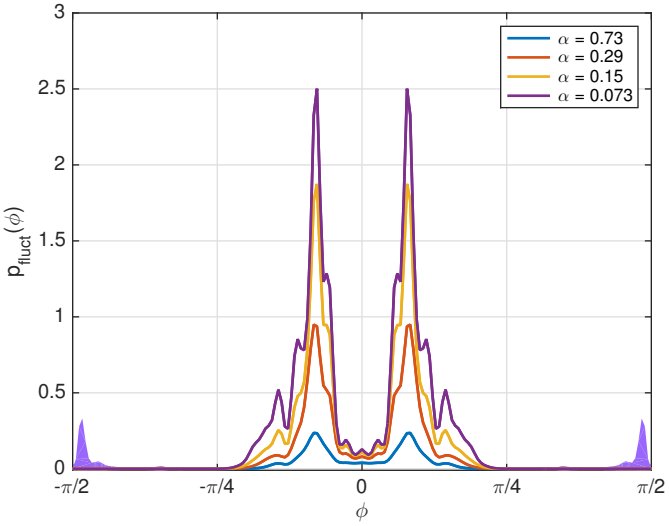
dissipative time scale $1/\alpha$ as soon as $1/\alpha$ is much larger than one, justifying the time scale separation hypothesis.

VI. LARGE DEVIATIONS OF REYNOLDS STRESSES

In section V, we studied the effective energy balance for the zonal flow $U(\phi)$ using numerical simulations of the linearized barotropic dynamics (15). This effective description of zonal jet dynamics takes into account the low-order statistics of Reynolds' stresses: average and covariance. In order to study rare events in zonal jet dynamics, we must employ the



(a)



(b)

FIG. 9. Energy injection rate into the zonal flow (a) by the mean Reynolds' stresses p_{mean} (first term in the r.h.s of (44)) and (b) by the fluctuations of Reynolds' stresses p_{fluct} (last term in the r.h.s of (44)), as functions of latitude ϕ, estimated from direct numerical simulations of the linearized barotropic equation (52) with parameters given in section II B, and with different values of the damping rate α. Shaded areas represent the statistical error bars. In Figure (a), we observe the convergence of p_{mean} to a finite function of ϕ as $\alpha \to 0$, in agreement with the theoretical predictions. In Figure (b), we observe that the values of p_{fluct} are relatively weak close the jet maximum $\phi = 0$, while they keep increasing as $\alpha \to 0$ in other locations, as expected from theory.

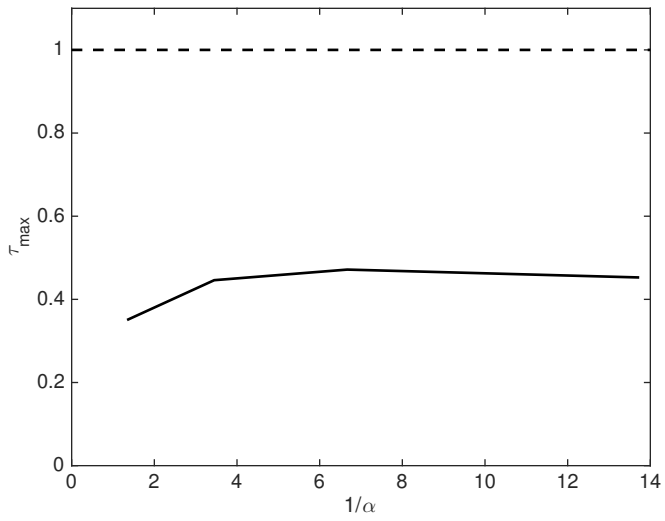


FIG. 10. Solid line: maximum correlation time of the Reynolds' stress divergence (47) as a function of the damping rate α. We clearly see the convergence of τ^{α}_{max} to a finite value as $\alpha \to 0$. The correlation time is of the order of the inertial time scale (equal to one by definition of the units, here represented by the dashed line), and much smaller than the dissipative time $1/\alpha$ (not represented here), showing the time scale separation between dissipative and inertial processes in the quasi-linear barotropic dynamics.

large deviation principle. The goal of this section is to apply the theoretical tools presented in sections IV C and IV D to the study of rare events statistics in zonal jet dynamics.

A. Large Deviation Principle for the time-averaged Reynolds' stresses

We first formulate the Large Deviation Principle for the quasi-linear barotropic equations (14) in the regime $\alpha \ll 1$, and present some properties of the large deviations functions. The numerical results are presented in section VI C. The Large Deviation Principle presented here is exactly equivalent to the one presented in a more general setting in section IV C.

Consider the evolution of ω_z from the first equation of (14). Over a time scale Δt much smaller than $1/\alpha$ but much larger than the correlation time τ we can write

$$\frac{\Delta\omega_z}{\Delta t} \equiv \frac{1}{\alpha}\frac{\omega_z(t+\Delta t)-\omega_z(t)}{\Delta t} \simeq \frac{1}{\Delta t}\int_t^{t+\Delta t} R(u)\,\mathrm{d}u - \omega_z(t)\,, \tag{48}$$

where we have used the fact that ω_z has not evolved much from t and $t+\Delta t$ (because

$\Delta t \ll 1/\alpha$), while $R(u)$ has evolved according to (15) with a fixed ω_z (or equivalently a fixed U). We also neglect hyper-viscosity in the evolution of ω_z, which is natural in the turbulent regime we are interested in. Note however that some hyper-viscosity is still present in the numerical simulations of (15), in order to ensure numerical stability. For consistency, we make sure that the hyper-viscous terms have no influence on the numerical results (see Figure 11).

We denote by $P_{\Delta t}\left[\frac{\Delta\omega_z}{\Delta t}\right]$ the probability distribution function of $\frac{\Delta\omega_z}{\Delta t}$, with a fixed t (and thus a fixed $\omega_z(t)$), but with an increasing Δt. This regime is consistent with the limit of time scale separation $\alpha \to 0$, where ω_z is nearly frozen while $\delta\omega$ keeps evolving. From (48), $P_{\Delta t}\left[\frac{\Delta\omega_z}{\Delta t}\right]$ is also the probability density function of the time-averaged advection term $\frac{1}{\Delta t}\int_t^{t+\Delta t} R(u)\,\mathrm{d}u$. The Large Deviation Principle gives the asymptotic expression of $P_{\Delta t}\left[\frac{\Delta\omega_z}{\Delta t}\right]$ in the regime $\Delta t \gg \tau$, namely

$$\ln P_{\Delta t}\left[\frac{\Delta\omega_z}{\Delta t}\right] \underset{\Delta t\to\infty}{\sim} -\Delta t\,\mathcal{L}\left[\frac{\Delta\omega_z}{\Delta t}\right]. \tag{49}$$

The function $\mathcal{L}$ is called the large deviation rate function. It characterizes the whole distribution of $\frac{\Delta\omega_z}{\Delta t}$ in the regime $\Delta t \gg \tau$, including the most probable value and the typical fluctuations.

Our goal in the following is to compute numerically $\mathcal{L}\left[\frac{\Delta\omega_z}{\Delta t}\right]$. This can be done through the scaled cumulant generating function (30). Using (48), the definition (30) can be reformulated as

$$H[\theta] = \lim_{\Delta t\to\infty}\frac{1}{\Delta t}\ln\int \mathrm{d}\dot{\omega}_z\, P_{\Delta t}\left[\dot{\omega}_z\right]\exp\left(\theta\cdot\Delta t\,\dot{\omega}_z\right) \tag{50}$$

Because ω_z is a field, here θ is also a field depending on the latitude ϕ, and H is a functional. For simplicity, we stop denoting the dependency in ω_z in H. In (50), we also have used the notation $\theta_1\cdot\theta_2 \equiv \int \mathrm{d}\phi\,\cos\phi\,\theta_1(\phi)\theta_2(\phi)$ for the canonical scalar product on the basis of spherical harmonics.

Using (49) in (50) and using a saddle-point approximation to evaluate the integral in the limit $\Delta t \to \infty$, we get $H[\theta] = \sup_{\dot{\omega}_z}\left\{\theta\cdot\dot{\omega}_z - \mathcal{L}\left[\dot{\omega}_z\right]\right\}$, i.e. H is the Legendre-Fenschel transform of $\mathcal{L}$. Assuming that H is everywhere differentiable, we can invert this relation as

$$\mathcal{L}\left[\frac{\Delta\omega_z}{\Delta t}\right] = \sup_{\theta}\left\{\theta\cdot\frac{\Delta\omega_z}{\Delta t} - H[\theta]\right\}. \tag{51}$$

The scaled cumulant generating function $H[\theta]$ can be computed either from a time series of $\delta\omega$ (see section IV D) or solving the Ricatti equation (see section IV C 2). Then the large deviation rate function $\mathcal{L}$ can be computed using (51), and this gives the whole probability distribution of $\frac{\Delta\omega_z}{\Delta t}$ (or equivalently of the time-averaged Reynolds' stresses) through the Large Deviation Principle (49).

In the following, we implement this program and discuss the physical consequences for zonal jet statistics. We first give a simpler expression of $H[\theta]$, that makes its numerical computation easier.

B. Decomposition of the scaled cumulant generating function

Using the Fourier decomposition (6), we can decompose the perturbation vorticity as $\delta\omega(\lambda,\phi) = \sum_m \omega_m(\phi)\mathrm{e}^{im\lambda}$, where ω_m satisfies

$$\frac{\partial \omega_m}{\partial u} = -L_{U,m}\left[\omega_m\right] + \sqrt{2}\eta_m, \tag{52}$$

where the Fourier transform of the linear operator (12) reads

$$L_{U,m}\left[\omega_m\right](\phi) = -\frac{im}{\cos\phi}\left(U(\phi)\omega_m(\phi) + \gamma(\phi)\psi_m(\phi)\right) - \alpha\omega_m(\phi) - \nu_n\left(-\Delta_m\right)^n\omega_m(\phi). \tag{53}$$

In (52), $\eta_m(\phi,t)$ is a Gaussian white noise such that $\eta_{-m} = \eta_m^*$, with zero mean and with correlations

$$\mathbb{E}\left[\eta_m(\phi_1,t_1)\,\eta_m^*(\phi_2,t_2)\right] = c_m(\phi_1,\phi_2)\,\delta(t_1-t_2),$$

$$\mathbb{E}\left[\eta_m(\phi_1,t_1)\,\eta_m(\phi_2,t_2)\right] = 0,$$

where c_m is the m-th coefficient in the Fourier decomposition of C in the zonal direction.

Using the Fourier decomposition, the zonally averaged advection term can be written $R(\phi) = \sum_m R_m(\phi)$ with $R_m(\phi) = -\frac{im}{\cos\phi}\partial_\phi\left(\psi_m\cdot\omega_{-m}\right)$. Using this expression and the fact that ω_{m_1} and $\omega_{m_2}^*$ are statistically independent for $m_1 \neq m_2$, the scaled cumulant generating function (50) can be decomposed as[46]

$$\begin{aligned} H[\theta] &\equiv \lim_{\Delta t\to\infty}\frac{1}{\Delta t}\ln\mathbb{E}_U\left[\exp\left(\theta\cdot\int_0^{\Delta t}\left(R(u)-\omega_z\right)\mathrm{d}u\right)\right] \\ &= -\theta\cdot\omega_z + \sum_m H_m\left[\theta\right], \end{aligned} \tag{54}$$

with

$$H_m\left[\theta\right] = \lim_{\Delta t\to\infty} \frac{1}{\Delta t} \log \mathbb{E}_U \exp\left[\int \mathrm{d}\phi\, \cos\phi\, \theta\left(\phi\right) \int_0^{\Delta t} R_m\left(\phi, u\right)\, \mathrm{d}u\right]. \tag{55}$$

We recall that $\mathbb{E}_U$ is the average in the statistically stationary state of (52).

In the following, we consider the case where only one Fourier mode m is forced, for simplicity and to highlight deviations from Gaussian statistics. If several modes are forced, their contibutions to the scaled cumulant generating function add up, according to (54).

Finally, consider the decomposition of the zonally averaged advection term into spherical harmonics (5), $R_m(\phi) = \sum_\ell R_{m,\ell}\, P_\ell^0(\sin\phi)$. Using $\theta(\phi) = \theta_\ell P_\ell^0(\sin\phi)$ in (55), we investigate the statistics of the ℓ-th coefficient $R_{m,\ell}$. The associated scaled cumulant generating function (55) is denoted $H_{m,\ell}\left(\theta\right) \equiv H_m\left[\theta P_\ell^0(\sin\phi)\right]$, and the large deviation rate function is denoted

$$\mathcal{L}_{m,\ell}\left(\dot\omega_\ell\right) = \sup_{\theta_\ell} \left\{\theta_\ell\, \dot\omega_\ell - H_{m,\ell}(\theta_\ell)\right\}. \tag{56}$$

C. Numerical results

The function $H_{m,\ell}$ defined in previous section can be computed either from a time series of $\omega_m(\phi, u)$ using the method described in section IV D, or solving the Ricatti equation as described in section IV C 2. Then, the large deviation rate funtion is computed using (56). We now show the results of these computations and discuss the physical consequences. We describe the results obtained by solving numerically the linearized barotropic equation (15), where we use the mean flow U the flow obtained from a quasilinear simulation as described in the end of section II B, and represented in Figure 1.

1. Scaled cumulant generating function

An example of computation of $H_{m,\ell}\left(\theta\right)$ is shown in Figure 11, with $m = 10$, $\ell = 3$ and $\alpha = 0.073$. The linearized barotropic equation (52) is integrated over a time $T_{max} = 54,500$, with fixed mean flow given in Figure 1, and the value of $R_{m,\ell}$ is recorded every 0.03 time units (the units are defined in section II A 1).

The scaled cumulant generating function (55) is then estimated following the procedure described in section IV D (thick black curve in Figure 11). Because the time series of $R_{m,\ell}$ is finite, $H_{m,\ell}(\theta)$ can only be computed with accuracy on a restricted range of values of θ

(see section IV D 2 for details), here $\theta \in [\theta_{min}/2, \theta_{max}/2] = [-0.6, 1.1]$.

The scaled cumulant generating function (55) is also computed solving numerically the Ricatti equation (37) and using (36) (yellow curve in Figure 11). We observe near perfect agreement between the direct estimation of $H_{m,\ell}$ (black curve in Figure 11) and the computation of $H_{m,\ell}$ using the Ricatti equation (yellow curve). The integration of the Ricatti equation was done with a finer resolution and a lower hyper-viscosity than in the simulation of the linearized barotropic equation (52), the agreement between both results in Figure 11 thus shows that the resolution used in the simulation of (52) is high enough, and that the effect of hyper-viscosity is negligible.

We stress that the computation of $H_{m,\ell}(\theta)$ using the Ricatti equation (37) does not require the numerical integration of the linear dynamics (52). Typically, the integration of (52) over a time $T_{max} = 54,500$ takes about one week, while the resolution of the Ricatti equation (37) for a given value of θ is a matter of a few seconds. This enables the investigation of the statistics of rare events (large values of $|\theta|$ in Figure 11) extremely easily, as we now explain in more detail.

2. *Rate function and departure from Gaussian statistics*

The main goal of this study is to investigate the statistics of rare events in zonal jet dynamics, that cannot be described by the effective dynamics studied in section V. Using the previous numerical results, we now show how to quantify the departure from the effective description.

The large deviation rate function $\mathcal{L}_{m,\ell}$ entering in the Large Deviation Principle (49) can be computed from $H_{m,\ell}$ using (56). The result of this calculation[47] is shown in Figure 12 (yellow curve).

Because of the relation (48), $\mathcal{L}_{m,\ell}$ can also be interpreted as the large deviation rate function for the time-averaged advection term, denoted $\bar{R}_{m,\ell,\Delta t} \equiv \frac{1}{\Delta t}\int_0^{\Delta t} R_{m,\ell}(u)\,\mathrm{d}u$. In other words, the probability distribution function of $\bar{R}_{m,\ell,\Delta t}$ in the regime $\Delta t \gg \tau$ satisfies

$$\ln P_{m,\ell,\Delta t}\left(\bar{R}\right) \underset{\Delta t \gg \tau}{\sim} -\Delta t\, \mathcal{L}_{m,\ell}\left(\bar{R}\right). \tag{57}$$

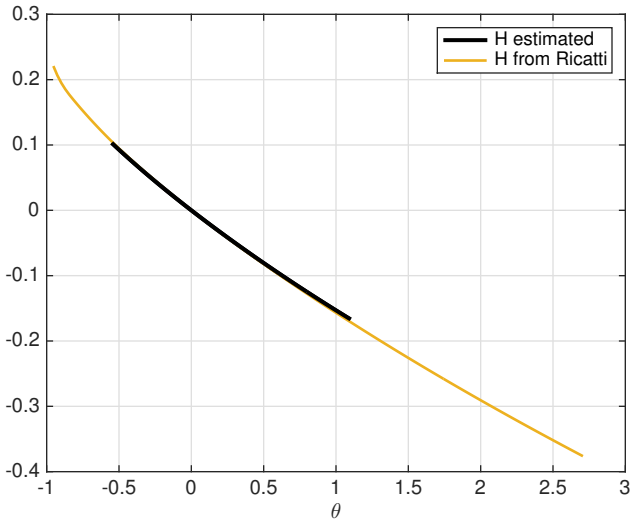


FIG. 11. Thick black line: scaled cumulant generating function $H_{10,3}(\theta)$ estimated from the numerical simulation of the linearized barotropic dynamics (52), with parameters defined in section II B and $\alpha = 0.073$. Statistical error bars are smaller than the width of this curve. Yellow curve: scaled cumulant generating function $H_{10,3}(\theta)$ computed from numerical integration of the Ricatti equation (37), using (36). The spectral cutoff in the Ricatti calculation is $L = 120$ (compared to $L = 80$ for the simulation of (52)), and the hyper-viscosity coefficient is such that the smallest scale has a damping rate of 4 (i.e. it is half of the hyperviscosity coefficient in the case $L = 80$). The estimated scaled cumulant generating function is in agreement with the one computed from the Ricatti equation, showing that the finite spectral cutoff and hyperviscosity are negligible in the calculation of $H_{10,3}(\theta)$. The numerical integration of the Ricatti equation enables access to larger values of $|\theta|$ (rarer events) extremely easily, see also Figure 12.

The Central Limit Theorem states that for large $\Delta t \gg \tau$, the statistics of $\bar{R}_{m,\ell,\Delta t}$ around its mean $\mathcal{R}_{m,\ell} \equiv \mathbb{E}_U\left[\bar{R}_{m,\ell,\Delta t}\right] = \mathbb{E}_U\left[R_{m,\ell}\right]$ are nearly Gaussian. A classical result in Large Deviation Theory is that the Central Limit Theorem can be recovered from the Large Deviation Principle[30]. Indeed, using the Taylor expansion of $H_{m,\ell}$ in powers of θ (31) and computing the Legendre-Fenschel transform (51), we get

$$\mathcal{L}_{m,\ell}\left(\bar{R}\right) = \frac{1}{2\mathcal{Z}_{m,\ell}}\left(\bar{R} - \mathcal{R}_{m,\ell}\right)^2 + O\left(\left(\bar{R} - \mathcal{R}_{m,\ell}\right)^3\right) \tag{58}$$

with $\mathcal{Z}_{m,\ell} \equiv \lim_{\Delta t\to\infty} \Delta t\, \mathbb{E}_U\left[\left[\bar{R}^2_{m,\ell,\Delta t}\right]\right]$. Using the Large Deviation Principle (57), this means that the statistics of $\bar{R}_{m,\ell,\Delta t}$ for small fluctuations around $\mathcal{R}_{m,\ell}$ are Gaussian with

variance $\mathcal{Z}_{m,\ell}/\Delta t$, which is exactly the result of the Central Limit Theorem. Then, the difference between the actual rate function $\mathcal{L}_{m,\ell}\left(\bar{R}\right)$ and its quadratic approximation (right-hand side of (58)) gives the departure from the Gaussian behaviour of $\bar{R}_{m,\ell,\Delta t}$.

From (58), the Gaussian behaviour is expected to apply roughly for $\left|\bar{R}-\mathcal{R}_{m,\ell}\right| \leq \sigma_{m,\ell,\Delta t}$ with $\sigma_{m,\ell,\Delta t} \equiv \sqrt{\mathcal{Z}_{m,\ell}/\Delta t}$. The values of $\mathcal{R}_{m,\ell} \pm \sigma_{m,\ell,\Delta t}$ are represented by the black vertical lines in Figure 12[48]. The quadratic approximation of the rate function is also shown in Figure 12 (purple curve). As expected, both curves are indistinguishable between the vertical lines (typical fluctuations), and departures from the Gaussian behaviour are observed away from the vertical lines (rare fluctuations). Namely, the probability of a large negative fluctuation is much larger than the probability of the same fluctuation for a Gaussian process with same mean and variance as $\bar{R}_{m,\ell,\Delta t}$. On the contrary, the probability of a large positive fluctuation is much smaller than the the probability of the same fluctuation for a Gaussian process with same mean and variance as $\bar{R}_{m,\ell,\Delta t}$.

The kinetic description basically amounts at replacing $\bar{R}_{m,\ell,\Delta t}$ by a Gaussian process with same mean and variance. From the results summarized in Figure 12, we see that such approximation leads to a very inaccurate description of rare events statistics. Understanding the influence of the non-Gaussian behavior of $\bar{R}_{m,\ell,\Delta t}$ on zonal jet dynamics is naturally a very interesting perspective of this work.

VII. CONCLUSIONS AND PERSPECTIVES

In this work we carried out a first study of the typical and large fluctuations of the Reynolds stress in fluid mechanics. Reynolds stress is certainly one of the key quantity in studying the largest scales of turbulent flows. This is especially true whenever a time scale separation is present, as then one expects an effective slow evolution equation to govern the large scale flow (see equation (2)). Not only the averaged momentum flux (the Reynolds stress) and averaged advection terms are essential, but also their fluctuations (that we call the Reynolds stress fluctuations).

We studied the case of a zonal jet for the barotropic equation on a sphere, in a regime for which time scale separation is relevant. For this case, we show then that the probability

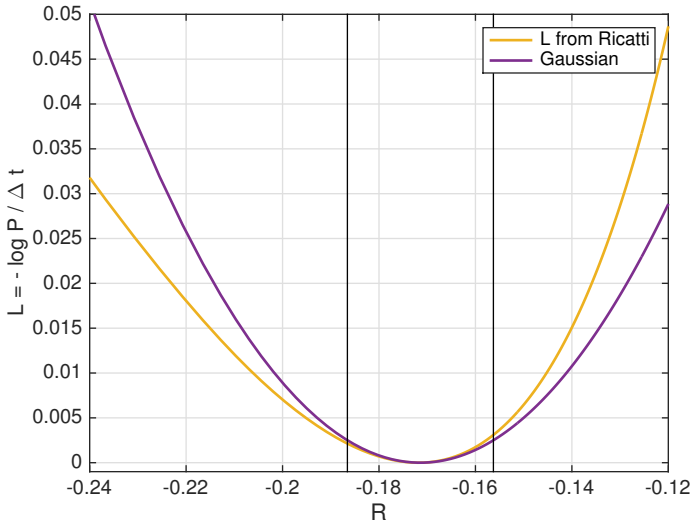


FIG. 12. Yellow curve: large deviation rate function $\mathcal{L}_{10,3}(\bar{R})$ computed from numerical integration of the Ricatti equation (37), using (36) and (51), with parameters defined in section II B and $\alpha = 0.073$. Purple curve: quadratic fit (58) that corresponds to a Gaussian process with same mean and variance as $\bar{R}_{10,3,\Delta t}$, the time-averaged advection term. Black vertical lines: standard deviation of $\bar{R}_{10,3,\Delta t}$ around its mean. Outside the vertical lines, we observe non-Gaussian behaviour of $\bar{R}_{10,3,\Delta t}$, in particular negative fluctuations are much more probable than positive ones.

distribution function of the equal-time (without time average) advection term has a distribution with typical fluctuations which are very large compared to the average, and with heavy tails. These probability distribution functions have exponential tails, both for the quasilinear and fully non-linear dynamics cases. For quasilinear dynamics we gave a simple explanation for these exponential tails.

When one is interested in the low frequency evolution of the jet, these high frequency fluctuations of the advection term and momentum fluxes are not relevant. We discussed that then the natural quantity to study is the large deviation rate function for the time averaged advection term (that we call the Reynolds stress large deviation rate function). We have proposed two methods to compute this rate function. First an empirical method, directly from the time series of the advection term, that could be applied to any dynamics. Second we show that for the quasilinear dynamics, the Reynolds stress large deviation rate function can be computed as the contraction of a solution of a matrix Riccati equation.

We demonstrated that such a computation can be performed by generalizing classical algorithms used to compute Lyapunov equations. Solving the matrix Riccati equation is much more computationally efficient, by several orders of magnitude, compared to accumulating statistics by numerical simulation, and gives direct and easy access to the probability of rare events. The approach is however limited to the quasilinear dynamics so far.

We discussed the Reynolds stress large deviation rate again for the specific case of a zonal jet that arises in turbulent barotropic flow on the rotating sphere. We illustrated the computation of the Reynolds stress large deviation rate, both using the empirical method and the Riccati equation. These two approaches give a very good agreement. This large deviation rate function clearly illustrate the existence of non-Gaussian fluctuations. The non Gaussian fluctuations are much more rare than Gaussian ones for positive values of the Reynolds stress component and much less rare than Gaussian for negative values.

Our work illustrates the possibility to compute Reynolds stress large deviation rate functions. It opens up a number of perspectives. A next step would be to study the spatial structure of the Reynolds stress fluctuation, and describe it from a fluid mechanics perspective. It would help to answer the following questions: What are the dominant spatial pattern for the fluctuations of the Reynolds stresses? What causes them? What is their effect on the low frequency variability of the large scale flow? The most interesting application of the Reynolds stress large deviation rate functions may be the study of rare long term evolutions of the large scale flow. For instance, in many examples, rare transitions between turbulent attractors have been observed, leading to a bistability phenomenology. In order to study quantitatively such a bistability phenomenology, for instance in order to compute transitions rates and transitions paths between attractors, one could consider equation (2) in the framework of Freidlin–Wentzell theory. The large deviation rate function we studied in this work would then be the basic building block, that would allow to define an action that should be minimized to compute transition paths and transition rates. In order to compute the action, the large deviation rate function should then be computed for any flow U along a possible transition path, as described in section VI C for a single example of a flow U.

An essential question, at a more mathematical level, is the validity of the quasilinear approximation as far as rare events are concerned. The self consistency of the quasilinear approach has been discussed theoretically by focusing on the average Reynolds stress[12]. This point has also been verified numerically in this work, through the study of properties of the

energy balance (see section V A) and through the verification of the fact that the linear equation correlation time has a limit when $\alpha \to 0$ (see section V B). However this does not necessarily imply that the quasilinear approximation is self-consistent as far as fluctuations, and more specifically rare fluctuations, are concerned. This could be addressed by studying the properties of solutions to the Ricatti equation in the limit $\alpha \to 0$ to assess whether or not the small scale dissipative mechanism (either viscosity or hyperviscosity) affects the statistics of the rare fluctuations. This problem is left as a prospect for future work.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013 Grant Agreement No. 616811) (F. Bouchet and T. Tangarife) and from the US NSF under Grant No. DMR-1306806 (J. B. Marston). J. B. Marston would also like to thank the Laboratoire de Physique de l'ENS de Lyon and CNRS for hosting a visit where some of this work was carried out.

REFERENCES

[1]H. Tennekes and J. L. Lumley, *A first course in turbulence* (MIT press, 1972).

[2]S. B. Pope, *Turbulent flows* (IOP Publishing, 2001).

[3]H. Touchette, "The large deviation approach to statistical mechanics," Physics Reports **478**, 1–69 (2009).

[4]F. Bouchet and E. Simonnet, "Random Changes of Flow Topology in Two-Dimensional and Geophysical Turbulence," Physical Review Letters **102**, 094504 (2009).

[5]C. M. Rohwer, F. Angeletti, and H. Touchette, "Convergence of free energy and large deviation estimators," Phys. Rev. E **92**, 052104 (2015).

[6]J. Herring, "Investigation of problems in thermal convection," Journal of the Atmospheric Sciences **20**, 325–338 (1963).

[7]R. H. Kraichnan, "Inertial Ranges in Two-Dimensional Turbulence," Phys. Fluids **10**, 1417–8 (1967).

[8]R. H. Kraichnan and D. Montgomery, "Two-dimensional turbulence," Rep. Prog. Phys. **43**, 547–619 (1980).

[9]F. Bouchet and A. Venaille, "Statistical mechanics of two-dimensional and geophysical flows," Physics Reports **515**, 227–295 (2012).

[10]B. F. Farrell and P. J. Ioannou, "Structural stability of turbulent jets," Journal of the atmospheric sciences **60**, 2101–2118 (2003).

[11]J. Marston, W. Qi, and S. Tobias, "Direct Statistical Simulation of a Jet," arXiv preprint arXiv:1412.0381 (to appear in the book "Zonal Jets: Phenomenology, Genesis, Physics," eds. Boris Galperin and Peter L Read, Cambridge University Press) (2014).

[12]F. Bouchet, C. Nardini, and T. Tangarife, "Kinetic theory of jet dynamics in the stochastic barotropic and 2d Navier-Stokes equations," Journal of Statistical Physics **153**, 572–625 (2013).

[13]K. Srinivasan and W. Young, "Reynolds stress and eddy diffusivity of β-plane shear flows," Journal of the Atmospheric Sciences **71**, 2169–2185 (2014).

[14]J. Laurie, G. Boffetta, G. Falkovich, I. Kolokolov, and V. Lebedev, "Universal profile of the vortex condensate in two-dimensional turbulence," Physical review letters **113**, 254503 (2014).

[15]E. Woillez and F. Bouchet, "Theoretical prediction of reynolds stresses and velocity profiles for barotropic turbulent jets," EPL **118**, 54002 (2017).

[16]F. Bouchet, C. Nardini, and T. Tangarife, "Kinetic theory and quasilinear theories of jet dynamics," arXiv preprint arXiv:1602.02879 (to appear in the book "Zonal Jets: Phenomenology, Genesis, Physics," eds. Boris Galperin and Peter L Read, Cambridge University Press) (2016).

[17]T. Tangarife, *Kinetic theory and large deviations for the dynamics of geophysical flows, https://tel.archives-ouvertes.fr/tel-01241523*, Theses, Ecole normale supérieure de lyon - ENS LYON, https://tel.archives-ouvertes.fr/tel-01241523 (2015).

[18]C. Nardini and T. Tangarife, "Fluctuations of large-scale jets in the stochastic 2d euler equation," arXiv:1602.06720 (2016).

[19]G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation* (Cambridge University Press, 2006).

[20]C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences* (Springer Series in Synergetics, Berlin: Springer, —c1994, 2nd ed. 1985. Corr. 3rd

printing 1994, 1994).

[21]K. Srinivasan and W. R. Young, "Zonostrophic Instability," Journal of the atmospheric sciences **69**, 1633–1656 (2011).

[22]C. C. Porco *et. al*, "Cassini imaging of jupiter's atmosphere, satellites, and rings," Science **299**, 1541 (2003).

[23]A program that implements spectral DNS for the non-linear and quasi-linear equations, solves the non-linear Riccati equation, and includes graphical tools to visualize statistics, is freely available. The application "GCM" is available for OS X 10.9 and higher on the Apple Mac App Store at URL http://appstore.com/mac/gcm.

[24]D. K. Lilly, "Numerical simulation of two-dimensional turbulence," Physics of Fluids (1958-1988) **12**, II–240 (1969).

[25]I. Grooms, "A Gaussian-product stochastic Gent-McWilliams parameterization," Ocean Modelling **106**, 27–43 (2016).

[26]N. A. Bakas, N. C. Constantinou, and P. J. Ioannou, "S3T stability of the homogeneous state of barotropic beta-plane turbulence," Journal of the Atmospheric Sciences **72**, 1689–1712 (2015).

[27]J. Marston, E. Conover, and T. Schneider, "Statistics of an unstable barotropic jet from a cumulant expansion," Journal of the Atmospheric Sciences **65**, 1955–1966 (2008).

[28]S. Tobias and J. Marston, "Direct statistical simulation of out-of-equilibrium jets," Physical Review Letters **110**, 104502 (2013).

[29]We can restrict ourselves to real ξ_m decomposing ω_m and ψ_m into real and imaginary parts.

[30]M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems, Second Edition*, Vol. 260 (Springer, 1998) Chap. 7.

[31]G. Pavliotis and A. Stuart, *Multiscale methods: averaging and homogenization*, Vol. 53 (Springer, 2008).

[32]F. Bouchet, T. Grafke, T. Tangarife, and E. Vanden-Eijnden, "Large deviations in fast-slow systems," Journal of Statistical Physics **162**, 793–812 (2016).

[33]S. Tobias, K. Dagon, and J. Marston, "Astrophysical fluid dynamics via direct statistical simulation," The Astrophysical Journal **727**, 127 (2011).

[34]M. E. Newman and G. T. Barkema, *Monte Carlo methods in statistical physics*, Vol. 13 (Clarendon Press Oxford, 1999).

[35]G. C. Papanicolaou, "Introduction to the asymptotic analysis of stochastic equations,"

Modern modeling of continuum phenomena (Ninth Summer Sem. Appl. Math., Rensselaer Polytech. Inst., Troy, NY, 1975) **16**, 109–147 (1977).

[36]The application "GCM" integrates the equation 34 and the effective dynamics 35.

[37]Note that the ordering of products with L_x and L_x^T differs between (34) and (37).

[38]W. Krauth, *Statistical mechanics: algorithms and computations*, Vol. 13 (Oxford University Press, 2006).

[39]Explicitely,

$$\mathbb{E}_{\frac{T}{\Delta t}}\left[\left(\int_{t_0}^{t_0+\Delta t}\left(R(s)-\bar{R}_T\right)\,\mathrm{d}s\right)^2\right]=\frac{\Delta t}{2T}\sum_{k=0}^{\frac{2T}{\Delta t}-2}\left(\int_{k\Delta t/2}^{k\Delta t/2+\Delta t}\left(R(u)-\bar{R}_T\right)\,\mathrm{d}u\right)^2, \quad (59)$$

assuming for simplicity that $T/\Delta t$ is an integer. Generalisations to any $T, \Delta t$ is straightforward, replacing $2T/\Delta t$ by its floor value. The 50% overlap is suggested by Welch's estimator of the power spectrum of a random process[49].

[40]D. T. Gillespie, "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral," Phys. Rev. E **54**, 2084–2091 (1996).

[41]F. Ait-Chaalal, T. Schneider, B. Meyer, and J. Marston, "Cumulant expansions for atmospheric flows," New Journal of Physics **18**, 025019 (2016).

[42]The statistical error bars for p_{fluct} are computed from the error in the estimation of $Z[U](\phi)$, which is similar to the estimation of the correlation time τ described in section IV D 1. The statistical error bars for p_{mean} are computed from the error in the estimation of the average F, given by $(\delta F)^2 = \frac{1+2\tau/\Delta t}{N}\mathrm{var}(F)$ where τ is the autocorrelation time of F, Δt the time step between measurements of the Reynolds' stress and N the total number of data points[34].

[43]J. B. Marston, "Statistics of the general circulation from cumulant expansions," Chaos **20**, 041107 (2010).

[44]F. Bouchet and H. Morita, "Large time behavior and asymptotic stability of the 2D Euler and linearized Euler equations," Physica D Nonlinear Phenomena **239**, 948–966 (2010), arXiv:0905.1551 [cond-mat.stat-mech].

[45]In this spherical geometry the maximum is taken over the inner jet region $\phi \in [-\pi/7, \pi/7]$.

[46]The time t in the upper and lower bounds of the integral in (54) are not relevant here, as we are considering the statistically stationary state of (52).

[47]Here the Legendre-Fenschel transform (51) is estimated as $\mathcal{L}_{m,\ell}(\dot{\omega}_z) = \theta^\star \cdot \dot{\omega}_z - H_{m,\ell}(\theta^\star)$ where $\theta^\star$ is the solution of $\dot{\omega}_z = \partial_\theta H_{m,\ell}(\theta^\star)$. Other estimators could be considered[5].

[48]The value of Δt used in this estimation is the optimal one $\Delta t^\star$, defined in section IV D.

[49]P. D. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," IEEE Transactions on audio and electroacoustics **15**, 70–73 (1967).