Eddy-mixing entropy and its maximization in forced-dissipative geostrophic turbulence

Tomos W. David,* Laure Zanna, and David P. Marshall

University of Oxford, Department of Physics,

Clarendon Laboratory, Parks Road, Oxford, OX1 3PU

(Dated: July 6, 2018)

Abstract

An equilibrium, or maximum entropy, statistical mechanics theory can be derived for ideal, unforced and inviscid, geophysical flows. However, for all geophysical flows which occur in nature, forcing and dissipation play a major role. Here, a study of eddy-mixing entropy in a forced-dissipative barotropic ocean model is presented. We heuristically investigate the temporal evolution of eddy-mixing entropy, as defined for the equilibrium theory, in a strongly forced and dissipative system. It is shown that the eddy-mixing entropy provides a descriptive tool for understanding three stages of the turbulence life cycle: growth of instability; formation of large scale structures; and steady state fluctuations. The fact that the eddy-mixing entropy behaves in a dynamically balanced way is not a priori clear and provides a novel means of quantifying turbulent disorder in geophysical flows. Further, by determining the relationship between the time evolution of entropy and the maximum entropy principle, evidence is found for the action of this principle in a forced-dissipative flow. The maximum entropy potential vorticity statistics are calculated for the flow and are compared with numerical simulations. Deficiencies of the maximum entropy statistics are discussed in the context of the mean-field approximation for energy. This study highlights the importance of entropy and statistical mechanics in the study of geostrophic turbulence.

* tomos.david@physics.ox.ac.uk

I. INTRODUCTION

Due to the highly chaotic nature of turbulence and its large number of degrees of freedom, the methods of statistical physics are an attractive approach for understanding the physics of turbulent systems. An equilibrium, or maximum entropy, theory has been proposed for ideal geophysical flows; by ideal we mean an isolated system with no forcing nor dissipation. However, it is not clear to what extent this equilibrium theory is useful for realistic, forced and dissipative, geophysical flows. With increasingly pressing concerns about the Earth's climate it is pertinent to consider whether a statistical mechanics approach can be used to improve our in sub-gridscale parameterizations of ocean turbulence at length scales of 10-100km.

The first example of the application of statistical mechanics to two-dimensional turbulence comes from [28], where a model of singular point vortices was proposed to characterize turbulent features. The statistical mechanics of point vortices has received much study since then. Studies such as [19] and [20] led the way towards continuous vorticity fields through utilizing the invariance of energy and enstrophy in variational problems. Geostrophic flows over topography were tackled using this methodology: in [38] via the maximization of an entropy; and independently in [7] via a phenomenological minimum enstrophy principle.

Work by Miller, Robert and Sommeria established a theory of equilibrium statistical mechanics of two-dimensional and simple geophysical flows [26, 27, 33–35], which we refer to as the 'Miller-Robert-Sommeria theory'. Miller-Robert-Sommeria theory produces an equilibrium statistical mechanics, by forming a variational problem to maximize an appropriate entropy, of two-dimensional or quasi-two-dimensional flow which: relies on the underlying Hamiltonian structure of the dynamics; applies to continuous vorticity fields; and conserves the invariants of motion of the flow [e.g. see reviews by 6, 8, 22, 40]. The power of the Miller-Robert-Sommeria theory is that the work of [28], [20], [38] and [7] are contained within this framework as particular limits or simplifications. This equilibrium/ideal (no forcing, no dissipation) theory has been used to suggest a statistical mechanical explanation for: the formation of ocean rings and jets [45]; the dynamics of the stratospheric polar vortex [31, 48]; Jupiter's Great Red Spot [5, 42]; bottom trapped ocean currents [44]; the vertical structure in stratified quasi-geostrophic flow [24, 39, 46]; as well as the global ocean circulation and its associated density profiles [37]. Miller-Robert-Sommeria statistical mechanics theory suffers

from the restrictive assumption of ideal flow; understanding how this statistical mechanics framework can be used in a forced-dissipative context is essential in order to apply these powerful ideas to realistic geophysical flows.

An application which is of particular interest to the authors, is ocean mesoscale (10-100km) turbulence. The eddies which make up this large-scale ocean turbulence lie beyond the computational reach of modern climate models when run to dynamic and thermodynamic steady state. These climate models typically have an ocean resolution of 50-100km while mesoscale eddies have typical length scales of 10-100km, meaning that it is necessary to parameterize the effect of eddies on the mean flow via turbulence closures. Although these motions are generated internally, from the advective term of the equations of motion, they are influenced by forcing and dissipation at wide range of scales. Recent studies have explored the possibility of a stochastic approach [3, 16, 30, 49] for ocean models. With this growing interest in the stochastic nature of mesoscale eddies, it is timely to study the statistics of vorticity, and the underlying organizing principles influencing these statistics, in simplified ocean models. Consequently, we search for theories of turbulence which are statistical in nature and include both forcing and dissipation.

In this study, inspired by the Miller-Robert-Sommeria theory we follow here a heuristic approach [17], in order to analyse entropy in forced and dissipative numerical simulations of simplified geophysical flows. Motivated by efforts to parameterize ocean mesoscale turbulence we consider the ability of a statistical mechanics approach in determining small-scale statistics given knowledge of the large-scale flow. In this study we will not determine nor characterize the equilibria of the ideal system as this will be unaccessible to us as we assume knowledge of the large-scale flow; this problem has, however, been approached in [9]. In this study we will consider a wide range of forcing and dissipation strength complimenting studies which have found that equilibrium statistical mechanics can give good predictions in the weak forcing and weak dissipation regime [4]. The specific aims of this study are as follows.

- To determine the impact of forcing and dissipation on the evolution of entropy in a turbulent barotropic jet, both analytically and numerically.

- To test the maximum entropy principle[1] and to understand the utility of this principle

[1] Not to be confused with the similarly named maximum entropy production principle.

in the context of a forced-dissipative turbulent jet.

- To use the maximum entropy principle as a means to formulate a relationship between dynamically balanced quantities and the small-scale statistics of the flow.

The paper is structured as follows. In Section II, we describe the barotropic model and the numerical experiments used in this study. In Section III, we introduce some key concepts, especially the eddy-mixing entropy. In Section IV, we derive analytical expressions for the influence of forcing and linear drag on the entropy. In Section V A and V B, we diagnose the entropy for a freely-decaying and a forced-dissipative turbulent jet, respectively to test the predictions of Section IV and consider entropy as balanced dynamical quantity. In Section VI, we derive a test for entropy maximization in a forced-dissipative system and compute the maximum entropy statistics, comparing with the numerical simulations. In Section VII, we discuss the mean-field approximation for energy and its relation to the maximum entropy statistics. In Section VIII, the study is concluded with some closing remarks.

II. MODEL AND EXPERIMENTS

A. Model Setup

We solve the barotropic vorticity equation on a β plane within a singly-periodic domain, recently used and described in [10]. The simplicity of this model allows us to perform many high resolution simulations, while the channel configuration provides an analogue to the Southern Ocean turbulent jet dynamics. The equation of motion is given by

$$\frac{\partial q}{\partial t} = -\{\psi, q\} - r\nabla^2\psi - \nu_h\nabla^6\psi - \partial_y\tau(y), \tag{1}$$

where the potential vorticity q is given by

$$q = \nabla^2\psi + \beta y; \tag{2}$$

$\nabla^2\psi$ is the relative vorticity and βy is the planetary vorticity making q equivalent to the absolute vorticity in this barotropic model; τ is the zonal wind stress which is defined to be a function of meridional distance only; r and ν_h are the linear drag coefficient and the hyper-viscosity respectively; the braces denote the horizontal Jacobian operator given as

$$\{A, B\} = \partial_x A\,\partial_y B - \partial_x B\,\partial_y A. \tag{3}$$

The biharmonic diffusion is used for numerical stability, preferentially dissipating the grid-scale noise compared to Laplacian diffusion. The hyper-viscosity is chosen to be as small as possible while allowing us to treat the linear drag as the dominant dissipative term in this study. Linear drag is an attractive choice of dissipative term due to its analytical tractability as well as its being analogous to oceanic bottom drag. A quadratic drag term could also be used but we leave this analysis for future study.

The periodicity in the zonal direction is employed to solve the model using a pseudo-spectral method and is modified from a pre-existing code [13–15]. The model domain is shown in Figure 1. The boundary conditions are free-slip

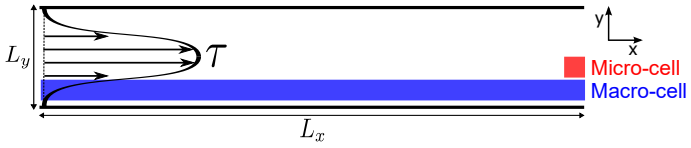

FIG. 1. Geometry of domain. wind stress, τ, in this study points from left to right. L_x and L_y are the longitudinal and latitudinal extents of the domain respectively. Macro-cells (blue) and micro-cells (red) in this study. Exploiting zonal symmetry zonal bands are used as macro-cells whilst micro-cells are the grid-points within.

$$\nabla^{2n}\psi|_{N,S} = 0, \tag{4}$$

where $n = 1, 3$; no normal flow

$$\partial_x\psi|_{N,S} = 0; \tag{5}$$

and global momentum conserving

$$\psi|_{N,S} = \pm\frac{\Gamma(t)}{2}. \tag{6}$$

We find Γ by solving the prognostic integral momentum balance,

$$\frac{d\Gamma}{dt} = -\iint \mathrm{d}^2\mathbf{x}\left[r\frac{\partial\psi}{\partial y} + \tau\right]. \tag{7}$$

This is the same condition used by [41] and is the barotropic (and rigid lid) limit of the general integral momentum balance derived in [23]. By applying this boundary condition we are able to impose a fixed wind stress forcing rather than relaxing to a background shear as is often done for models of this type [e.g. 29].

The ideal dynamics of this flow conserves various other quantities in addition to momentum as described below. In ideal flow the energy

$$E = \frac{1}{2}\iint \mathrm{d}^2\mathbf{x}\,(\nabla\psi)\cdot(\nabla\psi) \tag{8}$$

is conserved. This can be rewritten as

$$E = \frac{\Gamma}{4}(u_S - u_N) - \frac{1}{2}\iint \mathrm{d}^2\mathbf{x}\,\psi(q-\beta y), \tag{9}$$

exploiting the relationship between q and ψ as well as the boundary condition described above, where u_N and u_S are the velocities along the north and south boundaries respectively. When the the boundary flow is North-South asymmetric (e.g. when the wind stress is North-South asymmetric) the first term in will be non-zero, however, this is not the case in the flow realizations presented in this study where $u_N = u_S$.

In addition, ideal flow conserves the integral of any function of potential vorticity, called Casimirs [e.g. 36], such that

$$C = \iint \mathrm{d}^2\mathbf{x}\,c(q), \tag{10}$$

where c is an arbitrary function. We are primarily interested in the polynomial Casimirs which we will denote as

$$C_n = \iint \mathrm{d}^2\mathbf{x}\,q^n. \tag{11}$$

Two Casimirs of particular physical importance are the circulation, $n = 1$, and the enstrophy, $n = 2$. Alternatively, all Casimirs can be conserved simultaneously by conserving the global potential vorticity distribution, Π, given by

$$\Pi(q) = \frac{dA(q)}{dq}, \tag{12}$$

where $A(q)$ is the area of the domain occupied by points with a value of potential vorticity less than q. $A(q)$ is proportional to the global *cumulative* potential vorticity distribution function.

B. Numerical Experiments

We perform two sets of numerical experiments. The first set of experiments is based on a freely-decaying unstable jet in which the initial jet has a velocity profile

$$u(y) = u_0 \mathrm{sech}^2(y), \tag{13}$$

with $u_0 = 10$. The unstable jet evolves freely under the action of hyper-viscosity and varying strengths of linear drag. The drag coefficient is varied over a wide range from a lowest value where flow is just numerically stable to a highest value where the turbulence is beginning to be damped away.

The second set of experiments is based on a forced-dissipative turbulent jet spun up from rest with varied strength of the wind stress. The wind stress profile is kept the same for all simulations as,

$$\tau = \tau_0 \mathrm{sech}^2 \left(\frac{y}{\delta} \right), \tag{14}$$

but the magnitude of the jet is varied by changing the value of τ_0. The wind stress strength is doubled seven times to ensure a very wide range of turbulent flows. δ is the width parameter and is fixed for all simulations. Table I summarizes the values used in the different simulations. The parameters which are held constant for both experiments are given in Table II; these parameters were extensively tested for both this study and [10] to ensure numerical convergence.

III. EDDY-MIXING ENTROPY

The eddy-mixing entropy is not the same as the thermodynamic entropy associated with molecular motions. The eddy-mixing entropy is a measure of the disorder of the large scale turbulent flow, and depends on the choice of coarse-graining which distinguishes between the large scales and the small scales of the flow. The form of entropy we will use is chosen by analogy to the equilibrium entropy which has been shown to be the appropriate form for deriving an equilibrium statistical mechanics for *ideal* two-dimensional and geophysical flows and can be justified from a large deviation theory framework [25]. It is important to note that there is no a priori justification for there use of this entropy for non-equilibrium systems but follow empirically driven approach in this study.

To define the eddy-mixing entropy we follow the heuristic approach presented in [17]. To proceed we will define two sub-systems of the full flow:

- A *micro-cell* which is the smallest scale over which the details of the flow are important. The micro-cell is equivalent to the grid-cell for a high resolution numerical simulation. We think of each micro-cell as being characterized by only one value of the potential

Simulation ID	Wind stress strength, τ_0	Linear drag coefficient, r	Timestep, dt
Freely-decaying experiments			
D_1	0.000	0.0008	0.0005
D_2	0.000	0.0009	0.0005
D_3	0.000	0.0010	0.0010
D_4	0.000	0.0020	0.0010
D_5	0.000	0.0030	0.0010
D_6	0.000	0.0040	0.0010
D_7	0.000	0.0050	0.0010
Forced-dissipative experiments			
FD_1	0.005	0.0050	0.0010
FD_2	0.010	0.0050	0.0010
FD_3	0.020	0.0050	0.0010
FD_4	0.040	0.0050	0.0010
FD_5	0.080	0.0050	0.0010
FD_6	0.160	0.0050	0.0010
FD_7	0.320	0.0050	0.0010

TABLE I. List of experiments and non-dimensional parameters.

vorticity. We choose the grid-cell here for practical convenience but a more physically motivated choice could be used, for example the Batchelor scale.

- A *macro-cell* which is comprised of a number of micro-cells and is related to a choice of some coarse graining scale. The macro-cells should be chosen to exploit some dynamical symmetry of the system. In our case we, for the most part, choose zonal bands exploiting the zonal symmetry of the system apart from in Section VI A where we choose contours of instantaneous streamfunction. It important to note that the first choice of macro-cell cannot be used for zonal-symmetry-breaking flows (e.g. emergence of coherent eddies) while the second choice assumes knowledge of the answer we are seeking. To extend the analysis presented in this study to coherent eddy flows a

Parameter	Value
Meridional extent, L_y	$5\pi/2$
Zonal extent, L_x	20π
Number of zonal grid-points, n_x	1024
Number of meridional grid-points, n_y	128
Time-step, dt	1×10^{-3}
Output frequency	1
Total time of output	1×10^{4}
Hyper-viscosity, ν_h	2×10^{-6}
Wind stress width parameter, δ	0.4
Beta parameter, β	0.2

TABLE II. Fixed parameters of model simulations.

sophistical, partly Lagrangian, averaging method which respects the large-scale flow topology would be required.

The macro- and micro-cells used in this study are schematically illustrated in Figure 1.

Using this definition of the macro-cells, an eddy-mixing entropy is defined by counting the number of ways to arrange the micro-cells of value of potential vorticity into the macro-cells. The eddy-mixing entropy can be expressed as

$$S = \ln W = \ln \prod_I \frac{M^{(I)}!}{\prod_r M_r^{(I)}!}, \tag{15}$$

where $M^{(I)}$ is the number of micro-cells in the I^{th} macro-cell and $M_r^{(I)}$ is the number of micro-cells with the r^{th} value of potential vorticity in the I^{th} macro-cell. This counting method is adapted from [21].

For large numbers of micro- and macro-cells we can take the continuous limit to get

$$S[\rho] = -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\, \rho(\tilde{q}|\mathbf{x}) \ln(\rho(\tilde{q}|\mathbf{x})), \tag{16}$$

in terms of the probability distribution function, ρ. In words, *ρ is the probability of measuring a value, $\tilde{q}$, of the potential vorticity at the point* $\mathbf{x}$ *in the domain.* In this expression $\mathbf{x}$ has taken the place of I in labeling the macro-cell. The coordinate $\mathbf{x}$ should be interpreted

as a coarse-grained or smoothed coordinate. The $\tilde{q}$ is a random variable representing the result of a measurement of potential vorticity, and not the potential vorticity field, $q(\mathbf{x})$. The difference between the probability distribution, $\rho(q|\mathbf{x})$, used here and the full probability distribution (functional) of the system, $p[q(\mathbf{x})]$ should be noted. We interpret ρ as a marginal distribution of p and they only become equal when neighboring macro-cells become statistically independent, which is the case for the *ideal* theory which is in equilibrium; this is discussed in more detail in Section VII. However, we choose to heuristically study the distribution ρ which is accessible by numerical computation and, as will be shown, its behaviour is important in non-equilibrium systems. The eddy-mixing entropy is the sum over the continuous information entropies associated with the distribution of potential vorticity in each macro-cell. The method of numerically determining these entropies are given in an Appendix 1 and is used throughout this study.

IV. ANALYTICAL MODEL FOR EVOLUTION FOR ENTROPY

The fundamental quantity that we are interested in is the eddy mixing entropy given by (16) in Section III. In this section we derive a tendency equation for this entropy. We are not able to derive a full theory as the effects of the non-linear or non-local terms in the vorticity equation, (1), Section II, do not seem to be analytically tractable. Nevertheless it is possible to derive analytical expressions for the entropy evolution due to the remaining linear and local terms in the vorticity equation: the wind stress curl and the linear drag. We will now derive the influence of these terms on the entropy leaving the unknown tendency due to advection and hyper-viscosity as a residual, P.

Ignoring the non-local and non-linear terms, which include the hyper-viscous term and the advection, we have the following equation for the evolution of potential vorticity:

$$\frac{\partial q}{\partial t} = -r(q - \beta y) - g(y) \tag{17}$$

where $g(y) = -\partial_y \tau(y)$ is the constant (in time) forcing.

Equation (17) leads to the following partial differential equation for the probability distribution function, ρ,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \tilde{q}}[r(\tilde{q} - \beta y)\rho] - g(y)\frac{\partial \rho}{\partial \tilde{q}}. \tag{18}$$

The two terms on the right hand side of equation (18) were also derived in [18] but here we consider the effect of these terms on the entropy. We can write the entropy tendency as

$$\dot{S} = -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\dot{\rho}\ln\rho, \tag{19}$$

where the dot represents differentiation with respect to time. By substituting equation (18) into (19) we can derive (see Appendix 2) the influence of these terms on the entropy, yielding

$$\dot{S} = P - Ar, \tag{20}$$

where we have now included the effect of advection and hyper-viscosity as the residual, P; and A is the area of the domain. As the hyper-viscosity is small in our numerical calculations, we will take the liberty of referring to P as the *advective production of entropy*.

Notably, (20) does not have an explicit dependence on the zonally symmetric forcing as a constant wind stress only shifts the distribution in each zonal band and the entropy is invariant to these shifts. The linear drag leads to a remarkably simple term which is a perpetual and constant sink of entropy.

V. ENTROPY IN THE NUMERICAL SIMULATIONS

A. Eddy-mixing entropy in freely-decaying turbulence

As seen in (20) the entropy tendency has no explicit dependence on the forcing. However, the forcing will contribute to determining the behaviour of the advective production of entropy, P. To illuminate the effect of forcing we first consider the evolution of entropy in the absence of forcing: freely decaying unstable flow.

We begin by examining the evolution of entropy for short times as the instabilities grow then decay shown in Figure 2. In region A of Figure 2, the entropy increases very quickly concomitant with the exponential growth of eddy energy through shear instability in the jet. There is little spread in rate of the entropy growth in simulations $\mathrm{D}_1, ..., \mathrm{D}_7$ with changed drag parameter. This growth is arrested for all experiments at a maximum value in region B for the same time and where the maximum is also not greatly changed with the differing linear drag coefficient. The entropy then decreases, in the period B to C, towards its asymptotic behaviour. The rate of decrease is *greater* than can be explained by the sink of entropy due

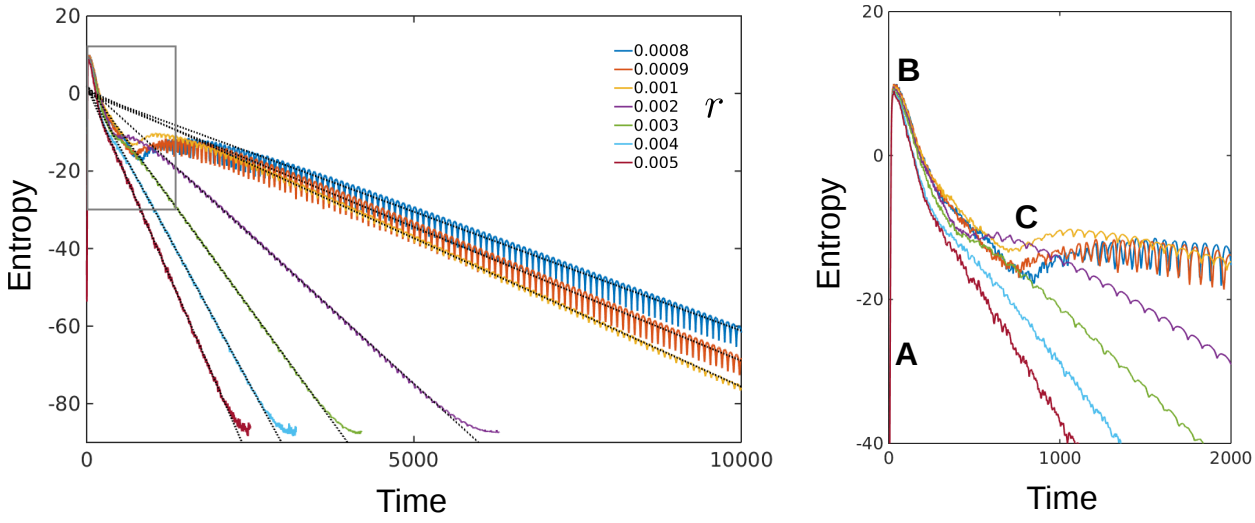

FIG. 2. Entropy as a function of time for different values of linear drag coefficient, simulations $D_1, ..., D_7$. A - We find that exponential growth of barotropic instabilities lead to a very fast growth in the entropy. B - The entropy reaches a maximum value which is insensitive to the value of the linear drag coefficient. This is followed by a decrease in entropy at a rate greater than the contribution of the linear drag. This implies that this decrease is due to the eddies themselves. C - At long time the entropy decay is linear and is explained by the linear drag. For the lowest values of linear drag coefficient a persistent Rossby wave emerges causing a secondary increase in the entropy before the long-time behaviour is seen.

to linear drag, Ar. This means that in the period between B to C the advective production of entropy must become negative and acts as a sink of entropy. As we shall see in the following sections this transient slump of entropy is concomitant with the emergence of large-scale flow structures and the decrease in disorder that this entails.

There is an interesting difference between simulations $D_1, ..., D_3$ and the other simulations. These low drag simulations see a second increase in entropy (Figure 2, near Time = 1000) toward the long-time behaviour as well as an oscillation about the long time behaviour. These effects can be illuminated by examining the flow, at low linear drag coefficient a persistent Rossby wave forms causing an increase in the disorder as compared to laminar flow as well as the observed oscillations. This however does not inhibit the asymptotic entropy decrease due to dissipation which erases small-scale fluctuation while, in the right parameter regime, leaving the large-scale flow intact.

In freely-decaying turbulence the eddies will ultimately die away through the action of

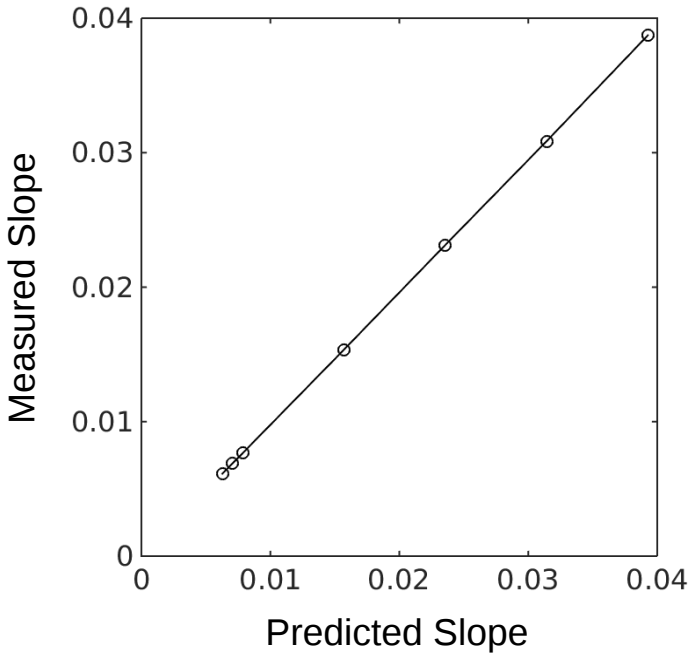

FIG. 3. Measured slope of long time linear decay of the entropy, diagnosed from simulations $\mathrm{D}_1, ..., \mathrm{D}_7$, plotted against the predicted slope. The agreement is near perfect.

linear drag and hyper-viscosity causing the advective production of entropy, P, to tend to zero at long times. In this case equation (20) tends to the asymptotic solution for the entropy evolution

$$S(t) \approx -Art + K \quad \text{for long times,} \tag{21}$$

where K is a constant of integration. We can test this hypothesis in a simulation of a freely decaying unstable jet as for long times we would expect the activity of the eddies to asymptotically decay to zero. We compare this prediction with the first numerical experiment described in Section II and shown in Figure 2. We see a striking agreement between the long time behaviour, to the right of C, predicted by (21) and the slopes diagnosed from the simulations. Figure 3 shows the agreement between the predicted and the measured long time slope of the linear entropy decrease which is found to be near exact.

B. Eddy-mixing entropy in forced-dissipative turbulence

We now turn our attention to the entropy evolution in the forced-dissipative experiments $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We start by considering the entropy for short times, comparing it to snapshots

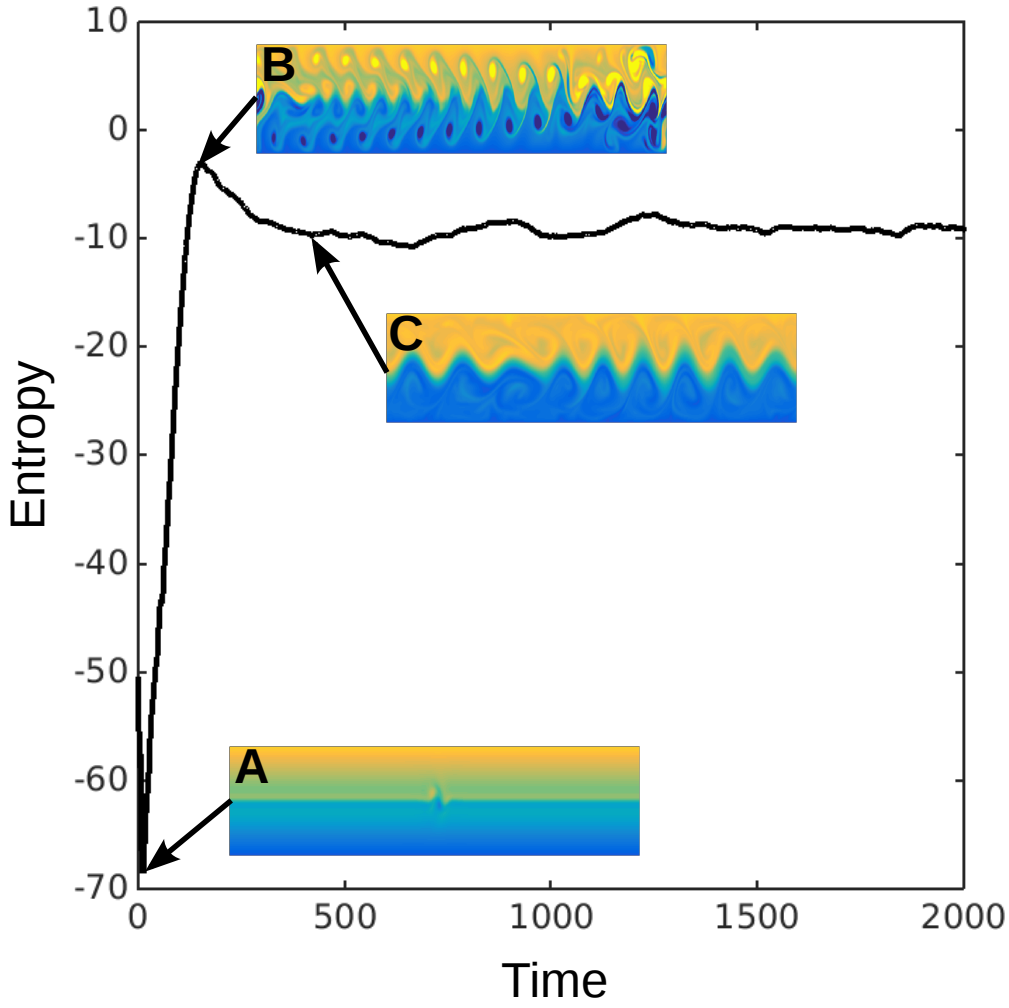

FIG. 4. Evolution of entropy during spin-up of forced-dissipative flow. A - Initially laminar flow with small perturbation. As the flow becomes unstable the eddy-mixing entropy increases rapidly. B - Entropy reaches a global maximum at the point where turbulence has spread across the whole domain. Disorder is at small scales. C - Coincident with the emergence of a large scale Rossby wave the entropy decreases. Subsequently the entropy fluctuates around a balanced time-mean value.

of the potential vorticity. As an example we consider experiment FD_3. Figure 4 shows the evolution of entropy as the system evolves to a statistically steady state. Initially, in region A, the flow is near laminar with only the small initial perturbation. We see that this corresponds to a low value of entropy. Once instabilities begin to grow the corresponding growth of entropy is fast and grows to a maximum value much like the evolution in the freely-decaying simulations. At the maximum of entropy, region B, the turbulence has covered the

whole domain with small scale eddies. As these eddies mix the potential vorticity we see a slump in the entropy. When we examine the flow at the bottom of the slump, region C, we see that a large scale Rossby wave has emerged propagating on a sharp potential vorticity gradient corresponding to a mixing barrier. This transient decrease of entropy, or disorder, in the system allows us to describe the way in which energy has condensed at large scales in an entropic sense. The concomitance of this decrease in entropy with the emergence of large scales leads to a novel interpretation of well known inverse transfer of energy phenomena: the emergence of coherent large scales can be described by the decrease of entropy in this system during the transient spin-up of the statistically steady state.

At longer times the entropy fluctuates around a balanced steady state value. This behaviour of entropy is the same for all the forced-dissipative simulations, $\mathrm{FD}_1, ..., \mathrm{FD}_7$, shown in Figure 5a, much like statistically steady state balance of energy, shown in Figure 5b. Both the time-mean entropy and energy increase with the wind stress strength in steady state as well as exhibit fluctuations about this mean (although the energy fluctuations are suppressed in Figure 5b due to the logarithmic scale). The balanced steady state behaviour of the entropy is explained, according to the reasoning of Section IV, by the competition between the advective production of entropy and the constant sink due to linear drag, that is

$$\overline{P} - Ar = 0, \tag{22}$$

where the over-line denotes the time-mean in statistically steady state. It is important to note that, because $-r$ is a merely a negative number, both the increase and decrease in the entropy fluctuations arise from the advective production, P. That is, *eddies can act as both a source and a sink of entropy.* It is important to note that the action of P as a source and a sink must be associated with the presence of dissipation which allows there to be fluctuations in otherwise conserved global quantities such as energy and entropy which exhibit inverse and direct transfers between scales.

Although the time derivative of entropy has no explicit dependence on forcing, the forcing does supply energy to the turbulent motions by sustaining the eddy production of entropy, unlike in the case of the freely-decaying experiment. The forcing implicitly sets the maximum and steady state value of entropy. We will further consider how the entropy is related to other well-known dynamical quantities in the following section.

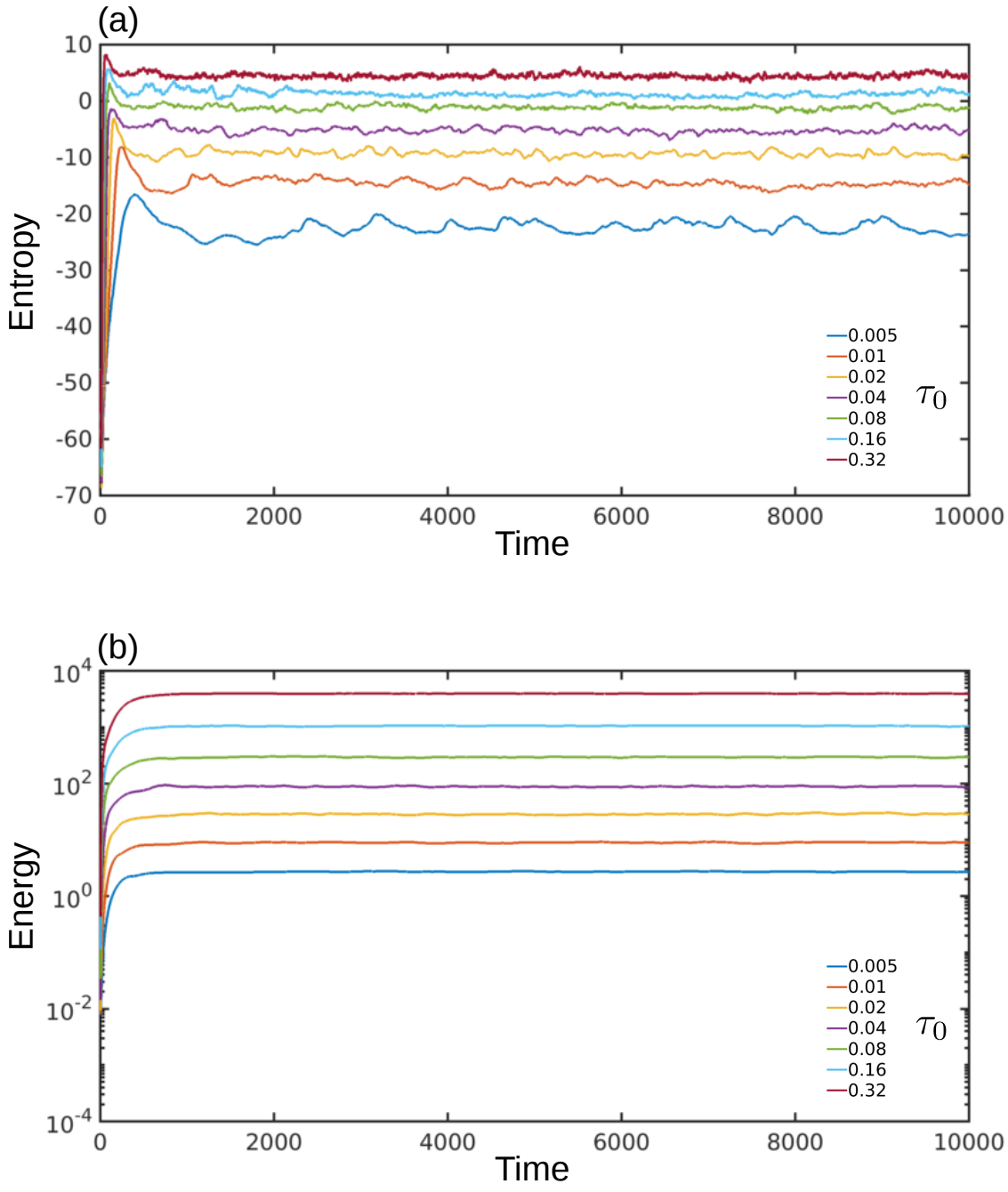

FIG. 5. (a) Evolution of entropy in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see that entropy behaves as a balanced quantity in a statistically steady state: in the time mean the sources of entropy are equal to the sinks of entropy. (b) Evolution of energy in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$ shown on a logarithmic scale. Like entropy the energy is balanced in a statistically steady state. The energy takes a longer time than the entropy to reach steady state balance.

VI. RELATION TO THE MAXIMUM ENTROPY PRINCIPLE

A. Time evolution of entropy and the maximum entropy principle

In what has been discussed so far we have considered the derivative of entropy with respect to time. Now we consider its relation to the maximum entropy principle which is at the core of the equilibrium statistical mechanics theory of ideal geophysical flow [17]. Although we will refer to the 'maximum' entropy principle it is important to note that the entropy may exhibit a minimum or stationary point and that classifying these stationary points lie beyond the scope of this study. In order to achieve clarity in our discussion we define the following terminology. For the purposes of this study it is useful to restrictively define *equilibrium* as a stationary entropy state of the ideal flow (i.e. MRS theory). For forced-dissipative systems, we refer to the long-time behaviour as a *(non-equilibrium) statistically steady state*.

The maximum entropy principle states that the entropy should be stationary with respect to variations in the probability distribution, ρ, given appropriate dynamical constraints. We can relate the time derivative of the entropy, S, to the functional derivative using the relation

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\frac{\delta S}{\delta\rho}. \tag{23}$$

Assuming that the system is in a stationary entropy state constrained by the value of energy and N polynomial Casimirs we have that the variational problem

$$\begin{aligned}\frac{\delta S}{\delta\rho} + \alpha(t)\frac{\delta}{\delta\rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho - E(t)\right)\\ - \sum_{n=1}^{N}\gamma_n(t)\frac{\delta}{\delta\rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho - C_n(t)\right) = 0,\end{aligned} \tag{24}$$

is satisfied, where α and γ_ns are Lagrange multipliers and where $E(t)$ and $C_n(t)$ are the time varying values of energy and Casimirs. Substituting this condition into (23), we relate the time derivative of entropy to that of energy and the Casimirs:

$$\frac{dS}{dt} = \alpha^*(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}, \tag{25}$$

where $\alpha^* \equiv -2\alpha$, and the detailed derivation of equation (25) is given in Appendix 3. This equation says that if the entropy is maximal under some constraints then the evolution of the entropy can be entirely explained by the evolution of the quantities constraining it. We

can split the time evolution of the Lagrange multipliers into temporal mean and fluctuations such that $\alpha^*(t) = \overline{\alpha^*} + \alpha^{*'}(t)$, and similarly for the other Lagrange multipliers. Assuming that the deviations in the Lagrange multipliers are small, which we will address in more detail in due course, and integrating (25), we can write an approximate relation for the time evolution of entropy in terms of the time evolution of energy and the Casimirs, giving

$$S(t) \approx \overline{\alpha^*} E(t) + \sum_{n=1}^{N} \overline{\gamma_n} C_n(t) + K, \tag{26}$$

where K is a constant of integration. This expression relies on two assumptions. Firstly, the entropy is maximized constrained by the value of energy and the Casimirs at any point in time; in other words, the system is in a *quasi-equilibrium state* defined as a non-equilibrium statistically steady state where the time scales for changes in the balanced quantities is larger that the time the eddies take to drive the system to stationary entropy state. Secondly, the fluctuations in the Lagrange multipliers (sensitivities of the entropy) are small. We now turn to testing the relation, (26), in order to test these assumptions.

B. Reconstruction of entropy evolution

We can test equation (26) by regressing the diagnosed time evolution of entropy onto the time evolution of energy and the other conserved quantities, the Casimirs, and comparing the reconstructed entropy time series against the diagnosed entropy time series. It was found the this procedure does not give a good reconstruction when the spin up is included in the time series, however there is good agreement when we only consider the statistically steady state: it is likely that the time dependence of the Lagrange multipliers is large during spin up but not in the statistically steady state.

An example is given in Figure 6 for simulation D_6. Figure 6 shows the reconstruction of the entropy time series using only the first two Casimirs, circulation and enstrophy, in addition to the energy as well as using Casimirs up to tenth order. The correlation for the first two Casimirs is 0.73, and when ten Casimirs are used the correlation is 0.93. This is a striking agreement and provides evidence that the turbulence acts to maximize entropy given time-varying constraints, according to equation (26), at each point in time in statistically steady state. That is, we find evidence that the system is in a quasi-equilibrium state as defined above. It must be noted that the quasi-equilibrium approximation is likely not to

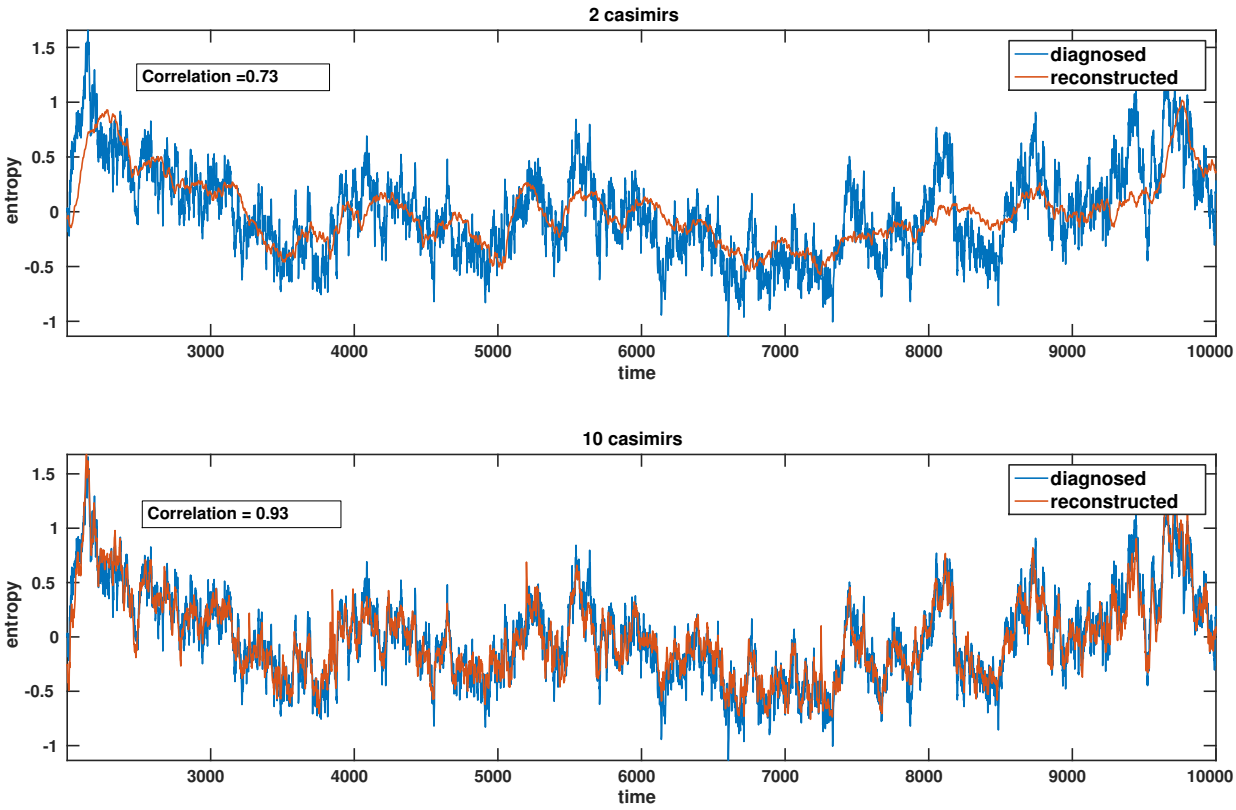

FIG. 6. Reconstruction of the entropy evolution using relation (26). *Top* - using only the first two Casimirs, circulation and enstrophy. *Bottom* - using first ten Casimirs. Both show significant correlation with the two Casimir reconstruction clearly matching the low frequency fluctuations, however but failing to capture the high frequency fluctuations. The ten Casimir reconstruction performs substantially better.

hold during spin-up.

The analysis has been repeated for all forced-dissipative simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$ and including differing numbers of Casimirs. Figure 7a shows the correlation between reconstructed and diagnosed entropy time series as a function of number of Casimirs for $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see a marked increase in the correlation with increased numbers of Casimirs. This shows the importance of higher order Casimirs in this statistical mechanics approach. It is also apparent that odd power Casimirs do not contribute significantly to increasing the correlation. We test the significance of these correlations by comparing with correlations produced by reconstructing the entropy from 'synthetic' time series for the energy and Casimirs. These 'synthetic' time series are produced randomly such that they have the same power spectra as the diagnosed time series. It is found that correlations derived from the model data lie

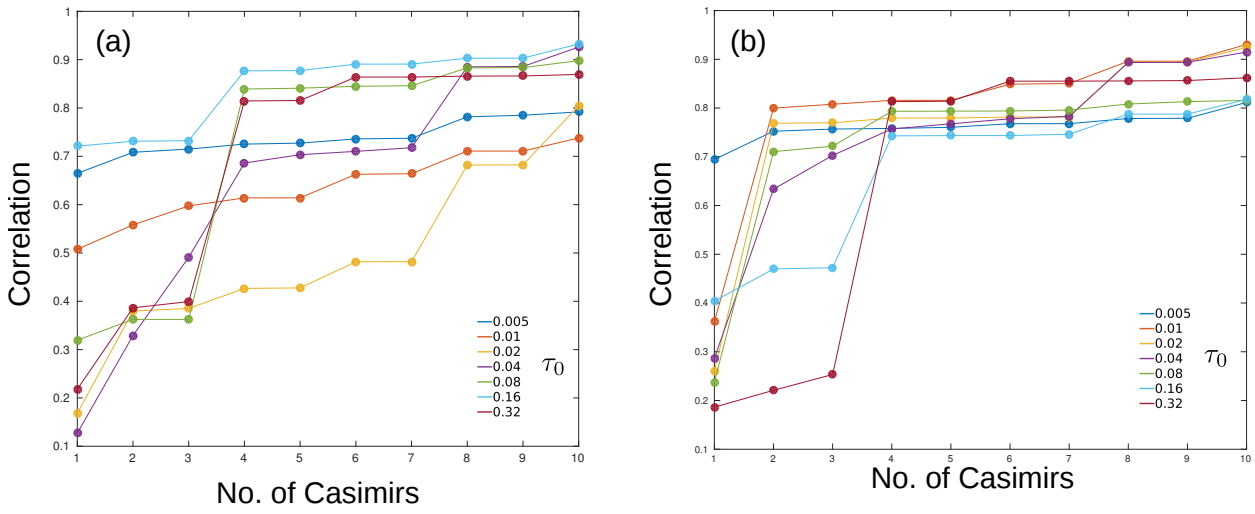

FIG. 7. (a) Correlation between the diagnosed and reconstructed entropy time series following the relationship (26). Entropy is calculated on zonal bands for simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation. (b) Correlation between the diagnosed and reconstructed entropy time series following the relationship (26). Entropy is calculated on instantaneous streamfunction for simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation but now all simulations show high correlation of greater than 0.7 with only four Casimirs. This is to be contrasted with a).

beyond the 99th percentile of an ensemble of 'synthetic' data. In this way we can reject the null-hypothesis that the observed correlations are due to statistical over-fitting. A detailed description of this significance test is given in Appendix 4.

We can repeat this analysis, now calculating the entropy along instantaneous streamfunction contours. This means that the macro-cells defined in Section III become contours of streamfunction rather than zonal band as has been used up to now. The entropy will change as a consequence of this transformation but the energy and Casimirs will not - this implies that only the projection in (26) of the entropy evolution onto the energy and Casimirs that will change. The correlation of the reconstructed entropy with the diagnosed entropy for this choice of macro-cell is shown in Figure 7b.

The main point to note here is that the correlation is higher, and converges quickly, for fewer Casimirs. The fourth order Casimir seems to be of particular importance with all simulations having a correlation of greater than 0.7 if four Casimirs are used. This

agrees with the observation of [10] that viewing potential vorticity distributions in a more of a Lagrangian sense acts to simplify the statistics, thus requiring a reduced number of moments to describe the local probability distribution of potential vorticity. The particular importance of the fourth order Casimir here is in contrast to the study of [1] who found that, in their numerical set-up, no Casimir above order three was relevant; this indicates that the particular Casimirs which are most important depend heavily on the particulars of any given system (i.e. domain geometry and nature of forcing).

The fact that it is possible, to a large extent, to reconstruct the statistically steady state time series for entropy using the corresponding dynamically balanced quantities of the flow is tantalizingly suggestive that in a statistically steady state of a forced-dissipative flow the turbulence pushes the system into a quasi-equilibrium: a non-equilibrium statistically steady state which nonetheless close to stationary entropy state at each point in time with time varying constraints. That is, although not in a true equilibrium (ideal flow with maximal entropy), the rate at which turbulence pushes the entropy to its allowed value is much faster than the time scales over which the conserved quantities are changing and is analogous to the quasi-static approximation in thermodynamics.

As we saw in Sections V A and V B, the eddies can act as a source or sink of entropy. Indeed, the fact that the balanced quantities such as energy (Figure 5b) fluctuate at all is due to the turbulence, if there were no non-linearities then we would have steady flow and no fluctuations. We suggest that the eddies play a double role, simultaneously maintaining the quasi-equilibrium and modulating its constraints. Further work, over a wider range of parameters, is required to obtain firmer evidence for the maximum entropy principle at work.

C. Solving for the Lagrange multipliers

Solving the variational problem, (24), gives us a probability distribution in terms of a set of unknown Lagrange multipliers, in this section we describe the method for determining these from knowledge of the energy and Casimirs of the flow. To determine the Lagrange multipliers it is necessary to solve the non-linear simultaneous equations

$$-2E = -\frac{\partial}{\partial \alpha} \int \mathrm{d}^2\mathbf{x} \, \ln \mathcal{Z}(\alpha, \gamma_1, ..., \gamma_N) \tag{27}$$

for the energy constraint and

$$C_n = -\frac{\partial}{\partial \gamma_n} \int \mathrm{d}^2\mathbf{x}\, \ln \mathcal{Z}(\alpha, \gamma_1, ..., \gamma_N) \tag{28}$$

for each Casimir constraint. Here $\mathcal{Z}$ is the local normalization, or the partition function, of the probability distribution given as

$$\mathcal{Z} = \int \mathrm{d}\tilde{q}\, \exp\left(-\alpha\langle\psi\rangle\tilde{q} - \sum_{i=n}^{N} \gamma_n \tilde{q}^n\right), \tag{29}$$

thus, $\mathcal{Z}$ is constructed such that

$$\int \mathrm{d}\tilde{q}\, \rho(\tilde{q}|\mathbf{x}) = 1. \tag{30}$$

Determining the Lagrange multipliers for given values of N constraints is numerically difficult and its solution is not tackled in this study.

However, we can proceed by reducing the dimensionality of the problem. Ironically this is achieved by first considering the case of infinite dimensions. Constraining the entropy of the flow the first N polynomial Casimirs of the flow is a truncated version of the exact constraint. To constrain by *all* Casimirs of the flow we constrain by the global potential vorticity distribution discussed in Section II. The constraint is given by

$$\Pi(\tilde{q}) = \int \mathrm{d}^2\mathbf{x}\, \rho(\tilde{q}|\mathbf{x}), \tag{31}$$

and the Lagrange multiplier becomes a function of $\tilde{q}$, $\gamma(\tilde{q})$. The corresponding variational problem

$$\frac{\delta S}{\delta \rho} + \alpha \frac{\delta}{\delta \rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\, \langle\psi\rangle\tilde{q}\rho - E\right) - \frac{\delta}{\delta \rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\, \gamma(\tilde{q})\rho - \Pi(\tilde{q})\right) = 0, \tag{32}$$

now gives the solution

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\left(-\alpha\langle\psi\rangle\tilde{q} - \gamma(\tilde{q})\right), \tag{33}$$

for the probability distribution. Substituting (33) into (31) we obtain the expression

$$\Pi(\tilde{q}) = e^{-\gamma(\tilde{q})} \int \mathrm{d}^2\mathbf{x}\, \frac{e^{-\alpha\langle\psi\rangle\tilde{q}}}{\mathcal{Z}}. \tag{34}$$

The integral here is a function of potential vorticity only and we can write $\gamma(\tilde{q})$ in terms of Π and the integral. This allows us to eliminate the Lagrange multiplier corresponding to the Casimir constraint leaving us with only α to find. Eliminating γ from (33), we obtain the

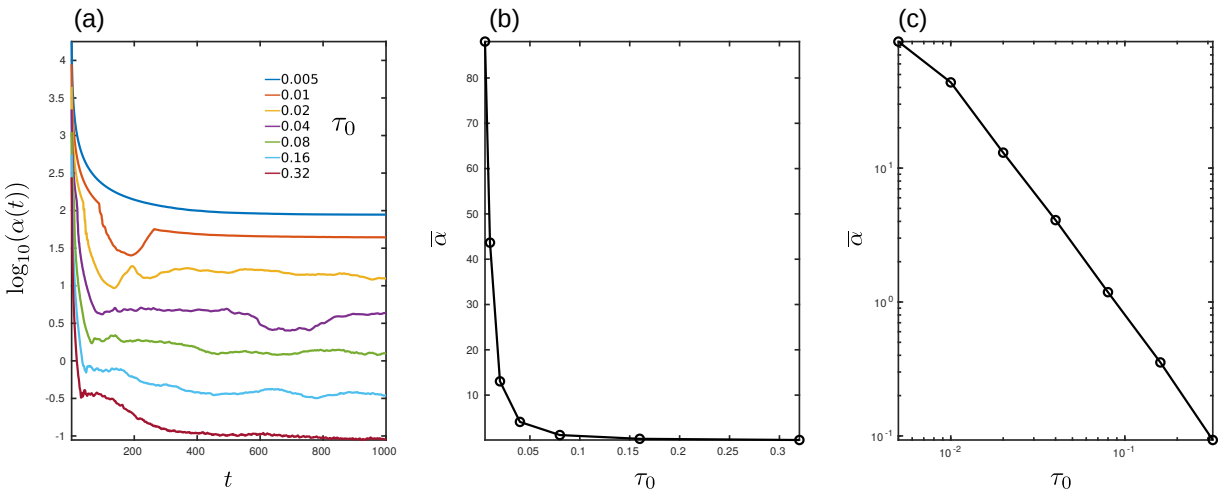

FIG. 8. (a) Shows the evolution of $\log_{10}\alpha$ with time. We see a very quick relaxation to a steady state value with fluctuation around this value in steady state being small. (b) Shows the steady state mean value of α against the strength of the wind stress. We see at low forcing have a very high sensitivity of entropy to energy while at high forcing the sensitivity is low. (c) Shows mean value of α against wind stress strength, now on a log-log scale. We see a near power-law behaviour with an approximate slope of -1.7.

probability distribution in terms of the Lagrange multiplier, α, and the global distribution, Π:

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}(\mathbf{x})\int \mathrm{d}^2\mathbf{x}\left(\frac{e^{-\alpha\langle\psi\rangle\tilde{q}}}{\mathcal{Z}(\mathbf{x})}\right)}\Pi(\tilde{q})e^{-\alpha\langle\psi\rangle\tilde{q}}. \tag{35}$$

A numerical method can be constructed for the two-dimensional problem of optimizing the values of α and $\mathcal{Z}$ simultaneously thus finding the maximum entropy distribution from only knowledge of global quantities. The details of the numerical method is given in Appendix 5 and the results of this methodology follow.

D. Reconstruction of statistics

In this section we reconstruct the statistics from the maximum entropy distribution, (35), by optimizing for the Lagrange multiplier, α. The resulting values for α in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$ are shown in Figure 8. We see that the value of α has a strong dependence on the strength of wind stress and time. Figure 8a shows the evolution of α during spin up, we see a strong reduction in the value of α at short times with the value of α fluctuating about

a statistically steady state value for long times. In steady state the fluctuation around the time-mean value is very small, this supports the assumption made in Section VI B to derive equation (26). Also shown is the dependence of the time-mean value of α on the wind stress strength in Figure 8b and c. The steady state sensitivity of entropy to energy is drastically decreased with the strength of wind stress suggesting that the entropy of the system becomes insensitive to perturbations in energy at high wind stress. Figure 8c displays a near power-law dependence of $\overline{\alpha}$ on τ_0 with an exponent of approximately -1.7, the explanation of this feature is left for future work.

Interpretations of the Lagrange multiplier, α, come with a caveat; the numerically determined value of α is only as accurate as the maximum entropy hypothesis and, in particular, the mean field approximation for the energy. To test the accuracy of these devices we compare the reconstructed statistics from the distribution, (35), and the diagnosed statistics from the numerical simulations.

Figure 9 shows diagnostics comparing the maximum entropy distribution evaluated using (35) with the statistics diagnosed from the numerical simulation itself, that is, the 'truth'. In short, despite displaying encouraging qualitative agreement with the simulations, the maximum entropy distribution does not fully capture the quantitative statistical details of simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$. Figure 9a shows a comparison between the reconstructed and diagnosed $\overline{\langle q \rangle} - \overline{\langle \psi \rangle}$ relation for simulation FD_6. The maximum entropy reconstruction shows good qualitative agreement but seems smoothed compared to the diagnosed relationship. Also for simulation FD_6, Figure 9b compares the $\overline{\mathrm{SD}(q)} - \overline{\langle \psi \rangle}$ relation and we see that the quantitative agreement is poorer. The standard deviation is underestimated in the centre of channel while being overestimated in the flanks. It must be noted that the maximum-entropy reconstruction of these statistical represents a drastic improvement over the energy-enstrophy constrained theory.

Figure 9c and d compare the diagnosed and reconstructed skewness and kurtosis for simulation FD_7. While giving qualitatively consistent features, the reconstructed skewness and kurtosis display significant deviations near to the boundary. In addition, Figure 9e and f compares the probability distribution, predicted and diagnosed, for simulations FD_4 and FD_5 at $y = 0$ and $y = 2.01$ respectively. In Figure 9e the trimodal nature of the distribution is successfully captured but we can see that the maximum entropy distribution overestimates the weight of the central peak compared to the side peaks. In Figure 9f, we again can see

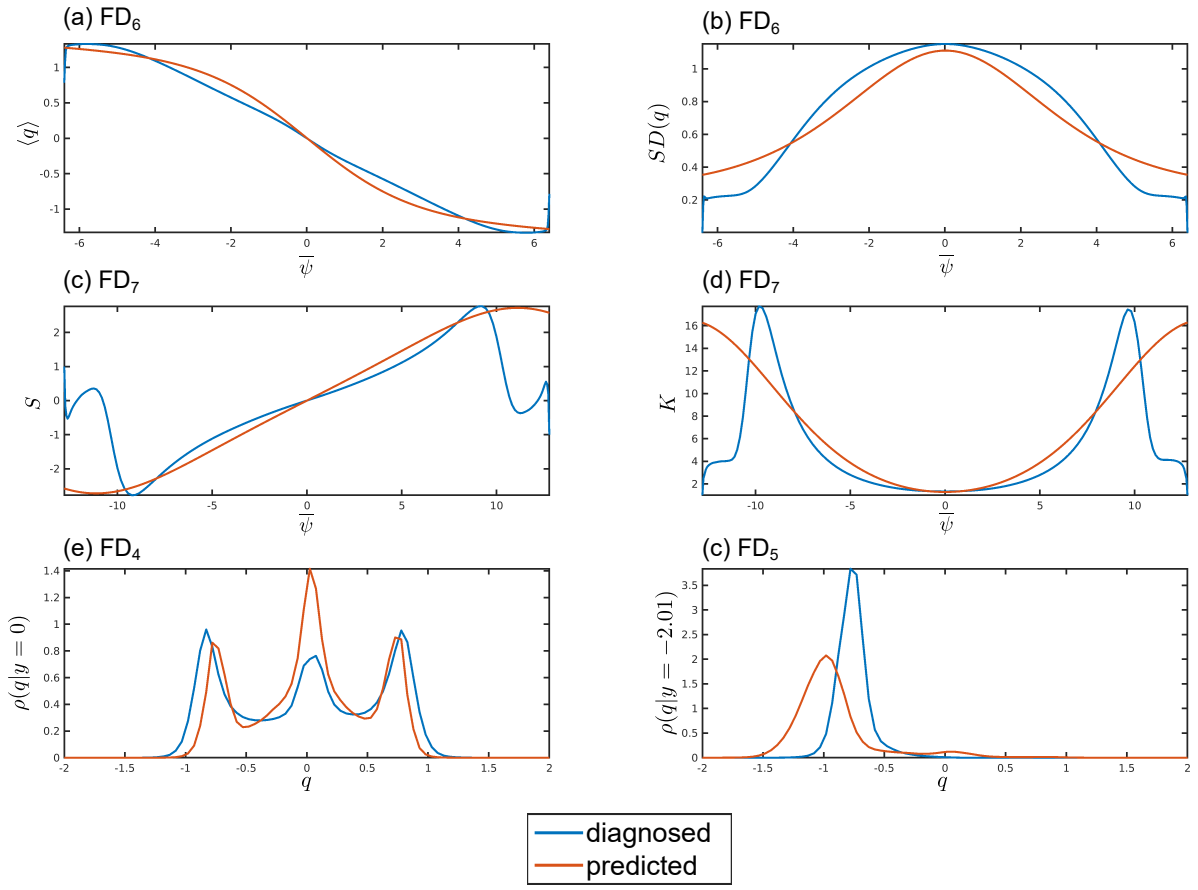

FIG. 9. Diagnostics testing reconstructed maximum entropy statistics. (a) Comparison of the time-averaged mean potential vorticity against time-averaged streamfunction for simulation FD_6. (b) Comparison of the time-averaged standard deviation of potential vorticity against time-averaged streamfunction for simulation FD_6. (c) Comparison of the time-averaged skewness of potential vorticity against time-averaged streamfunction for simulation FD_7. (d) Comparison of the time-averaged kurtosis of potential vorticity against time-averaged streamfunction for simulation FD_7. (e) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_4, distributions are evaluated for the centre of the channel, $y = 0$. (f) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_5, distributions are evaluated for the flank of the jet, $y = -2.01$.

a good qualitative agreement but in the details there are deficiencies. Many of the ways in which the maximum entropy statistics deviates from the simulations can be interpreted as a degradation of the strong persistent mixing barrier which the simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$ exhibit. The statistical nature and consequences of this mixing barrier has been extensively

studied, for this model, in [10].

In summary, the maximum entropy distribution, (35), produces an encouraging qualitative reproduction of the flow statistics. However, the quantitative reproduction is still lacking despite the suggestive evidence for the entropy being maximized presented in Section VI B. The derivation of the maximum entropy distribution relies on two main assumptions in the equilibrium, Miller-Robert-Sommeria, statistical mechanics: a) maximization of entropy; and b) the mean field approximation. We argue that the deficiencies of the maximum entropy distribution presented in this numerical experiment can be attributed to the breakdown of the mean field approximation in a forced dissipative system. This will be discussed in detail in the following section.

VII. THE MEAN-FIELD APPROXIMATION

A possible reason for the lack of success of the predicted probability distribution function, (35), is that the maximum entropy principle is not at work. Nevertheless, we believe that the analysis presented in Section VI B provides sufficient evidence to look for other reasons why (35) fails to quantitatively capture the statistics. In this section we consider the mean field approximation of Miller-Robert-Sommeria equilibrium statistical mechanics. The energy of the flow is given as

$$E[q] = -\frac{1}{2}\int \mathrm{d}\mathbf{x}\,\psi(\mathbf{x})(q(\mathbf{x}) - \beta y), \tag{36}$$

because of the presence of ψ it is not clear how make the necessary substitution to write E as a functional of ρ allowing us to tackle this constraint analytically. We rewrite the energy as

$$E[q] = -\frac{1}{2}\int\int \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'\,(q(\mathbf{x}) - \beta y)G(\mathbf{x},\mathbf{x}')q(\mathbf{x}') \tag{37}$$

where $G(\mathbf{x},\mathbf{x}')$ is the Green's function of the differential operator defined by, $q = \nabla^2\psi + \beta y$. Now, by swapping the potential vorticity field for its average value, and defining $\langle\tilde{q}\rangle \equiv \nabla^2\langle\psi\rangle + \beta y$, we get;

$$E_M[\rho] = -\frac{1}{2}\int \mathrm{d}\mathbf{x}\,\langle\psi\rangle(\langle\tilde{q}\rangle - \beta y). \tag{38}$$

This step is the *mean-field approximation*, often used in models of condensed matter physics. In essence, we are saying that, rather than considering the interaction energy between all

pairs of potential vorticity patches, we consider that each patch of potential vorticity only feels mean effect of all other patches.

For an ideal fluid in statistical equilibrium, the mean-field approximation ceases to be approximate and becomes exact due to the non-local aspect of the Green's function, G [6]; for example, two-dimensional vortex dynamics had a logarithmic Green's function leading to long-range vortex-vortex interactions. The mean-field formulation, however, necessitates two related properties of ideal equilibria: a) that the mean eddy potential vorticity flux is zero, $\nabla \cdot \langle \mathbf{u}' q' \rangle = 0$, which is a consequence of the equilibrium-state q-ψ relation, $\langle q \rangle = f(\langle \psi \rangle)$; and b) that neighbouring macro-cells of the flow are uncorrelated. These properties are fundamentally linked with the mean field approximation and are manifestly not satisfied in a forced-dissipative statistically steady state. Therefore, we suggest that while the maximization of entropy might be an useful organizing principle in forced-dissipative flow, the mean-field approximation remains only a crude approximation. We propose two potential avenues for future study to tackle this problem.

- It may be possible to find a coarse-graining (i.e., macro-cells) which is partially Lagrangian to reduce difference between $E[q(\mathbf{x})]$ and $E_M[\rho(\tilde{q}|\mathbf{x})]$. This is the approach taken in [17] which produces good agreement between experiment and equilibrium theory by moving into a frame of reference moving at the phase speed of a large scale Rossby wave. However, this is unlikely to work in the presence of multiple wave modes as is the case with simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$, see discussion in [10].

- Alternatively, we speculate that a perturbation method could be applied to the mean-field approximation to yield a more realistic probability distribution. However, precisely how this might be achieved remains to be determined and is subject of future work.

Indeed, it may be that a combination of the two approaches will provide the means of deriving forced-dissipative statistics from the maximum entropy principle.

VIII. CONCLUSION

In this study we have shown how an eddy mixing entropy can be used as a measure of turbulent disorder. By deriving the influence of forcing and linear drag, we were able

to use entropy to describe the turbulence in a freely-decaying and forced-dissipative flow. The evolution of entropy describes the three stages of the eddy life cycle and eddy-mean interaction: growth of instability, formation of large scale coherent structures and steady state fluctuations. In particular, the eddy production of entropy, which has been focus of much theoretic inquiry, can be explicitly computed from data. The fact that the eddy-mixing entropy behaves in a dynamically balanced way is not a priori clear and provides a novel means of quantifying turbulent disorder in geophysical flows. This study of the temporal evolution of entropy can inform work on stochastic parametrization by describing the disorder in a turbulent jet in a way that links to both statistical physics and information theory; we can begin to piece together a picture of the emergent physics of the total entropy whereas the local entropy of each macro-cell is related to the information content of the sub-gridscale statistics.

The relationship between the temporal evolution of entropy and the maximum entropy principle was considered in Section VI. Under the assumption of maximum entropy it was found that the time evolution of entropy was set by the time evolution of its constraints. Suggestive evidence was found that the entropy is maximized in the model simulations considered in this study. It is clear that if a variational problem can be used to infer the statistics then the number of Casimir constraints has to be large.

With this evidence for the maximum entropy principle being a physically meaningful candidate for describing the behaviour of turbulence in the system studied here, we considered the problem of inferring the sub-grid scale statistics. This is equivalent to inferring the Lagrange multipliers, used in the maximum entropy variational problem, from the constraints applied. We presented the mathematical formulation of this problem in Section VI and showed how the dimensionality could be reduced given knowledge of the global potential vorticity distribution. Further, we reconstructed the maximum entropy statistics from knowledge of the energy, global potential vorticity distribution, and zonal mean streamfunction as a functions of time. We find that although the maximum entropy statistics reproduces qualitatively representative features of the flow, quantitative agreement is lacking, especially for higher order statistical moments. In Section VII, the mean-field approximation was discussed as a potential culprit for the quantitative disagreement and avenues for future investigation were proposed.

In this study we have presented eddy-mixing entropy as both a descriptive tool and a

dynamically balanced quantity in a barotropic turbulent jet. We have also demonstrated the relationship between the statistical mechanics of forced-dissipative flow and well-known globally balanced quantities such as the energy and enstrophy of the flow. In doing so we were able to provide evidence for the action of the maximum entropy principle at work in a forced-dissipative system. The question of the usefulness of statistical mechanics theories, such as the Miller-Robert-Sommeria theory, in understanding the statistically steady states of two-dimensional and geophysical turbulence has received much attention [e.g. 11, 17, 32, 47] but remains somewhat unclear. By explicitly considering the evolution of eddy mixing entropy in a forced-dissipative model we are able to demonstrate the importance and utility of eddy mixing entropy in the study of forced-dissipative geophysical turbulence, opening the door to revisiting the application of statistical mechanics to ocean mesoscale eddy parameterizations [cf. 18].

ACKNOWLEDGMENTS

This work is funded by the UK Natural Environment Research Council. We would like to thank Antoine Venaille for insightful and helpful discussion of our work, as well as Chris O'Reilly for his assistance with the significance testing.

Appendix

1. Numerical computation of entropy

It is important to note the difference between the discrete, or Shannon, entropy

$$S = -\sum_i \rho_i \ln \rho_i, \tag{A.1}$$

and the continuous, or differential, entropy we use in this study

$$S = -\int \mathrm{d}x\, \rho(x) \ln \rho(x). \tag{A.2}$$

One of the clear differences between these two entropies is that the continuous entropy can become negative whereas the discrete entropy is never less than zero. On the other hand the continuous entropy can be negative and indeed tends to negative infinity for the asymptotic

limit of a delta-function. This means that we need to be careful when numerically evaluating an estimator for the continuous entropy. Naïvely using the standard discrete approximation for the integral in (A.2) leads to calculating a quantity proportional to the discrete entropy, (A.1). To find an approximation for, (A.2), we must evaluate the quantity

$$S \approx -\Delta x \sum_i \rho_i \ln \rho_i + \ln \Delta x, \tag{A.3}$$

where ρ_i becomes a histogram approximation to the distribution $\rho(x)$. However, this method of approximation was found to be biased and introduced a systematic error into the results presented in this study.

Instead, we used a sample-spacing estimator for the distribution leading to an improved numerical approximation for the continuous entropy [2, 43]. The sample-spacing estimator relies of the idea that when the data is ordered, from smallest to largest value and represented by the list $\{x^{(i)}\}$, then the reciprocal of the difference between two samples, separated by m, spaces is an estimator for the probability density. That is we approximate ρ by

$$\rho \propto \frac{1}{x^{(i+m)} - x^{(i)}}. \tag{A.4}$$

Substituting this into (A.2) and using appropriate normalizations, the following expression is found for the entropy

$$S \approx \frac{1}{N-m} \sum_{i=1}^{N-m} \ln\left(\frac{N+1}{m}(x^{(i+m)} - x^{(i)})\right). \tag{A.5}$$

Here, N is the number of samples; m is the spacing size; and $x^{(i)}$ represents the i^{th} ordered sample. Following [43] we use the fact that the optimal of m is around $\sqrt{N}$. This method was used to evaluate the entropy throughout this study and is found to be considerably better than more naïve methods. The simpler methods proved to have a strong dependence on the choice of histogram bin width rendering them unusable for quantitative comparison with theory.

2. Detailed derivation of the tendency equation for entropy

Ignoring the non-local and non-linear terms the potential vorticity equation is

$$\frac{\partial q}{\partial t} = -r(q - \beta y) - g(y), \tag{A.6}$$

and the corresponding probability distribution satisfies the equation,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \tilde{q}}[r(\tilde{q} - \beta y)\rho] - g(y)\frac{\partial \rho}{\partial \tilde{q}}. \tag{A.7}$$

We can derive the entropy tendency by substituting equation (A.7) into

$$\frac{dS}{dt} = -\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\rho \ln \rho, \tag{A.8}$$

$$= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left[\frac{\partial \rho}{\partial t}\ln\rho + \frac{\partial \rho}{\partial t}\right], \tag{A.9}$$

$$= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial \rho}{\partial t}\ln\rho - \frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\rho, \tag{A.10}$$

$$= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial \rho}{\partial t}\ln\rho, \tag{A.11}$$

where the final step comes from the normalization condition on the distribution, ρ. We rewrite (A.7) as

$$\frac{\partial \rho}{\partial t} = r\frac{\partial(\tilde{q}\rho)}{\partial \tilde{q}} - h(y)\frac{\partial \rho}{\partial \tilde{q}}, \tag{A.12}$$

where

$$h(y) = r\beta y + g(y). \tag{A.13}$$

Substituting, we obtain

$$\dot{S} = -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(\left(-h(y)\frac{\partial \rho}{\partial \tilde{q}} + r\frac{\partial(\tilde{q}\rho)}{\partial \tilde{q}}\right)\ln\rho\right). \tag{A.14}$$

We take each term separately in order to simplify the above expression. The first term is simplified, integrating by parts, as

$$\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(-h(y)\frac{\partial \rho}{\partial \tilde{q}}\ln\rho\right) = \int \mathrm{d}^2\mathbf{x}\,h(y)\int \mathrm{d}\tilde{q}\,\frac{\partial \rho}{\partial \tilde{q}}\ln\rho, \tag{A.15}$$

$$= -\int \mathrm{d}^2\mathbf{x}\,\left(h(y)\left([\rho\ln\rho]_{-\infty}^{+\infty} - \int \mathrm{d}\tilde{q}\,\frac{\partial \rho}{\partial \tilde{q}}\right)\right) \tag{A.16}$$

$$= 0. \tag{A.17}$$

We can see that the integral above is zero from the fact that both ρ and $\rho\ln\rho$ vanish at $\pm\infty$. The second term can also be simplified, integrating by parts twice, as

$$\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(r\frac{\partial(\tilde{q}\rho)}{\partial \tilde{q}}\ln\rho\right) = r\int \mathrm{d}^2\mathbf{x}\,\left([\tilde{q}\rho\ln\rho]_{-\infty}^{+\infty} - \int \mathrm{d}\tilde{q}\,\tilde{q}\frac{\partial \rho}{\partial \tilde{q}}\right) \tag{A.18}$$

$$= -r\int \mathrm{d}^2\mathbf{x}\,\left([\tilde{q}\rho]_{-\infty}^{+\infty} - \int \mathrm{d}\tilde{q}\,\rho\right) \tag{A.19}$$

$$= Ar, \tag{A.20}$$

where A is the area of the domain. Here, we have, in addition to the appropriate boundary conditions, used the normalization condition for ρ. Thus, the corresponding equation for the entropy tendency becomes

$$\dot{S} = P - Ar, \tag{A.21}$$

where we have reintroduced the non-local and non-linear terms as a residual via the advective production of entropy, P.

3. Time-evolution of maximal entropy

Suppose that entropy is maximized, constrained by time-varying energy and N polynomial Casimirs. The variational problem to be solved is given by

$$\frac{\delta S}{\delta \rho} + \alpha(t)\frac{\delta}{\delta \rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho - E(t)\right) - \sum_{n=1}^{N}\gamma_n(t)\frac{\delta}{\delta\rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho - C_n(t)\right) = 0. \tag{A.22}$$

taking the functional derivative of the constraints we obtain

$$\frac{\delta S}{\delta \rho} = \alpha(t)\langle\psi\rangle\tilde{q} + \sum_{n=1}^{N}\gamma_n(t)\tilde{q}^n\rho. \tag{A.23}$$

To find the entropy tendency, we substitute this variational problem into the relation

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\frac{\delta S}{\delta\rho}, \tag{A.24}$$

giving

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\left[\alpha(t)\langle\psi\rangle\tilde{q} + \sum_{n=1}^{N}\gamma_n(t)\tilde{q}^n\rho\right]. \tag{A.25}$$

Rearranging and pulling out time derivatives we obtain

$$\frac{dS}{dt} = \alpha(t)\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho + \sum_{n=1}^{N}\gamma_n(t)\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho, \tag{A.26}$$

and identifying the integrals with the constraints which they are, by construction, equal to, we obtain

$$\frac{dS}{dt} = -2\alpha(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}. \tag{A.27}$$

For ease of presentation we define $\alpha^* \equiv -2\alpha$, completing the derivation for the entropy tendency

$$\frac{dS}{dt} = \alpha^*(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}. \tag{A.28}$$

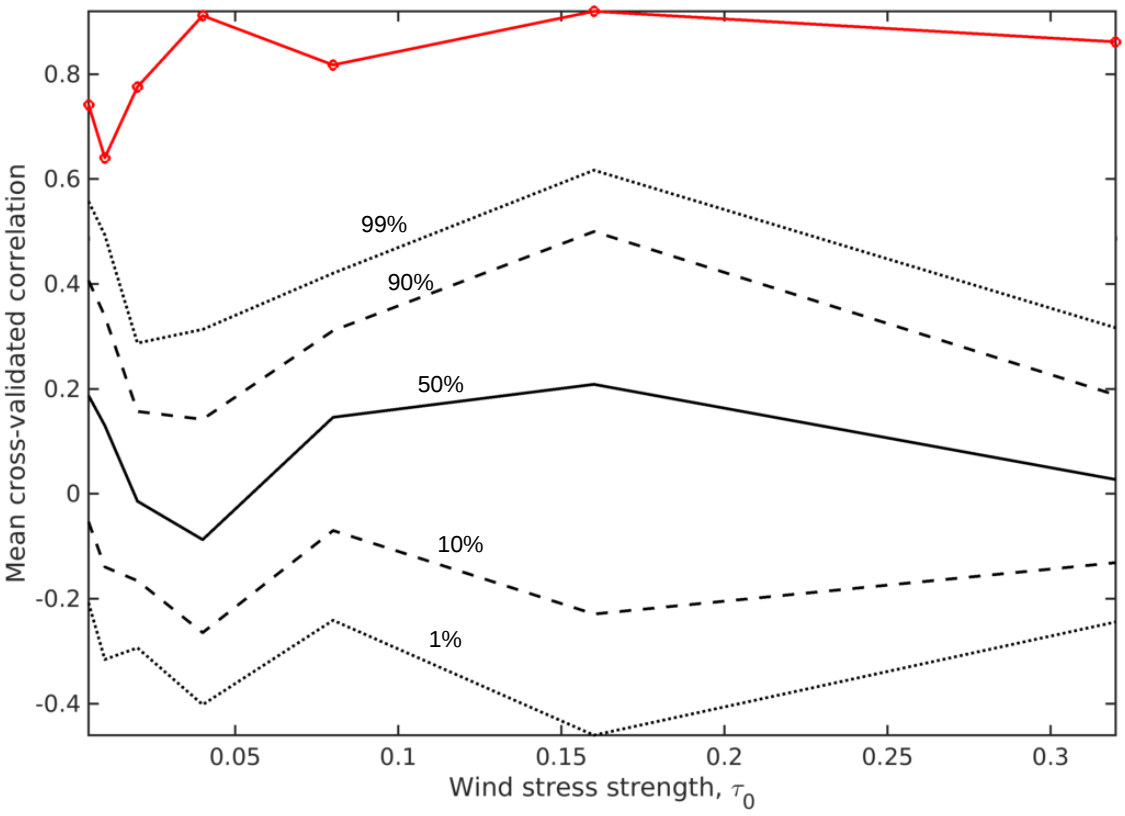

FIG. 10. *Red line* - mean cross-validated correlation between diagnosed and reconstructed entropy as a function of wind stress strength. *Black lines* - percentiles for the ensemble of mean cross-validated correlations from 'synthetic' time series of energy and Casimirs.

4. Significance testing

To exclude the possibility of statistical over-fitting of the entropy steady state fluctuations, in Section VI B, we perform a significance test for the correlations observed. Firstly, cross-validate the regression by using half the data to train the regression while validating on the other half. We also create an ensemble of 'synthetic' time series which have the same power spectra as the diagnosed energy and Casimir time series. This is done using a MATLAB function, ebisuzaki.m, which is available as part of the WEACLIM toolbox by Vincent Moron, available on the MathWorks file exchange. This method follows the significance test of [12]. We test the case using energy and 10 Casimirs to reconstruct to entropy evolutions, we cross-validate both the reconstruction by diagnosed time series as well as the ensemble of reconstructions from 'synthetic' time series. The results are given in Figure 10, where we

have plotted the mean cross-validated correlation for the diagnosed time series, $\overline{c}$, given as

$$\overline{c} = \frac{c_1 + c_2}{2}, \tag{A.29}$$

where c_1 is the correlation between the reconstructed and diagnosed entropy for the second half of the data, trained on the first half; c_2 is defined vice versa. Also plotted are the percentiles for the ensemble of 1000 mean cross-validated correlations produced from the 'synthetic' time series. This clearly shows that the observed correlations lie beyond the 99th percentile of the ensemble and excludes the possibility of statistical over-fitting.

5. Reconstructing maximum entropy statistics

We have seen from equation (35) that the partition function can be written as

$$\mathcal{Z}(\mathbf{x}) = \int \mathrm{d}\tilde{q} \left[\frac{\Pi(\tilde{q}) e^{-\alpha \langle \psi \rangle \tilde{q}}}{\int \mathrm{d}^2\mathbf{x} \left(\frac{e^{-\alpha \langle \psi \rangle \tilde{q}}}{\mathcal{Z}(\mathbf{x})} \right)} \right], \tag{A.30}$$

which is an implicit equation for $\mathcal{Z}$. Writing $R(\tilde{q}, \mathbf{x}) = \exp(-\alpha \langle \psi \rangle \tilde{q})$ we can recast equation (A.30) as an iterative relation for the $i+1$ approximation for the partition function

$$\mathcal{Z}_{i+1}(\mathbf{x}) = \int \mathrm{d}\tilde{q} \left[\frac{\Pi(\tilde{q}) R(\tilde{q}, \mathbf{x})}{\int \mathrm{d}^2\mathbf{x} \left[\frac{R(\tilde{q}, \mathbf{x})}{\mathcal{Z}_i(\mathbf{x})} \right]} \right]. \tag{A.31}$$

Starting with $\mathcal{Z}_1 = 1$, we can produce an approximation for the partition function which normalizes the probability distribution, ρ, to a better than 1% accuracy after 20 iterations.

Using this methodology we are able to reduce the problem of determining the Lagrange multipliers to a one-dimensional problem of determining α by optimizing for the mean-field energy, E_M. Using a MATLAB numerical minimization function (`fminsearch`), we can optimize α in order to match the energy of the flow, that is, by minimizing the cost function

$$L = \left| E_M - \frac{1}{2} \int \mathrm{d}^2\mathbf{x} \, \langle \psi \rangle \langle \tilde{q} \rangle \right|, \tag{A.32}$$

for given E_M, N and $\langle \psi \rangle$ which are diagnosed from the simulations.

---

[1] Abramov, R. V. and Majda, A. J. (2003). Statistically relevant conserved quantities for truncated quasigeostrophic flow. *PNAS*, 100(7):3841–3846.

[2] Beirlant, J., Dudewicz, E. J., Gyöfi, L., and van der Meulen, E. C. (1997). Nonparametric entropy estimation: an overview. *International Journal of Mathematical and Statistical Sciences*, 6(1):17–39.

[3] Berloff, P. S. (2005). Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529:71–95.

[4] Bouchet, F. and Simonnet, E. (2009). Random changes of flow topology in two-dimensional and geophysical turbulence. *Physical review letters*, 102(9):094504.

[5] Bouchet, F. and Sommeria, J. (2002). Emergence of intense jets and jupiter's great red spot as maximum-entropy structures. *Journal of Fluid Mechanics*, 464:165–207.

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

[7] Bretherton, F. P. and Haidvogel, D. B. (1976). Two-dimensional trubulence above topography. *Journal of Fluid Mechanics*, 78:129–154.

[8] Chavanis, P. H. (2009). Dynamical and thermodynamical stability of two-dimensional flows: variational principles and relaxation equations. *The European Physical Journal B*, 70(1):73–105.

[9] Corvellec, M. (2012). *Phase transitions in two-dimensional and geophysical turbulence*. Theses, Ecole normale supérieure de lyon - ENS LYON.

[10] David, T. W., Marshall, D. P., and Zanna, L. (2017). The statistical nature of turbulent barotropic jets. *Ocean Modelling*, 113:34–49.

[11] Dritschel, D. G., Qi, W., and Marston, J. B. (2015). On the late-time behaviour of a bounded, inviscid two-dimensional flow. *Journal of Fluid Mechanics*, 783:1–22.

[12] Ebisuzaki, W. (1997). A method to estimate the statistical signifiance of a correlation when the data are serially correlated. *Journal of Climate*, 10:2147–2153.

[13] Esler, J. G. (2008a). Robust and leaky transport barriers in unstable baroclinic flows. *Physics of Fluids*, 20(11):116602.

[14] Esler, J. G. (2008b). The turbulent equilibration of an unstable baroclinic jet. *Journal of Fluid Mechanics*, 599:241–268.

[15] Esler, J. G. and Haynes, P. H. (1999). Baroclinic wave breaking and the internal variability of the tropospheric circulation. *Journal of the Atmospheric Sciences*, 56(23).

[16] Grooms, I., Majda, A. J., and Smith, K. S. (2015). Stochastic superparameterization in a

quasigeostrophic model of the Antarctic Circumpolar Current. *Ocean Modelling*, 85:1–15.

[17] Jung, S., Morrison, P. J., and Swinney, H. L. (2006). Statistical mechanics of two-dimensional turbulence. *Journal of Fluid Mechanics*, 554:433–456.

[18] Kazantsev, E., Sommeria, J., and Verron, J. (1998). Subgrid-scale eddy parameterization by statistical mechanics in a barotropic ocean model. *Journal of Physical Oceanography*, 28:1017–1042.

[19] Kraichnan, R. H. (1967). Inertial ranges in two-dimentional turbulence. *The Physics of Fluids*, 10(7):299–303.

[20] Kraichnan, R. H. and Montgomery, D. (1980). Two-dimentional turbulence. *Reports on Progress in Physics*, 43:547–619.

[21] Lynden-Bell, D. (1967). Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. Astr. Soc.*, 136:101–121.

[22] Majda, A. and Wang, X. (2006). *Nonlinear dynamics and statistical theories for basic geophysical flows*. Cambridge University Press.

[23] McWilliams, J. C. (1977). A note on a consistent quasigeostrophic model in a multiply connected domain. *Dynamics of Atmospheres and Oceans*, 1:427–441.

[24] Merryfield, W. J. (1998). Effects of stratification on quasi-geostrophic inviscid equilibria. *Journal of Fluid Mechanics*, 354:345–356.

[25] Michel, J. and Robert, R. (1994). Large deviations for young measures and statistical mechanics of infinite dimensional dynamical systems with conservation law. *Communications in Mathematical Physics*, 159:195–215.

[26] Miller, J. (1990). Statistical mechanics of euler equations in two dimensions. *Physical Review Letters*, 65(17):2137–2140.

[27] Miller, J., Weichman, P. B., and Cross, M. C. (1992). Statistical mechanics, euler's equation, and jupiter's red spot. *Phys. Rev. A*, 45:2328–2359.

[28] Onsager, L. (1949). Statistical hydrodynamics. *Nuovo Cimento*, 6(2):249–286.

[29] Phillips, N. A. (1954). Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus*, 6:273–286.

[30] Porta Mana, P. G. L. and Zanna, L. (2014). Towards a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, 79.

[31] Prieto, R. and Schubert, W. H. (2001). Analytical predictions for zonally symmetric equilib-

rium states of the stratospheric polar vortex. *Journal of the atmospheric sciences*, 58(18):2709–2728.

[32] Qi, W. and Marston, J. B. (2014). Hyperviscosity and statistical equilibria of euler turbulence on the torus and the sphere. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(7):P07020.

[33] Robert, R. (1990). Etats d'équilibre pour l'écoulement bidimensionnel d'un fluide parfait. *Comptes Rendus de l'Academie des Sciences 1*, 311:481–515.

[34] Robert, R. (1991). A maximum-entropy principle for two-dimensional perfect fluid mechanics. *Journal of Statistical Physics*, 65:531–553.

[35] Robert, R. and Sommeria, J. (1991). Statistical equilibrium states for two-dimensional flows. *Journal of Fluid Mechanics*, 229:291–310.

[36] Salmon, R. (1998). *Lectures on Geophyical Fluid Dynamics*. Oxford University Press.

[37] Salmon, R. (2012). Statistical mechanics and ocean circulation. *Commun. Nonlinear Sci. Numer. Simulat.*, 17:2144–2152.

[38] Salmon, R., Holloway, G., and Hendershott, M. C. (1976). The equilibrium statistical mechanics of simple quasi-geostrophic models. *Journal of Fluid Mechanics*, 75:691–703.

[39] Schecter, D. A. (2003). Maximum entropy theory and the rapid relaxation of three-dimensional quasi-geostrophic turbulence. *Physical Review E*, 68(6):066309.

[40] Sommeria, J. (2001). Two-dimensional turbulence. In *New trends in turbulence Turbulence: nouveaux aspects*, pages 385–447. Springer.

[41] Treguier, A. M. (1989). Topographically generated steady currents in barotropic turbulence. *Geophysical and Astrophysical Fluid Dynamics*, 47:43–68.

[42] Turkington, B., Majda, A., Haven, K., and DiBattista, M. (2001). Statistical equilibrium predictions of jets and spots on jupiter. *Proceedings of the National Academy of Sciences*, 98(22):12346–12350.

[43] Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society, Series B*, 38:54–59.

[44] Venaille, A. (2012). Bottom-trapped currents as statistical equilibrium states above topographic anomalies. *Journal of Fluid Mechanics*, 699:500–510.

[45] Venaille, A. and Bouchet, F. (2011). Oceanic rings and jets as statistical equilibrium states. *Journal of Physical Oceanography*, 41:1860–1870.

774 [46] Venaille, A., Vallis, G. K., and Griffies, S. (2012). The catalytic role of the beta effect in
775 barotropization processes. *Journal of Fluid Mechanics*, 709:490–515.
776 [47] Wang, J. and Vallis, G. K. (1994). Emergence of Foffonoff states in inviscid and viscous ocean
777 circulation models. *Journal of Marine Research*, 52:83–127.
778 [48] Yasuda, Y., Bouchet, F., and Venaille, A. (2017). A new interpretation of vortex-split
779 stratospheric sudden warmings in terms of equilibrium statistical mechanics. *arXiv preprint*
780 *arXiv:1702.03716*.
781 [49] Zanna, L., Mana, P. P., Anstey, J., David, T., and Bolton, T. (2017). Scale-aware deterministic
782 and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, 111:66 – 80.