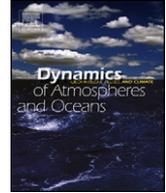




ELSEVIER

Contents lists available at ScienceDirect

Dynamics of Atmospheres and Oceans

journal homepage: www.elsevier.com/locate/dynatmoce

Uncertainty generation in deterministic flows: Theory and application with an atmospheric jet stream model

X. San Liang^{a,b,c,*}^a China Institute for Advanced Study, Central University of Finance and Economics, Beijing 100081, China^b State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou 310012, China^c Department of Mathematics and Systems Science, National University of Defense Technology, Changsha 410073, China

ARTICLE INFO

Available online 13 April 2011

Keywords:

Local entropy generation
Barotropic instability
Ensemble prediction
Optimal excitation
Frobenius–Perron operator

ABSTRACT

How uncertainties are generated in deterministic geophysical fluid flows is an important but mostly overlooked subject in the atmospheric and oceanic research. In this study, it is shown that the generating mechanisms include local entropy generation (LEG) and cumulant information transfer, both of which are explicitly expressed with the aid of a theorem established herein. To a system the former is intrinsic, representing the evolutionary trend of a marginal entropy and bringing connections between the two physical notions namely uncertainty and instability. The latter results from the interaction between different locations through dynamic event synchronization, and appears only in the course of state evolution. Although in practice it is a notoriously difficult task to estimate entropy and entropy-related quantities for atmospheric and oceanic systems, which are in general of large dimensionality, estimation of the LEG can be accurately fulfilled with ensembles of limited size. If, furthermore, the processes of a system under consideration are quasi-ergodic and quasi-stationary, its LEG actually can be fairly satisfactorily estimated even without appealing to ensemble predictions. These assertions are illustrated and verified in an application with two simulated quasi-geostrophic jet streams with compact chaotic attractors, one global over the whole domain and another highly localized. The LEG study provides an objective way of rapid assessment for predictions, which is important in the practical fields such as adaptive sampling and adaptive modeling.

© 2011 Elsevier B.V. All rights reserved.

* Current address: 7-321, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA.
Tel.: +1 617 324 5526.

E-mail address: xsliang@mit.edu

1. Introduction

Ever since the discovery of chaos by Lorenz (1963), uncertainty has become one of the major concerns in the field of numerical weather forecasts and ocean predictions. Numerous studies have been conducted in this regard; among them are Epstein (1969), Leith (1974), Carnevale and Holloway (1982), Farrell (1990), Nese et al. (1996), Ehrendorfer and Tribbia (1997), Toth and Kalnay (1993), Roache (1997), Shukla (1998), Schneider and Griffies (1999), Moore (1999), Smith et al. (1999), Palmer (2000), Kleeman (2002, 2008), Mu et al. (2002), Miller and Ehret (2002), Kirwan et al. (2003), DelSole (2004), Lermusiaux (2006), Tang et al. (2008), to name several. Given a model, uncertainties may result from boundary conditions, initial conditions, geometric specifications, dynamical parameters, numerical schemes, model inaccuracies, random external forcings, etc. But what makes uncertainty study such an extensive arena of research is the intrinsic source that is embedded in nonlinear dynamical systems. As observed by Lorenz (1963), for a highly nonlinear system like the atmosphere, even the model is deterministic, its outcome may appear chaotic and is hence intrinsically unpredictable beyond some time limit. The present study concerns about how uncertainties are generated within deterministic flows, a subject not only of theoretical interest from the viewpoint of ensemble dynamics, but also of practical importance in that, for example, it provides an objective way of model performance assessment, and offers guidance for adaptive sampling (Lermusiaux, 2007).

A natural way to quantify uncertainty is with entropy, i.e., entropy in the Shannon sense (see Cove and Thomas, 1991). (Note here we are not dealing with predictability, an area where one has to use relative entropy; see Kleeman, 2002.) Uncertainty generation is thus about how entropy is changed. Given a dynamical system, autonomous or non-autonomous,

$$\frac{d\mathbf{u}}{dt} = \mathbf{F}(\mathbf{u}, t), \quad \mathbf{u} = (u_1, u_2, \dots, u_n), \quad (1)$$

Liang and Kleeman (2005) showed and later on rigorously proved (Liang and Kleeman, 2007b) that the joint entropy H of (u_1, u_2, \dots, u_n) (units: nats) evolves as

$$\frac{dH}{dt} = E(\nabla \cdot \mathbf{F}), \quad (2)$$

where E stands for mathematical expectation, and $\nabla \cdot$ is the divergence operator with respect to \mathbf{u} . This remarkably concise law, which holds for systems of arbitrary dimensionality, states that the global uncertainty evolution of a system is completely controlled by the expansion/contraction of its phase space. Particularly, if the flow of the vector field of a system is divergence free, then its uncertainty is invariant. So if the whole ocean is taken as one sample space and particle trajectories are considered, then H must be conserved, as the sea water is incompressible. If the same study is performed with the atmosphere (e.g., pollutant dispersion), apparently this is not true. But if one examines the problem in an isobaric coordinate frame, in which the atmosphere is incompressible, H then must also be conserved.

For real atmospheric/oceanic problems, the systems of concern are usually not treated as a whole. Correspondingly, uncertainty in this sense is a local notion; that is to say, it has a distribution over the domain where it is defined. Given a state variable \mathbf{u} with components (u_1, u_2, \dots, u_n) , the uncertainty of u at location k is measured by the marginal entropy of u_k , written H_k . The uncertainty generation problem is thus essentially about how H_k changes as time goes on. Through studying the distribution of dH_k/dt in k we are able to identify the uncertainty sources/sinks, which will allow us to understand whether, when, and where a model may amplify the existing uncertainties for a state variable.

To find the time rate of change of a marginal entropy, one might think of first computing the marginal entropy directly with the probability density functions (pdf) from a given ensemble of predictions, and then computing its time variation. This is, however, notoriously difficult in atmosphere-ocean research, as an accurate estimation of the pdf requires a densely sampled state space, while atmospheric and oceanic systems are usually of huge dimensionality, allowing for only limited ensembles. A simple calculation may yield some hint about the difficulty: For a system with only 6 EOF modes, assuming 5 draws for each mode (a minimal requirement as argued by Kleeman, 2007) one obtains $5^6 = 15,625$ members for the ensemble, which is already beyond the present normal

computational capability. In fact, previous experiences (e.g., Kleeman, 2008) have shown that, even for a Gaussian distribution, it is difficult to have the covariance matrix accurately estimated.

The exponential computational complexity poses a continuing challenge to the studies of uncertainty and uncertainty-related problems; it becomes particularly severe in the practical fields such as adaptive sampling, where a rapid assessment of the prediction is needed to decide where to make the next observation. We therefore have to seek alternative approaches to the estimation of the marginal entropy rate dH_k/dt . Inspired by the above concise law namely (2), possibly we may first arrive at for dH_k/dt a formula that is as easy to handle as (2), and then use the formula to do the remaining work. But, as we will see in the next section, this is not the case. The resulting formula for dH_k/dt also requires an accurate estimation of the density functions and is hence impractical for large systems. Nonetheless, by carefully classifying the mechanisms that govern the increase of H_k , one will find that a key part that is intrinsic to the model itself and shows the trend of dH_k/dt actually can be fairly accurately estimated. That is to say, the uncertainty generation problem can be at least partially solved. The present study is devoted to the establishment and demonstration of this, using a barotropic quasi-geostrophic (QG) jet stream model.

An introduction of the theory is presented in the following section, where a theorem is proved with the proof placed in Appendix A. The application model is then set up (Section 3), and some solutions are briefly described (Section 4). In the interest of readership, some lengthy derivation is put in Appendix B. We choose two chaotic attractors for the uncertainty generation study, the results being presented in Sections 5 and 6, respectively. We then show that, remarkably, similar results can be obtained to a fairly satisfactory extent without resorting to ensemble predictions (Section 7). In Section 8 this study is concluded and some remaining problems are discussed.

2. Uncertainty generation

2.1. Local entropy generation and cumulant information transfer

Consider the system (1), i.e., $d\mathbf{u}/dt = \mathbf{F}(\mathbf{u}, t)$, where $\mathbf{u} = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$. We use \mathbf{u} instead of the conventional notation \mathbf{x} to denote the state variables in order to avoid confusion with the physical space dimension x . Let $\rho = \rho(\mathbf{u}) = \rho(u_1, u_2, \dots, u_n)$ be the joint probability density function (pdf) of \mathbf{u} . Then there is a Liouville equation (cf. Lasota and Mackay, 1994) governing the evolution of ρ :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{F}\rho) = 0. \quad (3)$$

Here the divergence is taken with respect to the n components of \mathbf{u} . We have shown in the introduction that the global uncertainty generation for the whole system, i.e., the evolution of the joint entropy

$$H = - \int_{\mathbb{R}^n} \rho(\mathbf{u}) \log \rho(\mathbf{u}) d\mathbf{u},$$

is governed by a law expressed in (2), as obtained by Liang and Kleeman (2005). Locally for each $k = 1, 2, \dots, n$, the uncertainty is respectively represented by the marginal entropies of u_k , i.e.,

$$H_k = - \int_{\mathbb{R}} \rho_k(u_k) \log \rho_k(u_k) du_k, \quad (4)$$

of which the evolution equation has also been obtained (Liang and Kleeman, 2005):

$$\frac{dH_k}{dt} = -E \left(F_k \frac{\partial \log \rho_k}{\partial u_k} \right). \quad (5)$$

(This is an equivalent expression of the Eq. (8) in Liang and Kleeman, 2005.) As mentioned in the introduction, this formula, though seemingly simple, is actually very difficult to compute, with the involvement of the derivative $\partial \log \rho_k / \partial u_k$. In evaluating the marginal density ρ_k , one needs to discretize the sample space and count the bins. The accuracy here is not much guaranteed, let alone the logarithm and, particularly, the derivative with respect to the coarse-grained sample space.

To circumvent the difficulty, look at the mechanisms that account for the marginal entropy evolution. For location k , the mechanisms may be classified into two categories by proximity to k . In other words, the uncertainty at location k has two distinctly different sources: one local source from k itself, one remote source from other locations. Correspondingly the right hand side of (5) can be decomposed into two parts: one resulting from the local mechanism, written dH_k^*/dt ; another from remote locations (written $T_{other \rightarrow k}$) through a mechanism called information transfer or information flow as may be referred to in the literature. Information transfer/flow accounts for the causal relation between different locations; it is a fundamental concept in general physics which has important applications in a wide variety of disciplines. It has been of interest in scientific communities for decades but has just been rigorously formulated (Liang and Kleeman, 2005, 2007a,b; Liang, 2008). Here $T_{other \rightarrow k}$ actually results from the cumulant information transferred to location k , i.e.,

$$T_{other \rightarrow k} = \sum_{\substack{\ell = 1 \\ \ell \neq k}}^n T_{\ell \rightarrow k}, \tag{6}$$

where $T_{\ell \rightarrow k}$ is the information transfer from ℓ to k as derived in Liang and Kleeman (2007b) (cf. Eq. (54) in their paper). We will henceforth refer to dH_k^*/dt and $T_{other \rightarrow k}$ respectively as the rate of local entropy generation and the rate of cumulant information transfer, or LEG and CIT for short. In principle, the information transfer between locations ℓ and k has been explicitly obtained in Liang and Kleeman (2007b), and the CIT to K follows from (6) accordingly. Unfortunately, the afore-mentioned issue of computational complexity still exists in the formula as it requires evaluation of the joint density and its derivatives.

In contrast, the LEG evaluation does not have this problem. For location k , Liang and Kleeman (2005) intuitively argued that $dH_k^*/dt = E(\partial F_k/\partial u_k)$ based on the concise form of Eq. (2). In the 2D case, this is indeed true, as has been established by Liang and Kleeman (2007b) in their Theorem 1. For a system of arbitrary dimensionality, we claim that this also holds, hence comes the following theorem:

Theorem 1. For an n -dimensional system $\mathbf{du}/dt = \mathbf{F}(\mathbf{u}, t)$, the rate of local entropy generation (LEG) at location k is

$$\frac{dH_k^*}{dt} = E\left(\frac{\partial F_k}{\partial u_k}\right), \quad k = 1, 2, \dots, n. \tag{7}$$

Remark. F_k is generally a function of all the components of \mathbf{u} ; E is hence the mathematical expectation with respect to the joint pdf of \mathbf{u} , not just with respect to that of u_k .

Proof. The proof is rather technically involved; it follows the strategy in Liang and Kleeman (2007b) but for different purpose. Refer to Appendix A for details. \square

It is interesting to observe that LEG, and hence local uncertainty generation, is intrinsic to physical models. This is especially clear in the case of linear models. To see it, consider a system with a vector field

$$\mathbf{F}(u) = \mathbf{A}u, \quad \text{where } \mathbf{A} = (a_{ij}) \text{ is some matrix.}$$

By (7) its LEG at location k is:

$$\frac{dH_k^*}{dt} = E\left(\frac{\partial F_k}{\partial u_k}\right) = E\left[\frac{\partial}{\partial u_k} \sum_{\ell=1}^n a_{k\ell} u_\ell\right] = a_{kk},$$

which has no dependence on \mathbf{u} . That is to say, once a model is set up, its local uncertainty generation is completely determined.

The intrinsicity of LEG is rooted in its underlying relation with instability. Uncertainty and instability are two different physical notions, though in the literature people often use them simultaneously (sometimes even mix them together in a confusing way). But they do connect to each other. For instance, one may argue that uncertainty increase implies instability, for a stable or asymptotically

stable system only leads to more certain states. The above theorem actually shows this connection in a quantitative way: Eq. (7) gives the average of the Lyapunov exponents for the corresponding linearized systems over all the possible perturbations localized at k . So if a system is stable/unstable under all the perturbations it must be certain/uncertain, i.e., with non-increasing/increasing uncertainties. But for more generic cases, the correspondence is not one-to-one anymore; it is complicated by the mean operator in (7), which tells that uncertainty is a “bulk” notion, not about individual unstable or stable events. That is to say, instability alone does not imply uncertainty increase. For a system, instabilities and stabilities may coexist, and the effect of the former may be canceled out by that of the latter, leading to no local uncertainty growth.

That LEG accounts for the intrinsic part of the marginal entropy variation is further substantiated by the observation that CIT occurs only in the course of state evolution, for information transfers through dynamic interactions between events. An example is shown in Liang (2008, Fig. 1c) where the transfer from one state to another spins up from zero to some limit, while the entropy production local to the latter is a constant in the absence of stochasticity. LEG is therefore ideal for characterizing the “static” part or potential of uncertainty generation, i.e., the potential of uncertainty generation embedded in the model itself.

The concise form of (7) implies that, in computing the LEG of a system, only involved is the expectation of some value from the model. By the law of large numbers, the expectation can be approximated by an ensemble average, and hence it can be easily evaluated without counting bins in the discretized sample space. Indeed, as we will see soon in the following two applications, even with an ensemble of very limited size (say, with only 20 members), the computed LEG is still satisfactory.

2.2. Local entropy generation over a collection of phase space locations

When talking about uncertainty generation, sometimes we may not refer to a specific location, but to a subspace. Besides, the governing equations in question usually have more than one dependent variables. When written in the form of a dynamical system, different locations in phase space, i.e. different indices in Eq. (1), may correspond to one single spatial location in physical space. For example, a 2D shallow water model with the rigid-lid assumption has two prognostic variables, namely, the horizontal velocity components v_x and v_y . If the governing equations are discretized on a grid with N mesh points, we obtain a $2N$ -dimensional system in the form of Eq. (1). If the state variables are arranged as $\mathbf{u} = (v_{x1}, v_{x2}, \dots, v_{xN}; v_{y1}, v_{y2}, \dots, v_{yN})^T$, then u_k and u_{N+k} , for $k = 1, 2, \dots, N$, correspond to the same location in physical space, say ι . So when we talk about the uncertainty generation at ι , we are actually referring to the joint uncertainty generation of u_k and u_{N+k} . All these motivate one to examine the joint uncertainty generation over a collection of some components of \mathbf{u} .

Theorem 2. For Eq. (1), the joint entropy of components u_k and u_ℓ with effects from all other components excluded evolves as

$$\frac{dH_{k\ell}^*}{dt} = E \left(\frac{\partial F_k}{\partial u_k} + \frac{\partial F_\ell}{\partial u_\ell} \right), \quad k, \ell = 1, 2, \dots, n. \quad (8)$$

Proof. The proof follows precisely the same track as that for Theorem 1, with the univariable Taylor series expansion replaced by multivariable Taylor series expansion. \square

Equation (8) can be straightforwardly generalized to the case with arbitrarily many components. The result is the sum of the contributions from those individual components. Correspondingly the CIT follows by subtracting the LEG from the marginal entropy rate.

2.3. Compact support and compact subspace

The above formulas are over the entire sample space, which may be too large to handle. The problem becomes particularly severe when the system, such as the atmosphere or ocean, has a large dimensionality, with which one can only generate ensembles of limited size. As mentioned before, this makes uncertainty estimation a challenging task.

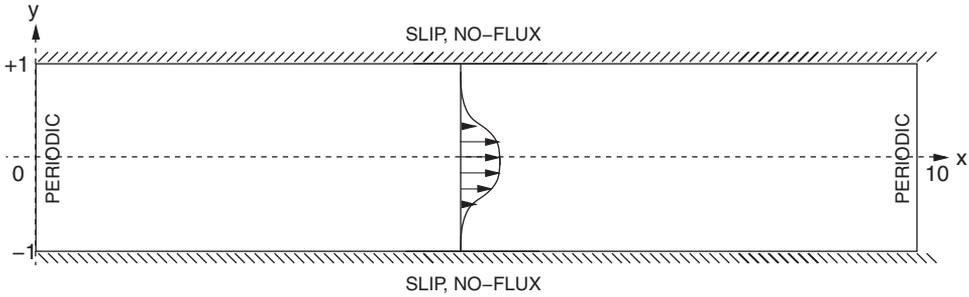


Fig. 1. Model configuration. Also shown is a schematic of the basic flow profile.

Fortunately, things are not that bad if one notices that, in establishing the theorems, the only “assumption” that we have used is that the pdf (and sometimes its derivatives) over a sample space Ω should vanish at its boundaries. In the above we deal with the entire sample space so this is of course true, otherwise it would not be the entire space. With this one may easily reason that, the theorems equally hold if the integration domain is a subspace of Ω bounded by the zero density (and density variation) contour(s). More formally, if ρ and its derivatives are compactly supported on Ω , we may focus on the support that makes a subspace. The uncertainty generation in this case thus follows precisely the above laws though here we are actually dealing with just a subspace. When the subspace has small dimensionality, the LEG and CIT can be conveniently computed by the above theorems. Fortunately again, for infinite dimensional nonlinear dynamical systems, it is a common feature that there usually exist some kind of compact attractors with finite dimensionality (e.g., Temam, 1997). Examples may be seen in the following applications. Note such is not the case for linear systems. In this sense, nonlinearity actually does a favor to our analysis, though people usually prefer linear systems to nonlinear systems when picking problems to work on.

3. The application model

The above theory is now applied to atmospheric studies. Consider a free-evolving and frictionless two-dimensional quasi-geostrophic jet stream confined within two latitudes (Fig. 1). Its dimensionless governing equation is (e.g., Pedlosky, 1987; Marshall and Plumb, 2008; Huang, 2010)

$$\frac{\partial \mathcal{L}\psi}{\partial t} + \frac{\partial \psi}{\partial x} \frac{\partial \mathcal{L}\psi}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial \mathcal{L}\psi}{\partial x} + \beta \frac{\partial \psi}{\partial x} = 0,$$

where ψ is the streamfunction, (x, y) are the spatial coordinates directing eastward and northward, respectively, and β is the meridional derivative of the Coriolis parameter (scaled by velocity scale over length scale squared). In the equation, we adopt the notation \mathcal{L} , instead of ∇^2 , for the Laplacian:

$$\mathcal{L} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2},$$

in order to avoid possible confusion with the divergence operator in (2), which is with respect to $\mathbf{u} = (u_1, \dots, u_n)$. With this the vorticity field is simply $\mathcal{L}\psi$. All these variables are dimensionless. Consider a rectangular domain $[0, 10] \times [-1, 1]$. The boundary is slip in y and periodic in x . That is to say,

$$\psi = \text{const}, \quad y = \pm 1 \tag{9}$$

$$\psi(10, y) = \psi(0, y), \quad -1 \leq y \leq 1. \tag{10}$$

Choose a basic zonal flow $(U(y), 0)$ and decompose the streamfunction ψ into a background part Ψ and a perturbation ψ'

$$\psi = \Psi + \psi'$$

such that

$$\frac{\partial \Psi}{\partial y} = -U(y), \quad \text{or} \quad \Psi = \Psi(y) = - \int U(y)dy + C, \tag{11}$$

with C an integral constant to be determined by the boundary conditions. The original governing equation becomes

$$\frac{\partial}{\partial t} \mathcal{L}\psi' + \frac{\partial \psi'}{\partial x} \frac{\partial \mathcal{L}\psi'}{\partial y} - \frac{\partial \psi'}{\partial y} \frac{\partial \mathcal{L}\psi'}{\partial x} + U \frac{\partial}{\partial x} \mathcal{L}\psi' + (\beta - U_{yy}) \frac{\partial \psi'}{\partial x} = 0.$$

For the sake of notation simplicity, drop the primes with the variables; ψ henceforth stands for the perturbation streamfunction. Without loss of generality, further let $\beta = 0$. This is justified by the observation that β and $-U_{yy}$ are acting the same role in the equation, and hence one for sure can take into its effect by adjusting the meridional structure of the background flow. The resulting governing equation for the perturbation streamfunction is, therefore,

$$\frac{\partial}{\partial t} \mathcal{L}\psi + \frac{\partial \psi}{\partial x} \frac{\partial \mathcal{L}\psi}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial \mathcal{L}\psi}{\partial x} + U \frac{\partial}{\partial x} \mathcal{L}\psi - U_{yy} \frac{\partial \psi}{\partial x} = 0.$$

Due to their linearity, the operators \mathcal{L} and $\partial/\partial t$ are commutable. The evolution of ψ thus can be obtained:

$$\frac{\partial \psi}{\partial t} = \mathcal{L}^{-1} \left[- \frac{\partial \psi}{\partial x} \frac{\partial \mathcal{L}\psi}{\partial y} + \frac{\partial \psi}{\partial y} \frac{\partial \mathcal{L}\psi}{\partial x} - U \frac{\partial}{\partial x} \mathcal{L}\psi + U_{yy} \frac{\partial \psi}{\partial x} \right]. \tag{12}$$

It is easy to see that the boundary conditions are the same in form as (9) and (10), with ψ now understood as the perturbation streamfunction. Throughout the study, the basic velocity profile $U = U(y)$ is fixed to be cosine:

$$U = \cos^2 \frac{\pi}{2} y. \tag{13}$$

This is what Kuo originally chose in his model (see Kuo, 1973), a well-known barotropic jet unstable to a spectrum of disturbances.

To solve, Eq. (12) is discretized using the finite-difference method. Here we adopt a leap-frog scheme for the differencing, which maintains an accuracy of the second order while keeping the features from being damped. This is important, as in many cases damping may destroy the unstable scenario and steer the state to a different (usually stable) equilibrium. But this scheme is also known to be time-splitting, and the developed computational modes in time may blow up the integration. We deal with this issue by applying a weak three-point time filter (with a coefficient 0.01) every 100 time steps. It successfully suppresses the numerical instability while not hurting the physical instability. In this study, the domain is partitioned into 50×51 mesh grids, and correspondingly, the horizontal spacings are $\Delta x = 0.2$ and $\Delta y = 0.04$. This resolution might be a little coarse, but for a manageable dimension size for ensemble prediction, we have to live with it. Anyhow, as we claimed in the beginning, the main purpose of this application is to demonstrate through the jet stream problem how the ideas can be utilized for uncertainty studies, rather than to perform a detailed study of the jet stream itself. In doing so the infinite state space for ψ is approximated by a finite-dimensional space, and a system of finite dimensionality results accordingly. In this case, the system has 2550 dimensions.

We will henceforth only deal with the so-obtained 2550-dimensional system, and forget about the original QG model. (This is because, though with the same physical model, different numerical schemes may result in different dynamical systems.) Clearly here the vector field \mathbf{F} has no explicit dependence on time; that is to say, what we have is an autonomous system for ψ at an array of discrete locations. We solve it with a uniform time step $\Delta t = 0.0005$. To compute its LEG, we need to find the differential of the vector field, $D\mathbf{F}$. For this problem, it is easier to begin with the vector field of the discretized vorticity equation, which we denote by \mathbf{F}_ζ . With a central difference scheme, obviously the matrix $D\mathbf{F}_\zeta$

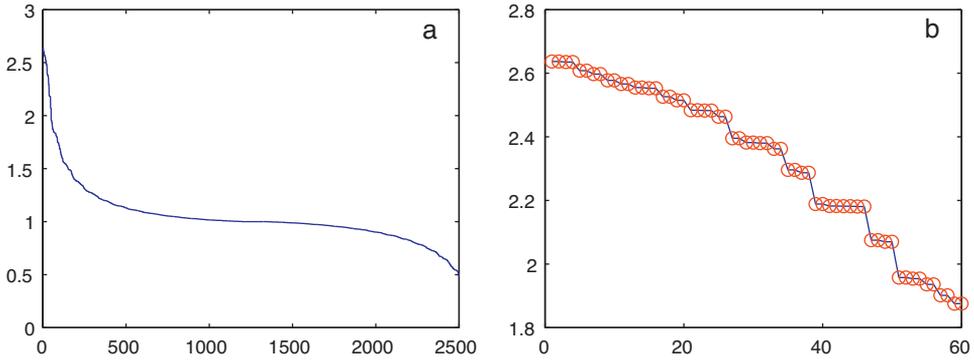


Fig. 2. (a) Singular value vs. SVD mode number. (b) A close-up of (a).

has at most twelve nonzero entries, which can be easily written out; see Appendix B for details. Once this is known, the differential of the vector field for ψ , i.e., DF , is

$$DF = \mathbf{L}^{-1}DF_{\zeta}. \quad (14)$$

where \mathbf{L} is the discretized Laplacian operator \mathcal{L} , and \mathbf{L}^{-1} its inverse. The diagonal entries of DF are what we need for the LEG evaluation.

4. Several attractors

4.1. Optimal excitation

To start our uncertainty study, first look at the unstable solutions of (12) that may lead to a chaotic scenario. Initially when the perturbation is small, (12) can be linearized, and the first-order dynamics is in the form of

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}\mathbf{u}, \quad \text{with } \mathbf{A} = \left(\frac{\partial F_k}{\partial u_\ell} \right)_{\mathbf{u}=0} \text{ being a linear operator.} \quad (15)$$

The optimal perturbation which leads to a maximal growth is given by the singular value decomposition (SVD) of $e^{\mathbf{A}t}$, as established in Farrell and Ioannou (1996). Here time t affects only the absolute size of the singular values; it does not have effect on the resulting optimal modes. We set it to be 1. Plotted in Fig. 2 are the singular values, ordered by magnitude. From it we see that the growth rate is almost the same for the first 26 modes (decreases from 2.63 to 2.5), and decreases to 1.7 at mode 100, indicating that the optimal basis has a large dimension. In Fig. 3, the structures of mode 1 and mode 100 are contoured. They are normalized, and mutually orthogonal. Though different in scale and orientation, they are similar in that they both have perturbations located in the neighborhood of the inflection points on the basic velocity profile. Besides, all the perturbations are oriented in opposition to the profile to allow the kinetic energy to be extracted from the background. For other modes, the structures are different. As shown in Fig. 4, one sees a completely different scenario on the patterns of mode 500 and mode 750. Although they still have a bending structure toward the left (west) against the stream, the perturbations are trapped around the axis ($y=0$) and near the boundaries ($y=\pm 1$).

That the first major optimal modes all have the perturbations around the inflection points of the basic velocity profile has its dynamical origin in the energy transfer structure with the Kuo model. Liang and Robinson (2007) demonstrated that on the canonical transfer distribution, the inflection regions act as two sinks of perturbation energy, though in all other regions the flow is unstable. One natural selection of the optimal perturbation is, therefore, to have the two sinks inhibited, and that is precisely what is shown above with the optimal modes.

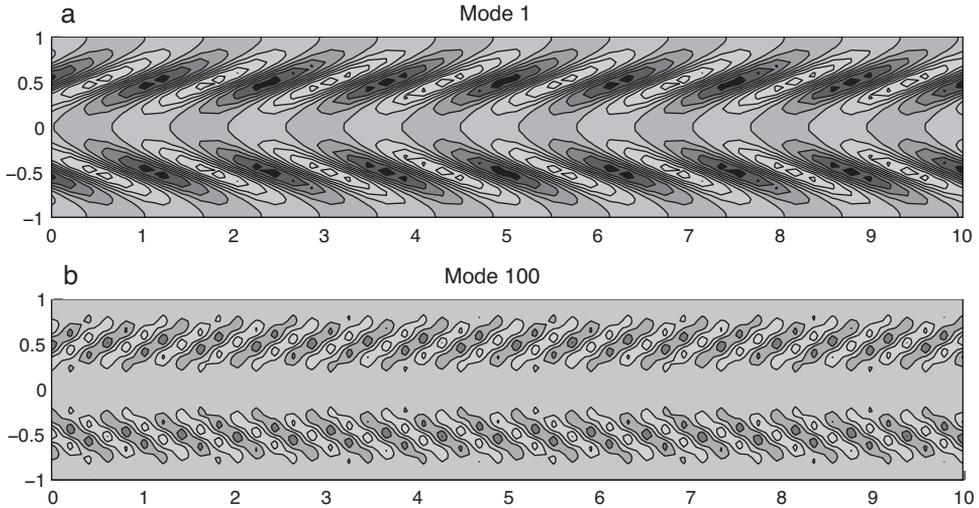


Fig. 3. The first and 100th optimal excitation modes.

Initialized with these modes, the governing equation (12) is integrated forward. We first take a look at the result with the most optimal perturbation mode, i.e., mode 1. In order not to shock the system too much, the initial condition is taken as the normalized mode multiplied by a factor 0.20. That is to say, the perturbation amplitude used for initialization is 0.20.

Just as predicted by Farrell's generalized stability theory, the system is very unstable. It undergoes a series of rapid changes in pattern before reaching a statistical equilibrium. The original reflection symmetry about the zonal axis is quickly broken, and, by time $t=5$, it has been organized into a pattern with negative and positive perturbations respectively dominating the south and north, and hence a strong perturbation flow counteracting the jet stream. This south–north anti-symmetry is reversed at $t=10$, when the strongest perturbation moves to the central axis. The emergence of this axially trapped perturbation changes the 2-peak structure into a 3-peak structure in the meridional direction. It then gradually connects the northern and southern peaks, making the whole perturbation field into a zonal wave form, with a meridional structure bending to the east (in opposition to the original orientation on the perturbation pattern). This scenario has been evident by $t=25$. But it does not last long; from $t=30$ to $t=40$, it is quickly reorganized into a wave with 8 peaks in x , and 3 peaks in y . After that, the energy is gradually transferred toward large scales, leaving a wave with 3 zonal wavenumbers ($t=50$). Later on at $t=75$, the 3-wave is changed into a form with 3 peaks plus 2 valleys and, finally, evolves into a zonal wave with 2 wavenumbers, as shown in Fig. 5.

An observation with this simulation is that, in the course of evolution, the patterns appear to be fairly regular. That is to say, though the system is very unstable, the reconstructed instability structures seem to be certain to a large extent. This is somehow not what we want for the application, as we would like to seek an example with enhanced feature of uncertainty to test our theory developed earlier on. For this reason, we turn to another optimal mode, mode 100.

Integration of the system (12) with an initial condition as shown in Fig. 3b (multiplied by 0.20) reveals a similar route for the perturbation field toward its statistical equilibrium. For example, it also begins with a breaking of the reflection symmetry about the zonal axis, followed by an oscillation of the negative/positive re-arrangement of the perturbation streamfunction in the south and north, which is accompanied by a reversing of the perturbation flow. The difference is that these events take slightly longer time. But afterwards, the duration of the 3-wavenumber stage is much shorter; as a result, the 2-wavenumber equilibrium is early reached by $t=60$ or so (see Fig. 6). Another difference is that the scenario in this case reveals much more chaotic details, as is easy seen when comparing

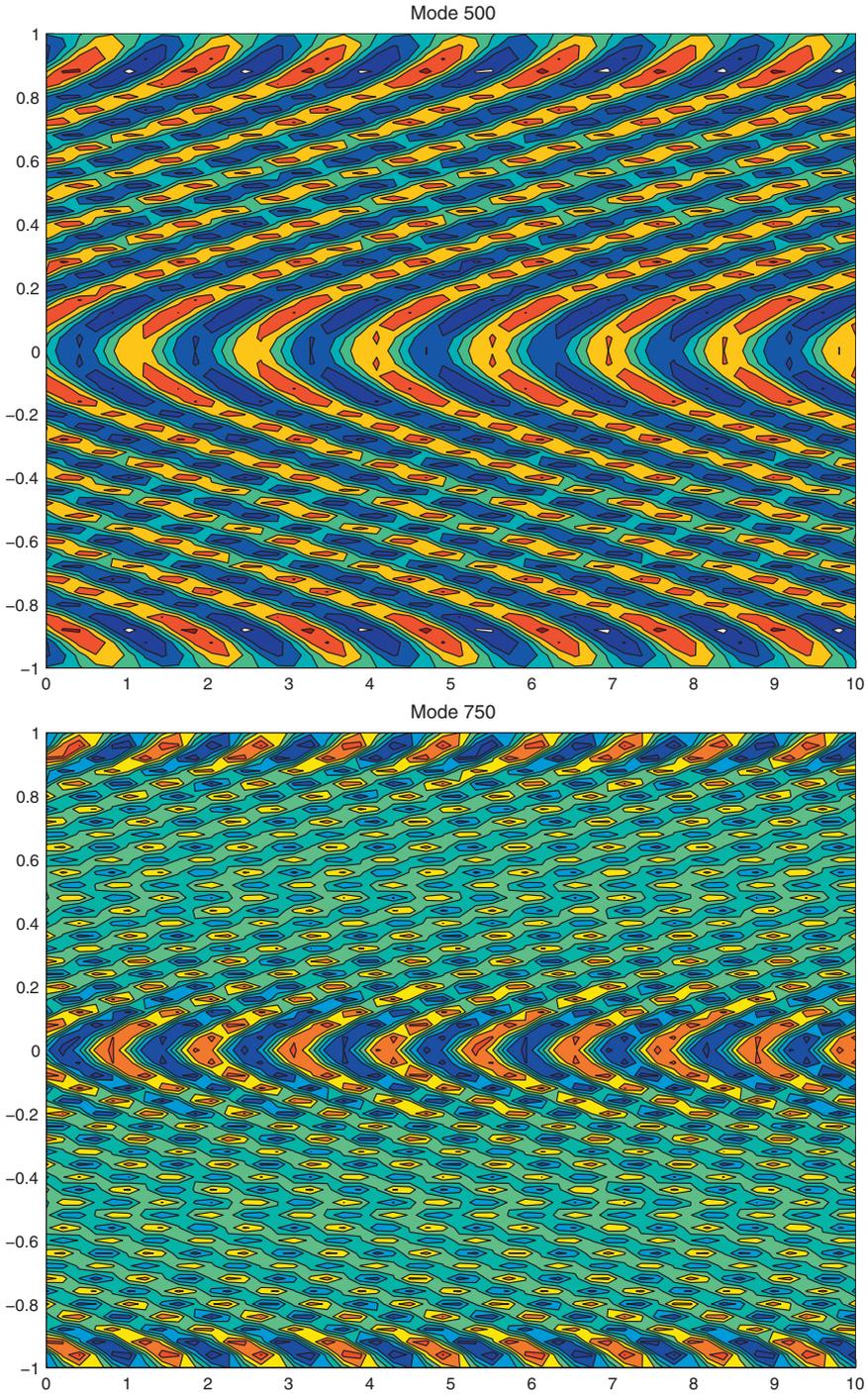


Fig. 4. Some other optimal excitation modes. The y-coordinates have been exaggerated for clarity.

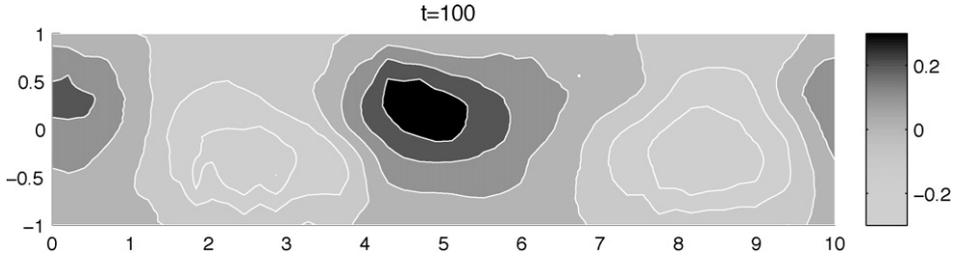


Fig. 5. The equilibrium solution initialized with the first optimal mode.

Fig. 6 to Fig. 5. Recall that it is uncertainty that we are interested in; we will hence use this result to build the dataset for our application.

4.2. Axially trapped excitation

The optimal excitation modes, though different in structures, all possess a large amount of perturbation energy in the low-speed regions distant from the jet axis. As we see previously, this usually leads to a rapid re-organization of structures into larger scales due to the inverse cascade of energy (e.g., Rhines, 1975; Haidvogel and Held, 1980; McWilliams and Chow, 1981). The process will not cease unless the growing energy is moved away via some mechanism and a balance is established on the energetics (Rhines, 1975). In a sluggish region, there is no mechanism to inhibit the growth (recall β is set to zero so there is no Rossby wave propagation in this study), so we always get a basin-scale equilibrium pattern with the optimal mode initialization.

The above interpretation motivates the consideration of an initial condition with perturbation energy trapped at the axis, in the hope of obtaining a chaotic attractor rich in small scale features. This makes sense, as the background velocity is maximized at the axis and, in the presence of a strong current, the upscale transferred energy will be carried downstream, leading the inverse cascade to a halt at some small scale. We search for such structures within a family of unstable normal modes. Go back to the matrix \mathbf{A} in (15). An eigenanalysis yields a list of eigenvalues λ which, if having positive real part, indicate (exponential) instability. Corresponding to these λ 's are the unstable eigenmodes. Recall that here growth rate is not of our concern; what we need to consider is the axially trapped structure. Fig. 7 presents such a mode, which corresponds to an eigenvalue of $\lambda = 3.12 \times 10^{-7} + 4.98i$. (Shown in the figure is the real part; the imaginary part is the same except for a phase lag of $\pi/2$.) Notice the large zonal structure among the small-scale features (see Fig. 7b). In the following, we will be working with this mode, though for reference several other modes may also be mentioned (see Fig. 8).

Initialized with Fig. 7 (multiplied by 0.20) Eq. (12) is integrated and the result contoured. Fig. 9 is a snapshot at the statistical equilibrium, which, for clarity, has been enlarged in the y direction. On the whole the pattern is very chaotic, though a large feature with two enhanced centers around $x=4$ and $x=9$ is visible. As will be seen in the following, this simulation serves well for our purpose.

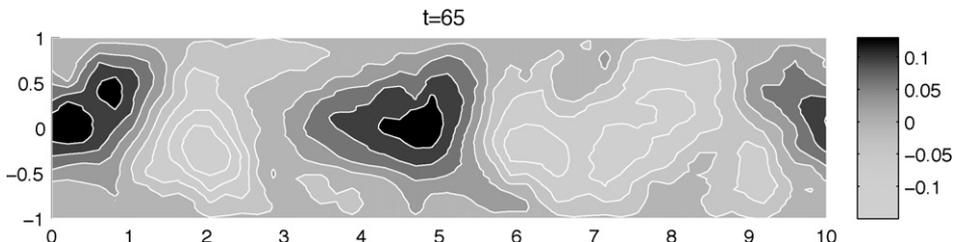


Fig. 6. The perturbation streamfunction at $t=65$ initialized with the 100th optimal mode as shown in Fig. 3b.

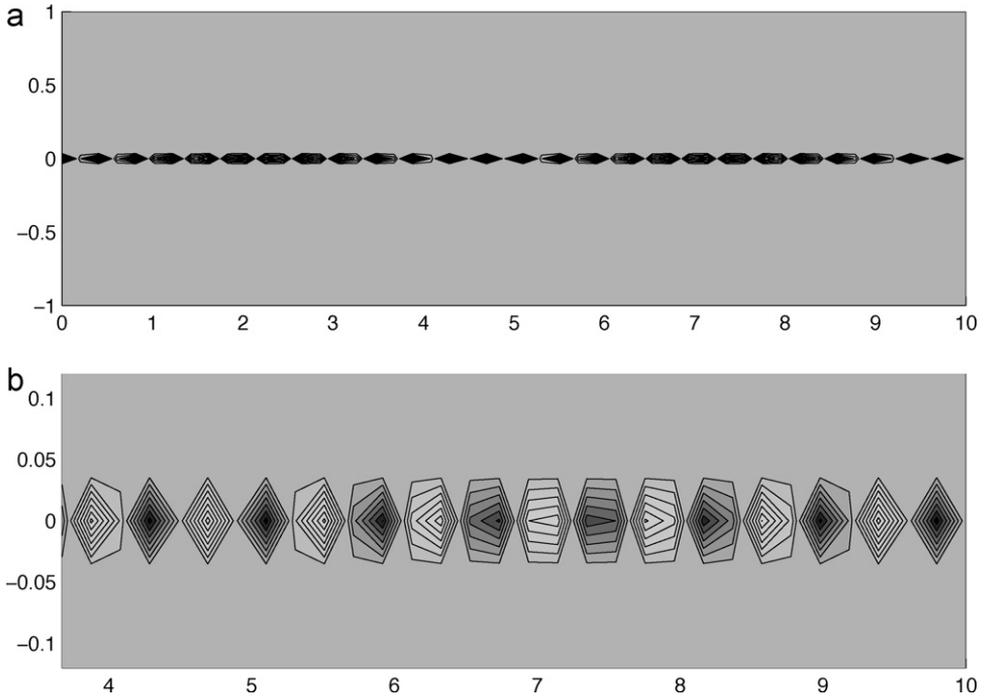


Fig. 7. (a) An unstable normal mode (corresponding to eigenvalue $3.12 \times 10^{-7} + 4.98i$). (b) A close-up of (a).

5. Uncertainty generation with the globally perturbed run

We need an ensemble of solutions of (12) to compute the LEG. As the model is deterministic, the randomness is limited within initial conditions, so a key point here is how an ensemble of initial conditions is formed. This section deals with the attractor shown in Fig. 6.

5.1. Ensemble formation

To generate an ensemble of initial conditions, integrate the system as shown in Fig. 6 from $t = 55$ for 40,000 steps to $t = 75$ (approximately one cycle for the observed large-scale feature). The structure of the attractor is found by performing an EOF analysis to the so-obtained data. From the result (Fig. 10), we see that the attractor is of low dimensionality. The first EOF mode dominates, accounting for 66% of the variance from that in Fig. 6 ($t = 65$). Over 90% of the variance is within the first 3 EOF modes, which are contoured in Fig. 11. We now use these three modes to generate the ensemble of initial conditions. (We have also tested with more modes but the final result is similar.)

For each EOF mode, compute from its corresponding projected series (denoted P1, P2, and P3 respectively) the mean and variance. Assuming that the time coefficients are normally distributed, we hence make random draws for the coefficients according to normal distributions with their respective means and variances. Specifically, the distributions are:

- P1: $N(-0.271, 0.0403)$,
- P2: $N(-0.0130, 0.0270)$,
- P3: $N(0.0556, 0.0099)$,

For each mode, Kleeman (2007) argued that 5 draws will be needed. We make 9, 7, and 5 draws for the first, second, and third modes, respectively. This gives an ensemble of $9 \times 7 \times 5 = 315$ members.

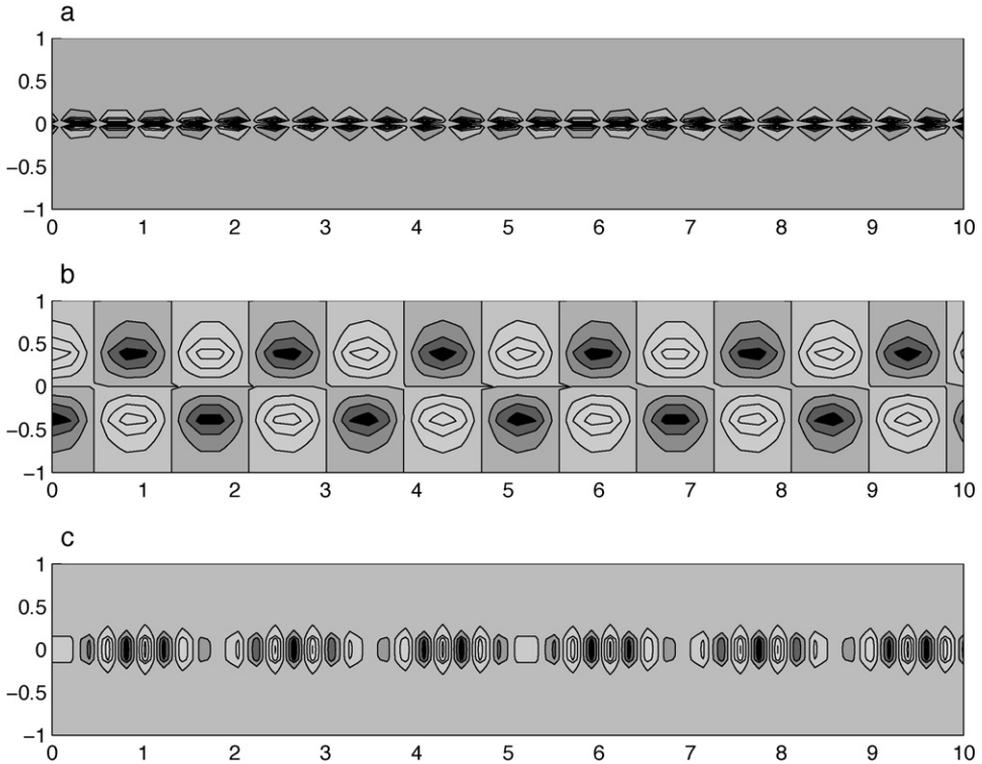


Fig. 8. Several other normal modes. (a) and (b) are stable (corresponding to eigenvalues $-1.83 \times 10^{-6} + 4.95i$ and $-1.36 \times 10^{-6} - 2.24i$ respectively); (c) is unstable (eigenvalue: $1.17 \times 10^{-6} - 1.75i$).

Another ensemble with $9 \times 9 \times 9 = 729$ members has also been generated and tested, and the result is essentially the same.

5.2. Uncertainty generation

With the 315 initial conditions, the system is steered forward to form ensembles of streamfunction at different time steps. This gives, in principle, all the needed information for the evaluation of the uncertainty generation rate.

First look at the LEG, i.e., dH_k^*/dt , for all k . Theorem 1 tells that it is equal to the expectation of $\partial F_k/\partial u_k$, which, given a model, can be analytically determined. By the law of large numbers, sample average converges to the expected value. So with a sizable ensemble, this part can be accurately estimated.

To illustrate this, assume that $\partial F_k/\partial u_k$ are Gaussian¹ and iid variables with mean μ and variance σ^2 . Then the sample average over an ensemble of size M is $N(\mu, \sigma^2/M)$ -distributed. Inverting the cumulant density function

$$\mathcal{F}(x, \mu, \frac{\sigma^2}{M}) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma/\sqrt{2M}} \right) \right]$$

¹ This assumption makes sense, thanks to the central limit theorem.

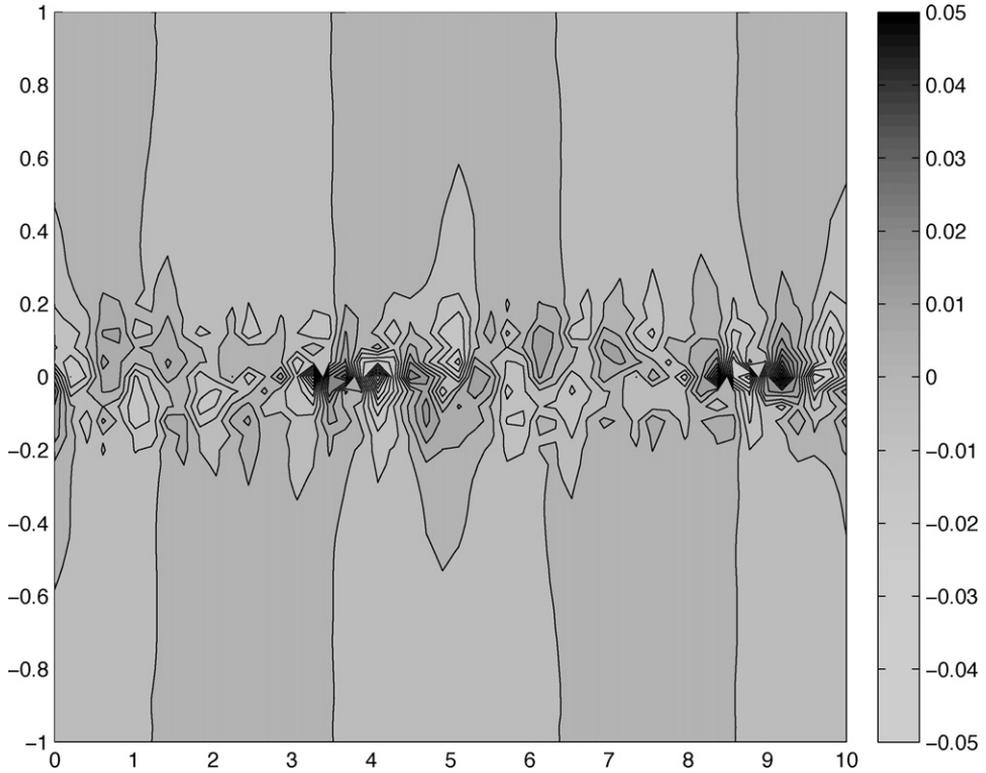


Fig. 9. A snapshot of the solution of (12) initialized with the axially trapped normal mode (Fig. 7). It is approximately at a statistical equilibrium. The aspect ratio is not plotted as it should be; the y scale is exaggerated to ensure a better visualization of the chaotic details.

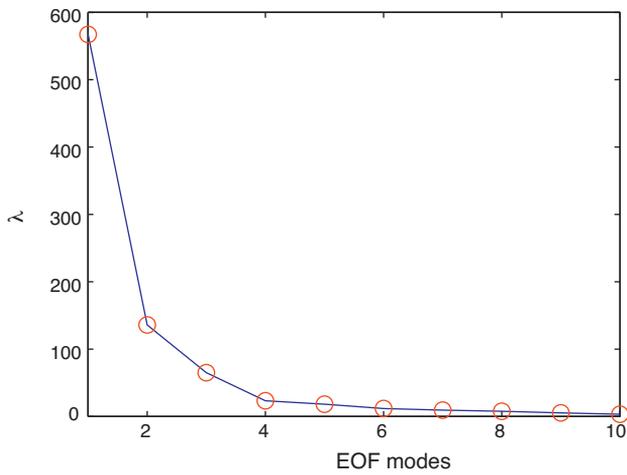


Fig. 10. Variance vs. mode number for the EOF analysis of the globally perturbed run.

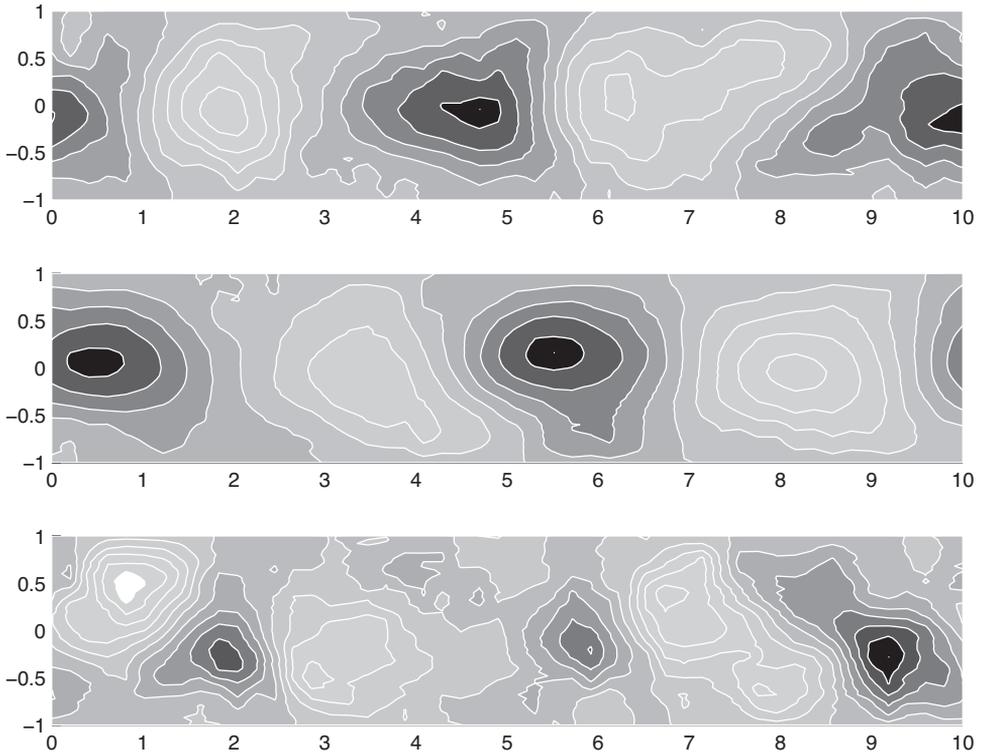


Fig. 11. The first three EOF modes (from top to bottom) for the globally perturbed run.

to get $(x - \mu)/(\sigma/\sqrt{2M}) \approx 1.3855$ for $\mathcal{F} = 97.5\%$. That is to say, the result is significant at a level of 95% if it falls within

$$\left[\mu - \frac{1.3855\sigma}{\sqrt{2M}}, \mu + \frac{1.3855\sigma}{\sqrt{2M}} \right].$$

Based on the result which we will show below, with $M = 315$ and conservative values $\mu = 0.1, \sigma \ll 0.1$, this tells that the relative error of the estimate is far less than 5%, at a significant level of 95%. So the local uncertainty generation can be fairly accurately estimated.

Presented in Fig. 12 are three LEG snapshots. Generally speaking, in this case the largest generation rate intrinsic to the model is near the two lateral boundaries. Also conspicuous on the distribution are two peaks tilting toward the west, which move downstream at a speed of roughly $1/(9000 \times 0.0005) = 0.22$, in agreement with the eastward propagation phase celerity in Fig. 6.

As we showed previously, though uncertainty and instability are two different notions in physics, they are intimately related indeed. Particularly LEG is connected through Eq. (7) to localized instability in the Lyapunov sense. With this the above structure is easily interpreted in physics. Comparing Fig. 12 to Fig. 6,² we see that the LEG is minimal along the axis where the perturbation takes its maximal value, while the largest LEG appears in the transition regions between the peaks and valleys where the perturbation vanishes. This makes sense. As the system we are dealing with is nonlinear and energy conserving, the disturbance growth ceases at its maxima. That is to say, the Lyapunov exponent at those points is zero, and so is the LEG if the scenarios are similar for all possible disturbances.

² To be precise, Fig. 6 corresponds to the LEG at step 1, which is not shown here. But roughly it is similar to that at step 1000 in Fig. 12.

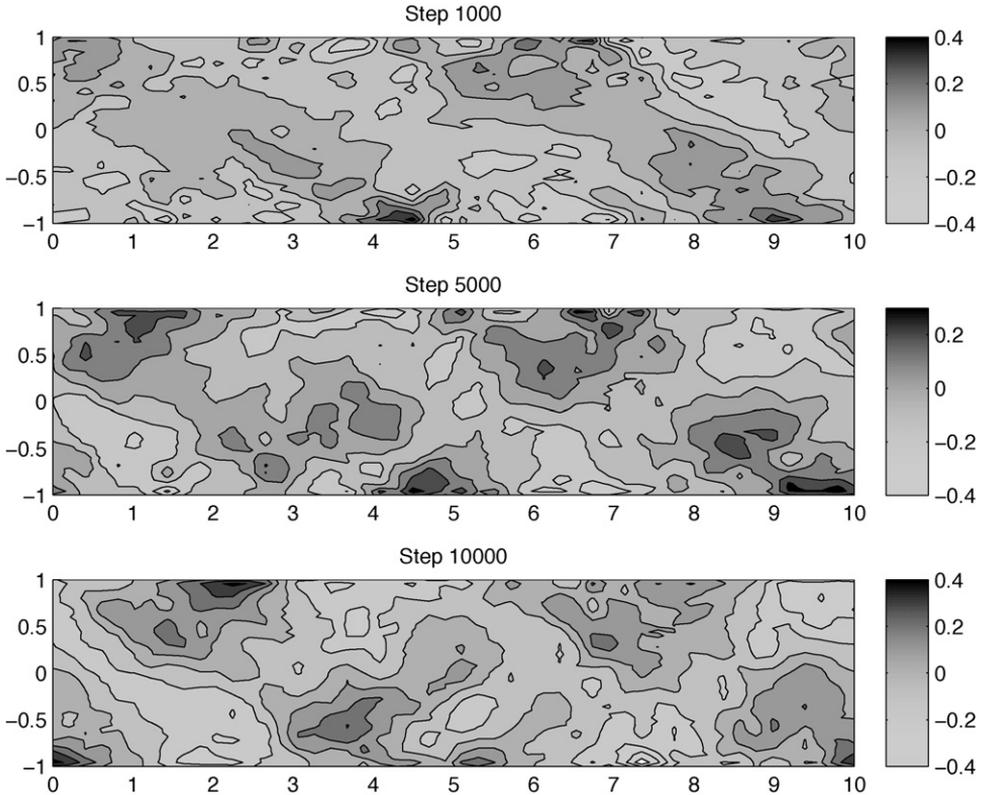


Fig. 12. The LEG at steps 1000 ($t=0.5$), 5000 ($t=2.5$), and 10,000 ($t=5$), for the globally perturbed run. The units are in nats per unit time.

On the other hand, when ψ at a location is first perturbed, disturbances grow the fastest (usually exponentially). So for a ψ distributed as in Fig. 6, the LEG must be maximized toward the two lateral boundaries.

Another observation of Fig. 12 is that, in the south ($y < 0$), the LEG takes positive (resp. negative) values where ψ makes transitions from negative to positive (resp. from positive to negative) as x increases, while in the north ($y > 0$) the scenario is reversed. This results from a joint effect of the flow and the solid boundaries. Take the southern half for an example. When ψ increases in x from negative to positive, $\partial\psi/\partial x > 0$, which gives a northward perturbation flow. This flow causes a divergence near the southern boundary, and hence locally a positive Lyapunov exponent. A positive LEG is therefore expected there. If the same gradient appears in the north, the northward flow will produce a convergence near the northern boundary, resulting in a local negative Lyapunov exponent and hence a negative LEG. This explains why from Fig. 6 we obtain a westward slanting structure for the LEG, and why LEG is maximized toward the boundaries.

In contrast to the simplicity in the LEG evaluation, the CIT, and the rate of marginal entropy $H_k = -\int \rho_k(u_k) \log \rho_k(u_k) du_k$, are difficult to compute, and it is not our intention to examine them in this study. Nonetheless, one may want to see through this application how LEG makes its part in dH_k/dt . We hence make a rough computation of H_k with the available ensemble members and then compute dH_k/dt . Here again problems arise when evaluating the time derivative. Any inaccuracy of the computed H_k , however small it is, may be amplified and lead to very large oscillation on the time series of dH_k/dt . As a result the spatial distributions are contaminated by random spots with huge spurious values. Here we deal with this by outputting H_k every twenty time steps, and then use central

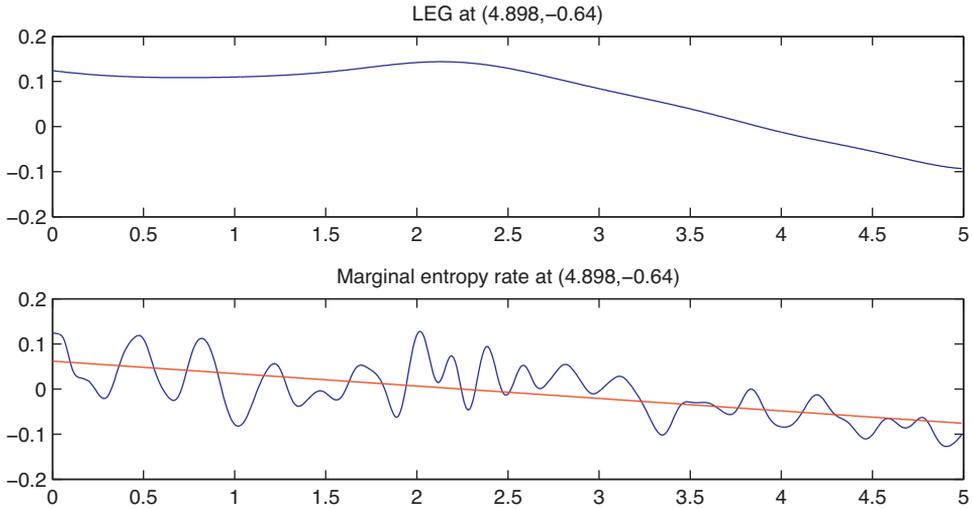


Fig. 13. Time series of the LEG (upper) and the marginal entropy rate (lower) at grid point (25,10) for the globally perturbed run. Also shown in the lower panel is the linearly fitted line.

differencing scheme to get a first estimate of dH_k/dt . The so-obtained time series is then smoothed by applying 50 times a three-point filter (1/4, 1/2, 1/4).

Shown in Fig. 13 are typical time series of the LEG (upper) and marginal entropy rates (lower). One sees that the former is smooth in general, while the latter has different kinds of time oscillations. As in this example, usually LEG accounts for the evolutionary trend of marginal entropy rate. This to some extent echoes our previous assertion about the intrinsicity of LEG.

5.3. LEG estimation with small ensembles

From the above analysis, it seems that the LEG of a system may be well estimated with an ensemble of rather limited size. To test this hypothesis, we generate an ensemble for the current prediction with only $5 \times 5 \times 5 = 125$ members. Just as expected, the resulting LEG is essential the same as that in Fig. 12 (not shown).

For practical reason, one is tempted to make more drastic reduction of the ensemble size and see what will happen with the LEG distribution. For this present prediction, we retain 5 draws for the first EOF mode, but make only 2 draws each for the remaining two. The resulting ensemble has only $5 \times 2 \times 2 = 20$ members, much less than $5 \times 5 \times 5 = 125$, i.e., the minimum size requirement. The computed result is shown in Fig. 14. As one may see, it agrees well with Fig. 12. Both have values maximized toward the lateral boundaries, and peaks/valleys tilt toward the west. The tilting slopes are essentially the same, and they propagate eastward at the same celerity. If there are anything different, the rate estimated here is slightly larger than that in Fig. 12, and some details superimposed on the large-scale structures may differ. Nevertheless, the similarity is striking, considering that the ensemble we are using is so small.

6. Uncertainty generation with the axially trapped perturbation run

6.1. Ensemble formation

Quite different from what we see in the previous section, the attractor from the axially trapped run appears very chaotic (Fig. 9). However, as for many GFD problems, it is actually of low dimensionality. Following the same procedure as in the preceding section, start from the state in Fig. 9 and integrate

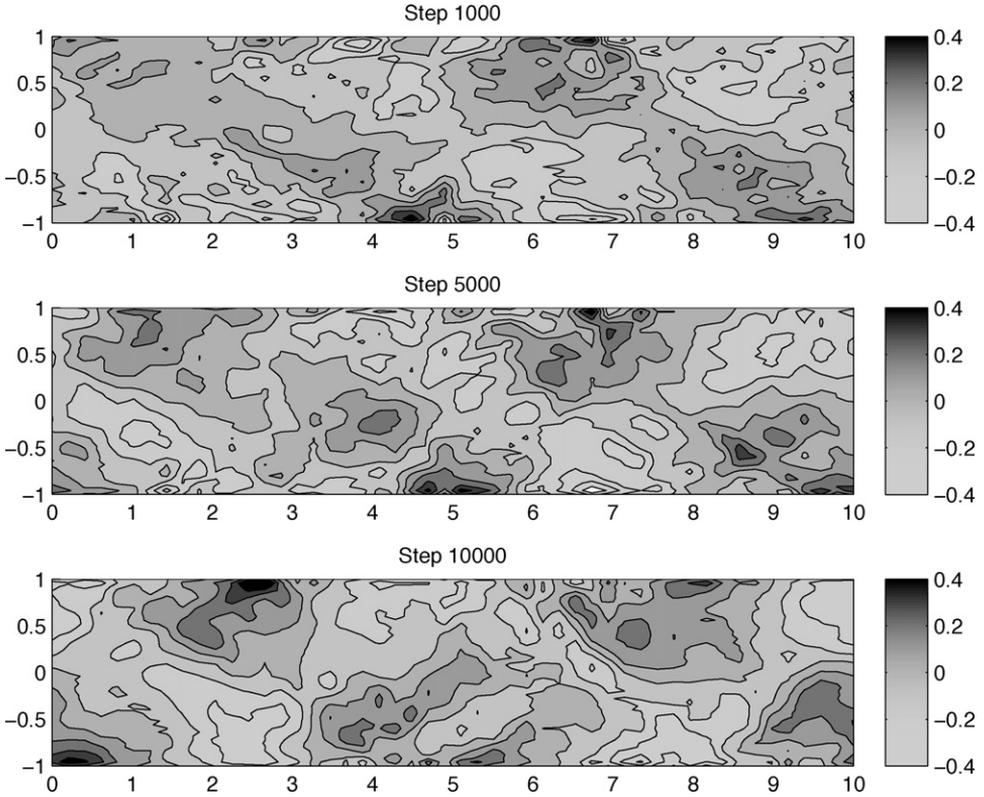


Fig. 14. Same as Fig. 12, but the estimation is with an ensemble of only 20 members.

the system for 10,000 steps (5 time units), with results being output every 2 steps. The obtained dataset is then EOF analyzed. From the result, over 90% of the variance from its initial field is covered within the first 5 EOF modes (cf. Fig. 15), among which the first accounts for 60%. We use these 5 modes to generate the ensemble of initial conditions.

As before, assume that the time coefficients corresponding to the 5 EOF modes are Gaussian, with means and variances computed from their respective time series P1, P2, and so forth. By computation the probability distributions are:

- P1: $N(-2.01 \times 10^{-2}, 7.7852 \times 10^{-5})$,
- P2: $N(-2.29 \times 10^{-4}, 8.43 \times 10^{-5})$,
- P3: $N(2.60 \times 10^{-3}, 6.57 \times 10^{-5})$,
- P4: $N(8.16 \times 10^{-5}, 5.00 \times 10^{-5})$,
- P5: $N(5.79 \times 10^{-5}, 3.60 \times 10^{-5})$.

Make 5 draws for each component and we obtain an ensemble with $5^5 = 3125$ members. Kleeman (2007) argued that 5 draws for each component suffice; this is also true here, we have doubled the ensemble size and obtained essentially the same result.

6.2. Uncertainty generation

Compared to the globally perturbed run in the previous section, the time scale in this case is much shorter. Generally speaking, a prediction of 500 steps ($t=0.25$) or less is enough for one to identify

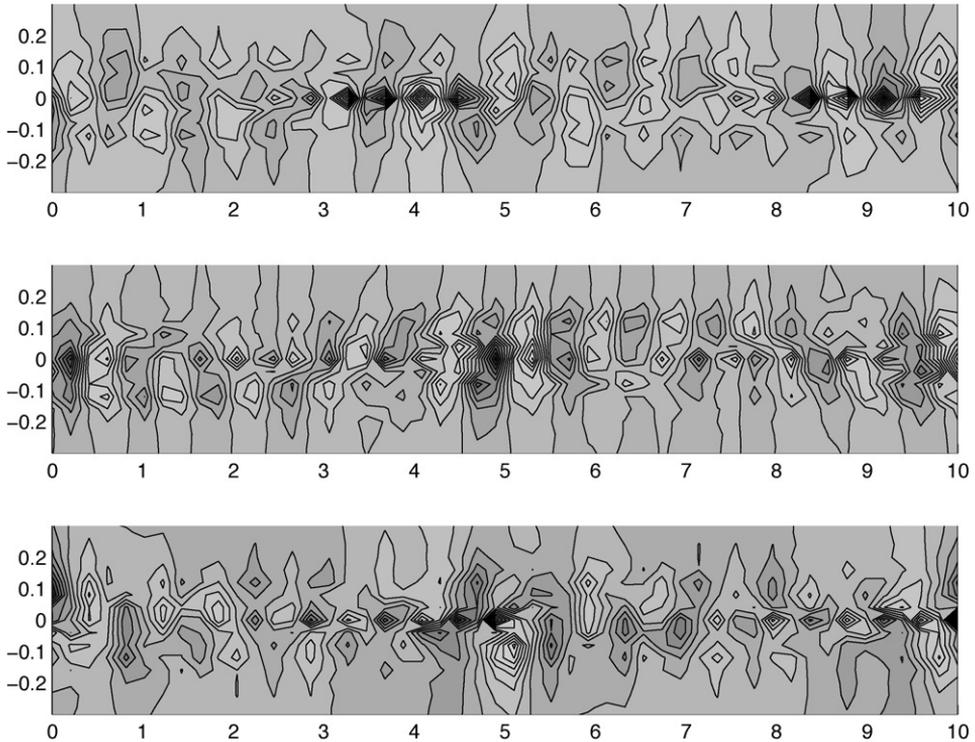


Fig. 15. The first three EOF modes (from top to bottom) for the run with axially trapped perturbation. The y direction has been enlarged for visual inspection.

the uncertainty generation patterns. A few of the 3125 members in the ensemble of initial conditions blow up the integration at 400 steps or so. These members are ruled out in estimating the uncertainty generation and, just in case of any numerical pollution, only examine the series up to 400 time steps.

Fig. 16 gives a snapshot of the computed LEG. We see from it that the rate is much larger than that in Fig. 12. That is to say, this run is much more uncertain than the previous globally perturbed run. This is within our expectation, as one may judge with naked eyes by comparing the attractors in Figs. 9 and 6. Another prominent feature is that it has a distribution with two clusters along the axis. Notice that dH_k^*/dt may be both positive and negative, indicating that the two centers are intrinsically both sources and sinks of uncertainty.

We also make a rough estimate of dH_k/dt to examine how LEG makes its part in the marginal entropy rate. The procedure is the same as that in the previous section. Fig. 17 are two such series at point (25,26) in the middle of the domain at the axis, or (4.9,0) in physical coordinates. The local

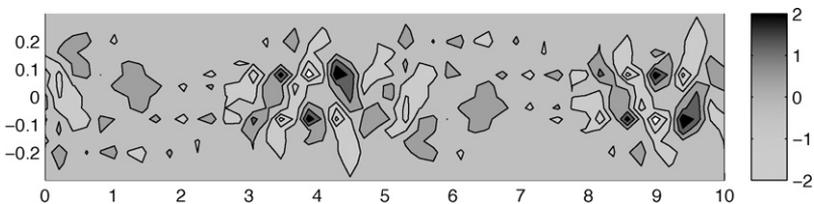


Fig. 16. The LEG at step 200 for the axially trapped perturbation run. The y direction has been enlarged for visual inspection; in the unshown region the LEG is negligible.

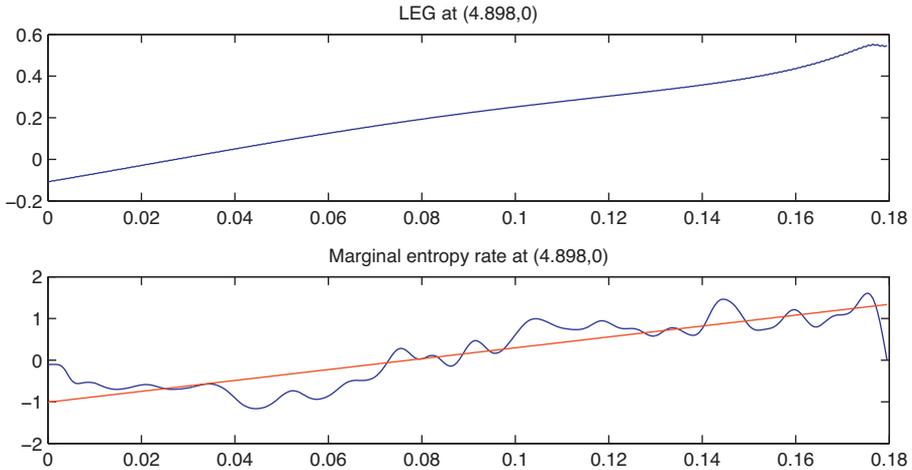


Fig. 17. Time series of the LEG and marginal entropy rate at point (25,26) for the axially trapped perturbation run. Also plotted in lower panel is a line showing the linear trend.

uncertainty generation dH_k^*/dt (LEG) is almost like a straight line, while others exhibit oscillations. Again, here dH_k^*/dt to a large extent explains the trend (see the fitted line in the middle subplot) of the marginal entropy rate; accordingly the time oscillation should be accounted for by the CIT. Of course, here we are not attempting to reach any conclusion, as the estimations of dH_k/dt and hence the CIT cannot be made accurate. More work is needed but we leave that to future studies.

6.3. LEG estimation with small ensembles

As before, it is tempting to reduce the ensemble size and see how the LEG estimate may vary. First consider only four EOF modes and form an ensemble of $5^4 = 625$ members. The result is essentially the same as Fig. 16. Further remove one EOF mode and retain only two draws each for modes 2 and 3, and form an ensemble with only $5 \times 2 \times 2 = 20$ members. The so-estimated LEG is shown in Fig. 18. As one may see, the basic structures in Fig. 16, such as the positioning of the positive/negative centers, have been reproduced here. The differences are, firstly, that the estimated rate here is slightly larger, and, secondly, that some features appear between the two clusters. But as a whole, the LEG has been fairly satisfactorily reproduced, even though the ensemble used is of such a small size.

7. LEG estimation without ensemble prediction

The successful LEG estimation with small ensembles motivates one to ask: is it possible to estimate LEG without doing ensemble prediction? This is a very tempting question; if it is true, the computational burden would be completely relieved in the study of uncertainty generation.

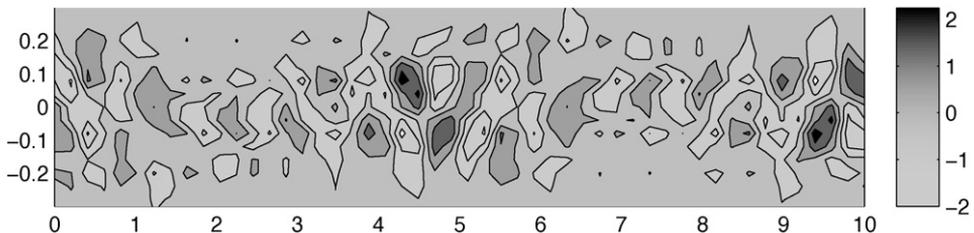


Fig. 18. Same as Fig. 16, but the LEG is estimated using an ensemble with only 20 members.

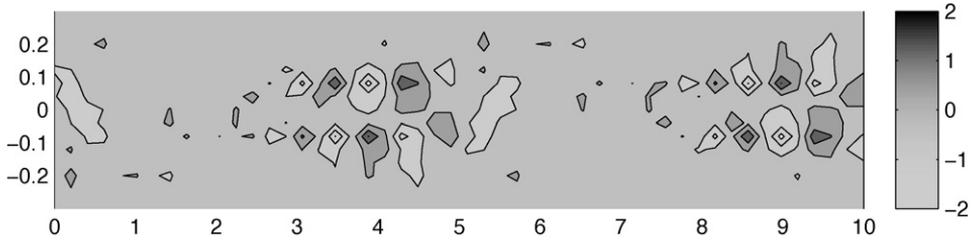


Fig. 19. Same as Fig. 16, but the LEG is estimated with a time average over 125 time steps centered around step 200.

In the above computations, we approximate the mathematical expectation in the LEG formula (7) with an ensemble mean. If the involved processes are ergodic and stationary, the ensemble mean can be replaced by a time average, and hence one may argue that the above assertion could indeed be true in this case. We put it to test with the standard prediction of the axially trapped run, which looks like quasi-stationary and quasi-ergodic. We have tried the time averages of $\partial F_k / \partial u_k$ with a time length of 125–401 steps. (We pick odd number steps to ensure a symmetry around the center point.) The results are essentially the same. Shown below in Fig. 19 is such an estimate with a time average over 125 steps centered around step 200. Compared to Fig. 16, the two clusters are evident, and the positive/negative centers are correctly located. Although the values may be slightly smaller, and the geometric shapes of the positive/negative centers may be a little different, as a whole the similarity between Figs. 19 and 16 is remarkable. Besides, we have also tried a 21-step averaging and a 3001-step averaging. The results, however, are not satisfactory. The 21-step one apparently has too short a time span to include all the major processes; the other one is too long and some longer variation appears, making the series no long stationary. A reasonable time span should be long enough to include all the processes, and short enough to ensure the stationary assumption to hold. In real applications, this may be set based on experience and physical intuition.

Of course, one cannot expect that real problems all have ergodic and stationary processes. For generic cases, we argue that the above strategy may still be worthwhile to try. In (7), the mathematical expectation acts, to some extent, a role like filtering, both in space and time. To test this conjecture, consider the globally perturbed run. For a time step, say 5000, we choose a series of $\partial F_k / \partial u_k$ centered around it with a time span of $2^{j_{\max}} = 1024$ ($j_{\max} = 10$) steps, and reconstruct it with an orthonormalized cubic-spline basis of scale level $j_0 = 5$ (see Liang and Anderson, 2007). The reconstruction is further filtered in space with a 2D orthonormalized cubic-spline basis (Liang and Anderson, 2007) to remove the 2-point features. Contoured in Fig. 20 is the result. Generally speaking, it looks similar to Fig. 12. The tilting structure is well reconstructed. Most of the positive and negative centers are at the right positions. Although improvements are still to be made, this is remarkable now, as we have gained a qualitatively satisfactory understanding of local uncertainty generation without even doing ensemble prediction!

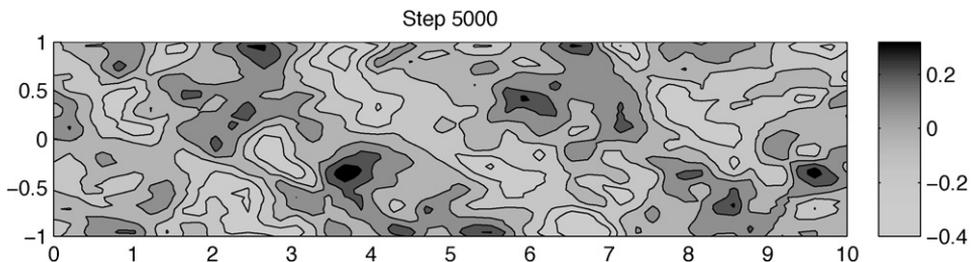


Fig. 20. The LEG for the globally perturbed run at step 5000 estimated with a time filter followed by a spatial filter.

8. Discussion and conclusions

We have studied the uncertainty generation within deterministic fluid flows, with preliminary applications to a barotropic quasi-geostrophic (QG) jet stream. For each location, say k , uncertainty arises along with the random distribution of some state variable, say $\mathbf{u} = (u_1, \dots, u_n)$, at the location; it is measured by the marginal entropy of u_k , or H_k as used in the text. The problem of uncertainty generation is thus about the time change of H_k , which is due to the entropy production local at k , and the information transfer (or information flow as may be referred to in the literature) from other locations. The former is referred to as local entropy generation, or LEG for short; the latter is a cumulant effect from all locations other than k , and is hence termed cumulant information transfer (CIT). We thus have, in short,

Marginal entropy change = local entropy generation (LEG) + cumulant information transfer (CIT).

Given a dynamical system

$$\frac{d\mathbf{u}}{dt} = \mathbf{F}(\mathbf{u}, t), \quad \mathbf{u} = (u_1, u_2, \dots, u_n),$$

the LEG at k , dH_k^*/dt , proves to be

$$\frac{dH_k^*}{dt} = E \left(\frac{\partial F_k}{\partial u_k} \right), \quad k = 1, 2, \dots, n,$$

where E is the mathematical expectation with respect to all the components of \mathbf{u} . This information-theoretic law is a rigorous mathematical result (refer to Appendix A for a detailed proof). The corresponding CIT can be computed, in principle, by subtracting the LEG from the marginal entropy rate.

To atmosphere and ocean models (including the governing equation, discretizing scheme, etc.) LEG is a "static" and intrinsic property, showing their potentials to generate uncertainties. Physically it is the mean of localized Lyapunov exponents, bringing connection between the two physical notions namely instability and uncertainty. For linear systems, LEG is independent of states, and can be computed analytically. In contrast, CIT arises in the course of state evolution; it is a result of the interaction between different locations through dynamic event synchronization.

A major challenge in the study of atmospheric/oceanic uncertainty generation is the evaluation of entropy-related quantities. Ensembles of enough size are in general computationally intractable with the computing capabilities so far. This study, however, shows that LEG can be accurately computed from the concise formula (7) with ensembles of very limited size. Furthermore, if the process of concern is ergodic and stationary, the ensemble mean in (7) actually can be well approximated with a time average. That is to say, in this case, the LEG can be estimated with a single realization. In more generic cases, though ergodicity and stationarity may not hold, we have shown that the LEG may still be evaluated to one's satisfaction with temporal and spatial filters. In other words, one may study local uncertainty generation without even doing ensemble prediction! This is remarkable as real atmospheric and oceanic systems are usually of huge dimensionality and hence ensemble predictions are expensive.

The above theory and hypotheses have been illustrated and tested with a QG barotropic atmospheric jet stream. Two predictions have been used, each with a compact low-dimensional attractor to ensure an ensemble prediction with a feasible but large enough ensemble. The first is a two-wavenumber wave superimposed by chaotic small-scale features. Its LEG also exhibits a wave form with the same phase celerity, but with a structure quite different from the attractor. Spatially it is intensified toward the lateral boundaries, and is dominated by a pattern tilting toward the west. Another prediction has a very chaotic attractor, which is trapped along the zonal axis. Its LEG is significant only within two clusters. Time series analysis shows that the LEG evolves smoothly, and more often than not it gives the trend of the uncertainty evolution. In both cases, their LEG rates have been satisfactorily estimated without doing ensemble prediction.

The physical importance and computational tractability make LEG a good measure to assess model performance in terms of uncertainty generation. In doing this, one should bear in mind that LEG includes inputs from both physics and numerics. That is to say, different discretizations may result in different LEG distributions, as one may check with the expression in Appendix B. Separation of these two sources should be straightforward, but it is not our intention to do it here; we intend to perform such a separation in a real ocean application with a sophisticated ocean prediction system, and present it in a forthcoming paper.

It should be noted that, in this study, we did not pay attention to the CIT namely the cumulant information transfer. Information transfer/flow is a very important mechanism to produce uncertainties within a system, but here what is considered is just the cumulant effect. That is to say, given a location k , what we obtain is just the effect to k from all other locations, without knowing where the effect is from. Clearly, differentiation of the cumulant transfer is important, as it helps to identify the causal relation between different locations. In this regard, a pioneering study was conducted by Kleeman (2007), using the empirical formalism available in his days. Since then, a rigorous formalism has been established and explicit formulas derived for both deterministic systems (Liang and Kleeman, 2007b) and stochastic systems (Liang, 2008), and has been applied to some low-dimensional dynamical systems. Applications to large dimensional systems like the atmosphere and oceans, however, encountered problems because of the involvement of joint probability density function in the formulas, whose evaluation is notoriously difficult. We expect more progress in this field in the near future.

Acknowledgments

The comments from two anonymous reviewers have helped to improve this manuscript, for which the author is grateful.

Appendix A. Proof of Theorem 1

Proof. By definition dH_k^*/dt means the time rate of change of the marginal entropy of u_k at time t with effects from all other components of \mathbf{u} excluded, i.e., with all other components instantaneously frozen as parameters at t . Without loss of generality, let $k=1$. If not, one can always re-arrange the indices of $\mathbf{u}=(u_1, u_2, \dots, u_n)$ to make it so. Discretization of Eq. (1) on the interval $[t, t+\Delta t]$ results in a mapping

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \mathbf{u}(t) \mapsto \mathbf{u}(t+\Delta t) = \mathbf{u}(t) + \mathbf{F}(\mathbf{u}(t), t)\Delta t.$$

To avoid confusion, denote $\mathbf{u}(t+\Delta t)$ by \mathbf{v} , and suppress the time dependence for clarity. The mapping is then

$$\begin{aligned} v_1 &= u_1 + F_1(\mathbf{u})\Delta t, \\ v_2 &= u_2 + F_2(\mathbf{u})\Delta t, \\ &\vdots \\ v_n &= u_n + F_n(\mathbf{u})\Delta t. \end{aligned}$$

Correspondingly there is a Frobenius–Perron operator (F–P operator henceforth)

$$\mathcal{P} : L^1(\mathbb{R}^n) \rightarrow L^1(\mathbb{R}^n), \quad \rho(\mathbf{u}) \mapsto \mathcal{P}\rho(\mathbf{u}),$$

which is defined to be, in a loose sense,

$$\int_{\omega} \mathcal{P}\rho(\mathbf{u})d\mathbf{u} = \int_{\Phi^{-1}(\omega)} \rho(\mathbf{u})d\mathbf{u}, \quad \text{for all } \omega \subset \mathbb{R}^n.$$

When the mapping Φ is nonsingular and invertible, \mathcal{P} can be explicitly evaluated (cf. Lasota and Mackay, 1994):

$$\mathcal{P}\rho(\mathbf{u}) = \rho \left[\Phi^{-1}(\mathbf{u}) \right] |J^{-1}|, \tag{A.1}$$

where

$$J = \det \left[\begin{array}{c} \frac{\partial(v_1, v_2, \dots, v_n)}{\partial(u_1, u_2, \dots, u_n)} \end{array} \right]$$

is the Jacobian determinant of Φ . For our problem, when $\Delta t \rightarrow 0$, Φ is nonsingular and always invertible. In fact, the inversion is $\Phi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto u$:

$$\begin{aligned} u_1 &= v_1 - F_1(\mathbf{v})\Delta t + O(\Delta t^2), \\ u_2 &= v_2 - F_2(\mathbf{v})\Delta t + O(\Delta t^2), \\ &\vdots \\ u_n &= v_n - F_n(\mathbf{v})\Delta t + O(\Delta t^2). \end{aligned}$$

Now let all the components of \mathbf{u} except u_1 be frozen as parameters at t . The system thus modified evolves from t to $t + \Delta t$, resulting a mapping: $\Phi_{\hat{u}_2 \dots \hat{u}_n} : u_1 \mapsto v_1$:

$$v_1 = u_1 + F_1(u_1, \hat{u}_2, \dots, \hat{u}_n)\Delta t,$$

where a hat signifies that the variable is frozen at t , and $\hat{\cdot}$ (for all i) indicates that the effect of u_i is excluded. Let $\Delta t \rightarrow 0$, again $\Phi_{\hat{u}_2 \dots \hat{u}_n}$ is invertible and $\Phi_{\hat{u}_2 \dots \hat{u}_n}^{-1} : v_1 \mapsto u_1$ is:

$$u_1 = v_1 - F_1(v_1, \hat{u}_2, \dots, \hat{u}_n)\Delta t + O(\Delta t^2).$$

The Jacobian of $\Phi_{\hat{u}_2 \dots \hat{u}_n}^{-1}$ is

$$J^{-1} = \det \left[\frac{\partial}{\partial v_1} \Phi_{\hat{u}_2 \dots \hat{u}_n}^{-1} \right] = 1 - \frac{\partial F_1}{\partial v_1}(v_1, \hat{u}_2, \dots, \hat{u}_n)\Delta t + O(\Delta t^2). \tag{A.2}$$

So the corresponding F-P operator

$$\begin{aligned} \mathcal{P}_{\hat{u}_2 \dots \hat{u}_n} \rho_1(v_1) &= \rho_1 \left[v_1 - F_1(v_1, \hat{u}_2, \dots, \hat{u}_n)\Delta t + O(\Delta t^2) \right] |J^{-1}| \\ &= \rho_1(v_1 - F_1\Delta t) \cdot \left(1 - \frac{\partial F_1}{\partial v_1}\Delta t \right) + O(\Delta t^2) \\ &= \left[\rho_1(v_1) - \frac{\partial \rho_1}{\partial v_1} F_1\Delta t \right] \left(1 - \frac{\partial F_1}{\partial v_1}\Delta t \right) + O(\Delta t^2) \\ &= \rho_1(v_1) - \frac{\partial(\rho_1 F_1)}{\partial v_1}\Delta t + O(\Delta t^2). \end{aligned} \tag{A.3}$$

For simplicity, all the variables with no arguments explicitly written are understood as functions of $(v_1, \hat{u}_2, \hat{u}_3, \dots, \hat{u}_n)$.

By the definition of Shannon entropy, $H_{1\hat{u}_2 \dots \hat{u}_n}(t + \Delta t)$ is the expectation of the function $f \equiv -\log \mathcal{P}_{\hat{u}_2 \dots \hat{u}_n} \rho_1(v_1)$. That is to say, it is equal to the integration of f times some pdf over the corresponding sample space. The first density to be multiplied is $\mathcal{P}_{\hat{u}_2 \dots \hat{u}_n} \rho_1(v_1)$, but f also depends

on the frozen components, namely \hat{u}_2, \hat{u}_3 , and so forth. So we need also multiply some joint density of $(\hat{u}_2, \hat{u}_3, \dots, \hat{u}_n)$. Here one should be cautioned that the freezing is performed on $[t, t + \Delta t]$, given u_1 at t . So the joint density is actually a conditional density on u_1 at t . Based on these arguments, the entropy of u_1 at $t + \Delta t$, with u_2, u_3, \dots, u_n frozen as parameters since t , is:

$$H_{1\mathcal{P}_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}}(t + \Delta t) = - \int_{\mathbb{R}^n} \mathcal{P}_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}\rho_1(v_1) \cdot \log \mathcal{P}_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}\rho_1(v_1) \cdot \rho(\hat{u}_2, \hat{u}_3, \dots, \hat{u}_n | u_1) dv_1 d\hat{u}_2 d\hat{u}_3 \dots d\hat{u}_n. \tag{A.4}$$

In (A.4), one thing that should be noted is the involvement of u_1 in $\rho(\hat{u}_2, \hat{u}_3, \dots, \hat{u}_n | u_1)$ while the integration is with respect to v_1 . We must convert u_1 to v_1 . Recall that they are related through the mapping $\Phi_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}^{-1}$, i.e.,

$$u_1 = v_1 - F_1 \Delta t + O(\Delta t^2).$$

We hence have

$$\begin{aligned} \rho(\hat{u}_2, \dots, \hat{u}_n | u_1) &= \frac{\rho(u_1, \hat{u}_2, \dots, \hat{u}_n)}{\rho_1(u_1)} = \frac{\rho(v_1, \hat{u}_2, \dots, \hat{u}_n) - (\partial\rho/\partial v_1)F_1 \Delta t}{\rho_1(v_1) - (\partial\rho_1/\partial v_1)F_1 \Delta t} + O(\Delta t^2) \\ &= \frac{1}{\rho_1(v_1)} \left[\rho(v_1, \hat{u}_2, \dots, \hat{u}_n) - \frac{\partial\rho}{\partial v_1} F_1 \Delta t \right] \left[1 + \frac{1}{\rho_1} \frac{\partial\rho_1}{\partial v_1} F_1 \Delta t \right] + O(\Delta t^2) \\ &= \frac{1}{\rho_1(v_1)} \left[\rho - \frac{\partial\rho}{\partial v_1} F_1 \Delta t + \frac{\rho}{\rho_1} \frac{\partial\rho_1}{\partial v_1} F_1 \Delta t \right] + O(\Delta t^2), \end{aligned}$$

where, again, all dependent variables without arguments are in $(v_1, \hat{u}_2, \dots, \hat{u}_n)$.

For the remaining parts in (A.4), by Eq. (A.3) we have

$$\begin{aligned} &\mathcal{P}_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}\rho_1(v_1) \cdot \log \mathcal{P}_{\mathcal{Q}\mathcal{R}\dots\mathcal{K}}(v_1) \\ &= \left(\rho_1(v_1) - \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right) \cdot \log \left(\rho_1(v_1) - \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right) + O(\Delta t^2) \\ &= \left(\rho_1(v_1) - \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right) \cdot \left[\log \rho_1 + \log \left(1 - \frac{1}{\rho_1} \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right) \right] + O(\Delta t^2) \\ &= \left(\rho_1 - \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right) \cdot \left[\log \rho_1 - \frac{1}{\rho_1} \frac{\partial(\rho_1 F_1)}{\partial v_1} \Delta t \right] + O(\Delta t^2) \\ &= \rho_1 \log \rho_1 - \left[\frac{\partial(\rho_1 F_1)}{\partial v_1} + \frac{\partial(\rho_1 F_1)}{\partial v_1} \log \rho_1 \right] \Delta t + O(\Delta t^2), \end{aligned}$$

where for notational simplicity, we have also omitted the argument v_1 in ρ_1 . So

$$\begin{aligned}
 H_{1\mathbb{V}\mathbb{V}\dots\mathbb{V}}(t + \Delta t) &= - \int_{\mathbb{R}^n} \left\{ \rho_1 \log \rho_1 - \left[\frac{\partial(\rho_1 F_1)}{\partial v_1} + \frac{\partial(\rho_1 F_1)}{\partial v_1} \log \rho_1 \right] \Delta t \right\} \\
 &\quad \cdot \frac{1}{\rho_1} \left[\rho - \frac{\partial \rho}{\partial v_1} F_1 \Delta t + \frac{\rho}{\rho_1} \frac{\partial \rho_1}{\partial v_1} F_1 \Delta t \right] dv_1 d\hat{u}_2 \dots d\hat{u}_n + O(\Delta t^2) \\
 &= - \int_{\mathbb{R}^n} \left[\rho \log \rho_1 + \log \rho_1 \left(-\frac{\partial \rho}{\partial v_1} F_1 + \frac{\rho}{\rho_1} \frac{\partial \rho_1}{\partial v_1} F_1 \right) \Delta t \right. \\
 &\quad \left. - \frac{\rho}{\rho_1} \left(\frac{\partial(\rho_1 F_1)}{\partial v_1} + \frac{\partial(\rho_1 F_1)}{\partial v_1} \log \rho_1 \right) \Delta t \right] dv_1 d\hat{u}_2 d\hat{u}_3 \dots d\hat{u}_n + O(\Delta t^2) \\
 &= - \int \rho(v_1, \hat{u}_2, \dots, \hat{u}_n) \log \rho_1(v_1) dv_1 d\hat{u}_2 \dots d\hat{u}_n - \int \left[\log \rho_1 \left(-\frac{\partial \rho}{\partial v_1} + \frac{\rho}{\rho_1} \frac{\partial \rho_1}{\partial v_1} \right) F_1 \Delta t \right. \\
 &\quad \left. - \frac{\rho}{\rho_1} \frac{\partial(\rho_1 F_1)}{\partial v_1} (1 + \log \rho_1) \Delta t \right] dv_1 d\hat{u}_2 \dots d\hat{u}_n + O(\Delta t^2) \\
 &= H_1(t) + \int \left[\frac{\partial(\rho F_1)}{\partial v_1} \log \rho_1 + \rho \frac{\partial F_1}{\partial v_1} + (\rho F_1) \frac{\partial \log \rho_1}{\partial v_1} \right] \Delta t dv_1 d\hat{u}_2 \dots d\hat{u}_n + O(\Delta t^2) \\
 &= H_1(t) + \int \left[\frac{\partial(\rho F_1 \log \rho_1)}{\partial v_1} + \rho \frac{\partial F_1}{\partial v_1} \right] \Delta t dv_1 d\hat{u}_2 \dots d\hat{u}_n + O(\Delta t^2).
 \end{aligned}$$

The first term in the bracket integrates to zero as the pdf vanishes on the boundaries of the sample space. Changing the variables $(v_1, \hat{u}_2, \dots, \hat{u}_n)$ back to (u_1, u_2, \dots, u_n) , the above integration results in

$$H_{1\mathbb{V}\mathbb{V}\dots\mathbb{V}}(t + \Delta t) = H_1(t) + E \left(\frac{\partial F_1}{\partial u_1} \right) + O(\Delta t^2). \quad (\text{A.5})$$

Therefore,

$$\frac{dH_1^*}{dt} = \lim_{\Delta t \rightarrow 0} \frac{H_{1\mathbb{V}\mathbb{V}\dots\mathbb{V}}(t + \Delta t) - H_1(t)}{\Delta t} = E \left(\frac{\partial F_1}{\partial u_1} \right). \quad \square$$

Appendix B. Differentiation of the vector field for the vorticity equation

For easy reference, rewrite the vorticity equation here as:

$$\mathcal{L} \frac{\partial \psi}{\partial t} = -\frac{\partial \psi}{\partial x} \frac{\partial \mathcal{L} \psi}{\partial y} + \frac{\partial \psi}{\partial y} \frac{\partial \mathcal{L} \psi}{\partial x} - U \frac{\partial \mathcal{L} \psi}{\partial x} + U_{yy} \frac{\partial \psi}{\partial x}. \quad (\text{B.1})$$

Discretizing with central differences, we get

$$\begin{aligned}
 \mathcal{L} \frac{\partial \psi_{i,j}}{\partial t} &= -\frac{\psi_{i+1,j} - \psi_{i-1,j}}{2\Delta x} \times \frac{(\mathcal{L}\psi)_{i,j+1} - (\mathcal{L}\psi)_{i,j-1}}{2\Delta y} + \frac{\psi_{i,j+1} - \psi_{i,j-1}}{2\Delta y} \times \frac{(\mathcal{L}\psi)_{i+1,j} - (\mathcal{L}\psi)_{i-1,j}}{2\Delta x} \\
 &\quad - U_j \times \frac{(\mathcal{L}\psi)_{i+1,j} - (\mathcal{L}\psi)_{i-1,j}}{2\Delta x} + U_{yy,j} \times \frac{\psi_{i+1,j} - \psi_{i-1,j}}{2\Delta x},
 \end{aligned} \quad (\text{B.2})$$

where

$$(\mathcal{L}\psi)_{i,j} \equiv \frac{\psi_{i+1,j} - 2\psi_{i,j} + \psi_{i-1,j}}{\Delta x^2} + \frac{\psi_{i,j+1} - 2\psi_{i,j} + \psi_{i,j-1}}{\Delta y^2}, \quad (\text{B.3})$$

for all indices i and j . Clearly, each interior point is related to 12 points in its neighborhood, as shown in Fig. B.1 around point (i, j) . The 2D indices are arranged into a vector form through a mapping $(i, j) \mapsto k$

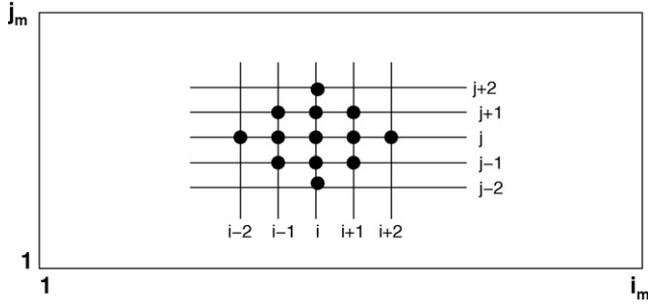


Fig. B.1. Mesh grid.

such that $k = k(i, j) = (j - 1) \times i_m + i$. Let F_ζ be the vector field of the vorticity equation (i.e., the equation for $\mathcal{L}\psi$), and $F_{\zeta, k(i, j)}$ be its k th component, then

$$\begin{aligned} \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{k(i, j)}} &= \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{i, j}} = 0; \\ \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{k(i+1, j)}} &= \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{i+1, j}} = -\frac{1}{2\Delta x} \left(\frac{\mathcal{L}\psi}{\partial y} \right)_{i, j} \\ &\quad - \left(\frac{\partial \psi}{\partial x} \right)_{i, j} \times \frac{1}{2} \left(\left(-\frac{1}{\Delta x^2} \right) / \Delta y + \left(\frac{1}{\Delta x^2} \right) / \Delta y \right) \\ &\quad + \left(\frac{\partial \psi}{\partial y} \right)_{i, j} \times \frac{-\frac{2}{\Delta x^2} - \frac{2}{\Delta y^2}}{2\Delta x} - U_j \times \frac{-\frac{2}{\Delta x^2} - \frac{2}{\Delta y^2}}{2\Delta x} + U_{yy, j} \times \frac{1}{2\Delta x} \\ &= \frac{1}{2\Delta x} \left[U_{yy, j} - \left(\frac{\partial \mathcal{L}\psi}{\partial y} \right)_{i, j} \right] + \frac{1}{\Delta x^2 + \Delta y^2} \left[U_j - \left(\frac{\partial \psi}{\partial y} \right)_{i, j} \right], \end{aligned}$$

where the following shorthands

$$\begin{aligned} \left(\frac{\partial \psi}{\partial x} \right)_{i, j} &\equiv \frac{\psi_{i+1, j} - \psi_{i-1, j}}{2\Delta x}, \\ \left(\frac{\partial \psi}{\partial y} \right)_{i, j} &\equiv \frac{\psi_{i, j+1} - \psi_{i, j-1}}{2\Delta y}, \end{aligned}$$

have been used and will be used throughout. Continuing the derivation,

$$\begin{aligned} \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{k(i+2, j)}} &= \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{i+2, j}} = \left(\frac{\partial \psi}{\partial y} \right)_{i, j} \cdot \frac{1}{2\Delta x^3} - U_j \cdot \frac{1}{2\Delta x^3} = \frac{1}{2\Delta x^3} \left[\left(\frac{\partial \psi}{\partial y} \right)_{i, j} - U_j \right]; \\ \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{k(i-1, j)}} &= \frac{\partial F_{\zeta, k(i, j)}}{\partial \psi_{i-1, j}} = \frac{1}{2\Delta x} \left(\frac{\partial \mathcal{L}\psi}{\partial y} \right)_{i, j} - \left(\frac{\partial \psi}{\partial x} \right)_{i, j} \frac{1}{2} \left(-\frac{1}{\Delta x^2 \Delta y} + \frac{1}{\Delta x^2 \Delta y} \right) \\ &\quad + \left(\frac{\partial \psi}{\partial y} \right)_{i, j} \frac{\frac{2}{\Delta x^2} + \frac{2}{\Delta y^2}}{2\Delta x} - U_j \frac{\frac{2}{\Delta x^2} + \frac{2}{\Delta y^2}}{2\Delta x} + U_{yy, j} \frac{-1}{2\Delta x} = \frac{1}{2\Delta x} \left[\left(\frac{\partial \mathcal{L}\psi}{\partial y} \right)_{i, j} - U_{yy, j} \right] \\ &\quad + \frac{1}{\Delta x^2 + \Delta y^2} \left[\left(\frac{\partial \psi}{\partial y} \right)_{i, j} - U_j \right]; \end{aligned}$$

$$\begin{aligned} \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i-2,j)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i-2,j}} = \left[U_j - \left(\frac{\partial \psi}{\partial y} \right)_{i,j} \right] \frac{1}{2\Delta x^3}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i,j+1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i,j+1}} = - \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \times \frac{1}{2} \left\{ \frac{-\frac{2}{\Delta x^2} - \frac{3}{\Delta y^2}}{\Delta y} + \frac{1}{\Delta y^3} \right\} + \left(\frac{\partial \mathcal{L} \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta y} \\ &= \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \times \left\{ \frac{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}}{\Delta y} \right\} + \frac{1}{2\Delta y} \left(\frac{\partial \mathcal{L} \psi}{\partial x} \right)_{i,j}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i,j+2)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i,j+2}} = - \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta y^3}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i,j-1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i,j-1}} = - \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \times \frac{1}{2} \left\{ -\frac{1}{\Delta y^3} + \frac{\frac{3}{\Delta y^2} + \frac{2}{\Delta x^2}}{\Delta y} \right\} - \frac{1}{2\Delta y} \left(\frac{\partial \mathcal{L} \psi}{\partial x} \right)_{i,j} \\ &= \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \left\{ -\frac{1}{\Delta y} \left(\frac{1}{\Delta y^2} + \frac{1}{\Delta x^2} \right) \right\} - \frac{1}{2\Delta y} \left(\frac{\partial \mathcal{L} \psi}{\partial x} \right)_{i,j}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i,j-2)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i,j-2}} = \frac{1}{2\Delta y^3} \left(\frac{\partial \psi}{\partial x} \right)_{i,j}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i+1,j-1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i+1,j-1}} = \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta y \Delta x^2} + \left(\frac{\partial \psi}{\partial y} \right)_{i,j} \frac{1}{2\Delta x \Delta y^2} - U_j \frac{1}{2\Delta x \Delta y^2}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i+1,j+1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i+1,j+1}} = - \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta x^2 \Delta y} + \left(\frac{\partial \psi}{\partial y} \right)_{i,j} \frac{1}{2\Delta x \Delta y^2} - U_j \frac{1}{2\Delta x \Delta y^2}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i-1,j-1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i-1,j-1}} = \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta x^2 \Delta y} - \left(\frac{\partial \psi}{\partial y} \right)_{i,j} \frac{1}{2\Delta x \Delta y^2} + U_j \frac{1}{2\Delta x \Delta y^2}; \\ \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{k(i-1,j+1)}} &= \frac{\partial F_{\zeta, k(i,j)}}{\partial \psi_{i-1,j+1}} = - \left(\frac{\partial \psi}{\partial x} \right)_{i,j} \frac{1}{2\Delta x^2 \Delta y} - \left(\frac{\partial \psi}{\partial y} \right)_{i,j} \frac{1}{2\Delta x \Delta y^2} + U_j \frac{1}{2\Delta x \Delta y^2}. \end{aligned}$$

These form the $k(i, j)$ th row of the matrix DF_{ζ} ; all other entries at the same row are zero.

References

- Carnevale, G.F., Holloway, G., 1982. Information decay and the predictability of turbulent flows. *J. Fluid Mech.* 116, 115–121.
- Cove, T.M., Thomas, J.A., 1991. *Elements of Information Theory*. Wiley, New York.
- DelSole, T., 2004. Predictability and information theory. Part I. Measures of predictability. *J. Atmos. Sci.* 61 (20), 2425.
- Ehrendorfer, M., Tribbia, J.J., 1997. Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.* 54, 286–313.
- Epstein, E.S., 1969. Stochastic dynamic prediction. *Tellus* 21, 739–759.
- Farrell, B.F., 1990. Small error dynamics and predictability of atmospheric flows. *J. Atmos. Sci.* 47, 2409–2416.
- Farrell, B.F., Ioannou, P.J., 1996. Generalized stability theory. Part I. Autonomous operators. *J. Atmos. Sci.* 53, 2025–2040.
- Haidvogel, D.B., Held, I.M., 1980. Homogeneous quasi-geostrophic turbulence driven by a uniform temperature gradient. *J. Atmos. Sci.* 37, 2644–2660.
- Huang, R.X., 2010. *Ocean Circulation: Wind-Driven and Thermohaline Processes*. Cambridge University Press, 799 pp.
- Kirwan, A.D., Toner, M., Kantha Jr., L., 2003. Predictability, uncertainty, and hyperbolicity in the ocean. *Int. J. Eng. Sci.* 41, 249–258.
- Kleeman, R., 2002. Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.* 59, 2057–2072.
- Kleeman, R., 2007. Information flow in ensemble weather predictions. *J. Atmos. Sci.* 64 (3), 1005–1016.
- Kleeman, R., 2008. Limits, variability and general behavior of statistical predictability of the mid-latitude atmosphere. *J. Atmos. Sci.* 65, 263–275.
- Kuo, H.L., 1973. Dynamics of quasigeostrophic flows and instability theory. In: Yih, C.-S. (Ed.), *Advances in Applied Mechanics*, vol. 13. Academic Press, pp. 247–330.

- Lasota, A., Mackay, M.C., 1994. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer, New York.
- Leith, C.E., 1974. Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.* 102, 409–418.
- Lermusiaux, P.F.J., 2006. Uncertainty estimation and prediction for interdisciplinary ocean dynamics. Special issue on “Uncertainty Quantification”. In: Glimm, J., Karniadakis, G. (Eds.), *J. Comput. Phys.* 217, 176–199.
- Lermusiaux, P.F.J., 2007. Adaptive modeling, adaptive data assimilation, and adaptive sampling. *Physica D* 230, 172–196.
- Liang, X.S., Anderson, D.G.M., 2007. Multiscale window transform. *SIAM J. Multiscale Model. Simul.* 6, 437–467.
- Liang, X.S., 2008. Information flow within stochastic dynamical systems. *Phys. Rev. E* 78, 031113.
- Liang, X.S., Kleeman, R., 2005. Information transfer between dynamical system components. *Phys. Rev. Lett.* 95 (24), 244101.
- Liang, X.S., Kleeman, R., 2007a. A rigorous formalism of information transfer between dynamical system components. I. Discrete mapping. *Physica D* 231, 1–9.
- Liang, X.S., Kleeman, R., 2007b. A rigorous formalism of information transfer between dynamical system components. II. Continuous flow. *Physica D* 227, 173–182.
- Liang, X.S., Robinson, A.R., 2007. Localized multi-scale energy and vorticity analysis. II. Finite-amplitude instability theory and validation. *Dyn. Atmos. Oceans* 44, 51–76.
- Lorenz, E.N., 1963. Deterministic non-periodic flow. *J. Atmos. Sci.* 20, 130–141.
- Marshall, J., Plumb, R.A., 2008. *Atmosphere, Ocean, and Climate Dynamics: An Introductory Text*. Academic Press, 319 pp.
- McWilliams, J.C., Chow, J.H.S., 1981. Equilibrium geostrophic turbulence. I. A reference solution in a β -plane channel. *J. Phys. Ocean* 11, 921–949.
- Miller, R.N., Ehret, L.L., 2002. Ensemble generation for models of multimodal systems. *Mon. Wea. Rev.* 130, 2313–2333.
- Moore, A.M., 1999. The dynamics of error growth and predictability in a model of the Gulf Stream. II. Ensemble prediction. *J. Phys. Oceanogr.* 29, 762–778.
- Mu, M., Duan, W., Wang, J., 2002. The predictability problems in numerical weather and climate prediction. *Adv. Atmos. Sci.* 19, 191–204.
- Nese, J.M., Miller, A.J., Dutton, J.A., 1996. The nature of predictability enhancement in a low-order ocean-atmosphere model. *J. Clim.* 9, 2167–2172.
- Palmer, T.N., 2000. Predicting uncertainty in forecasts of weather and climate. *Rep. Prog. Phys.* 63, 71–116.
- Pedlosky, J., 1987. *Geophysical Fluid Dynamics*. Springer-Verlag, New York, 710 pp.
- Rhines, P.B., 1975. Waves and turbulence on a beta-plane. *J. Fluid Mech.* 69, 417–443.
- Roache, P.J., 1997. Quantification of uncertainty in computation fluid dynamics. *Ann. Rev. Fluid Mech.* 29, 123–160.
- Schneider, T., Griffies, S.M., 1999. A conceptual framework for predictability studies. *J. Clim.* 12, 3133–3155.
- Shukla, J., 1998. Predictability in the midst of chaos: a scientific basis for climate forecasting. *Science* 282, 728–731.
- Smith, L.A., Ziehmann, C., Fraedrich, K., 1999. Uncertainty dynamics and predictability in chaotic systems. *Q. J. R. Meteorol. Soc.* 125, 2855–2886.
- Tang, Y., Kleeman, R., Moore, A.M., 2008. Comparison of information-based measures of forecast uncertainty in ensemble ENSO prediction. *J. Clim.* 21, 230–247.
- Temam, R., 1997. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Springer, New York.
- Toth, Z., Kalnay, E., 1993. Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Bull. Am. Meteorol. Soc.* 74, 2317–2330.