Perturbation Growth and Structure in Uncertain Flows. Part II

Brian F. Farrell

Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts

Petros J. Ioannou

Department of Physics, National and Capodistrian University of Athens, Athens, Greece

ABSTRACT

Perturbation growth in uncertain systems associated with fluid flow is examined concentrating on deriving, solving, and interpreting equations governing the ensemble mean covariance. Covariance evolution equations are obtained for fluctuating operators and illustrative physical examples are solved. Stability boundaries are obtained constructively in terms of the amplitude and structure of operator fluctuation required for existence of bounded second-moment statistics in an uncertain system. The forced stable uncertain system is identified as a primary physical realization of second-moment dynamics by using an ergodic assumption to make the physical connection between ensemble statistics of stable stochastically excited systems and observations of time mean quantities. Optimal excitation analysis plays a central role in generalized stability theory and concepts of optimal deterministic and stochastic excitation of certain systems are extended in this work to uncertain systems. Remarkably, the optimal excitation problem has a simple solution in uncertain systems: there is a pure structure producing the greatest expected ensemble perturbation growth when this structure is used as an initial condition, and a pure structure that is most effective in exciting variance when this structure is used to stochastically force the system distributed in time.

Optimal excitation analysis leads to an interpretation of the EOF structure of the covariance both for the case of optimal initial excitation and for the optimal stochastic excitation distributed in time that maintains the statistically steady state. Concepts of pure and mixed states are introduced for interpreting covariances and these ideas are used to illustrate fundamental limitations on inverting covariances for structure in stochastic systems in the event that only the covariance is known.

1. Introduction

Dynamical equations for evolution of ensemble mean fields under uncertain dynamics were obtained in Farrell and Ioannou (2002, hereafter Part I). Specifically, for the uncertain linear system

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}\boldsymbol{\psi} + \epsilon\xi(t)\mathbf{B}\boldsymbol{\psi}, \qquad (1)$$

an exact equation for the ensemble mean state $\langle\boldsymbol{\psi}(t)\rangle$ initialized at $t = 0$ was obtained, where the bracket $\langle\cdot\rangle$ denotes the ensemble mean over realizations of the $O(1)$ random variable $\xi(t)$. In (1), $\mathbf{A}$ is the ensemble mean operator, ϵ is the amplitude of the operator fluctuations, and $\mathbf{B}$ is the matrix of the fluctuation structure. The random variable $\xi(t)$ is assumed to be stationary with zero mean, unit variance, and autocorrelation time $t_c = 1/\nu$, that is, $\langle\xi(t + \tau)\xi(t)\rangle = e^{-\nu\tau}$. For general $\mathbf{A}$ and $\mathbf{B}$

the equation for the ensemble mean state initialized at $t = 0$ is

$$\frac{d\langle\boldsymbol{\psi}\rangle}{dt} = \left(\mathbf{A} + \epsilon^2\mathbf{B}\int_0^t e^{\mathbf{A}s}\mathbf{B}e^{-\mathbf{A}s}e^{-\nu s}\,ds\right)\langle\boldsymbol{\psi}\rangle, \quad (2)$$

which for $\mathbf{A}$ and $\mathbf{B}$ commuting reduces to

$$\frac{d\langle\boldsymbol{\psi}\rangle}{dt} = \left[\mathbf{A} + \frac{\epsilon^2}{\nu}(1 - e^{-\nu t})\mathbf{B}^2\right]\langle\boldsymbol{\psi}\rangle. \qquad (3)$$

Dynamical equations (2) and (3) are exact for Gaussian $\xi(t)$ and when the fluctuations are Gaussian we replace the general random variable ξ by its Gaussian counterpart η. It was shown in Part I of this paper that (2) and (3) remain accurate approximations for the physically important class of ξ that are Gaussian up to fluctuation amplitude Ξ_0, with zero probability of values greater than Ξ_0, provided that the Kubo number associated with the fluctuations, defined as $K = \epsilon t_c$, is less than Ξ_0. This is important because physical quantities are typically bounded in variation while the Gaussian is not, so that predictions depending on unbounded variation inherent

Corresponding author address: Dr. Brian F. Farrell, Division of Engineering and Applied Sciences, Harvard University, Pierce Hall, 29 Oxford St., Cambridge, MA 02138.
E-mail: farrell@deas.harvard.edu

in an assumed Gaussianity of the operator fluctuations are not physical.

Notable is the short autocorrelation time limit of (2) (Arnold 1992; Sardeshmukh et al. 2001):

$$\frac{d\langle \boldsymbol{\psi} \rangle}{dt} = \left(\mathbf{A} + \frac{\epsilon^2}{\nu} \mathbf{B}^2 \right) \langle \boldsymbol{\psi} \rangle. \qquad (4)$$

This equation is generally valid when the autocorrelation time of the fluctuations is short enough and the Kubo number $K = \epsilon t_c \ll 1$ (Van Kampen 1992). In particular, this equation governs the evolution of $\langle \boldsymbol{\psi} \rangle$ for fluctuations that are temporally white (Arnold 1992). For nonwhite fluctuations, (4) may be regarded as an approximate evolution equation valid for small Kubo numbers that will be referred to as the equivalent white noise approximation. It should be noted that Eqs. (2) and (3) are valid only for evolution from $t = 0$ and in particular cannot be directly used to obtain the response to continuous forcing, while (4) may be so used.

Having obtained equations for the ensemble mean field, $\langle \boldsymbol{\psi} \rangle$, under uncertain dynamics in Part I, we wish now to obtain equations for the ensemble mean covariance, $\langle \mathbf{C} \rangle = \langle \boldsymbol{\psi}\boldsymbol{\psi}^\dagger \rangle$, under similarly uncertain dynamics.

2. Covariance dynamics for operators that are certain but with uncertain initial conditions or uncertain additive forcing

We begin with a review of covariance dynamics of certain systems, which forms the background for analyzing covariance dynamics of uncertain systems and provides an opportunity to introduce the tensor product notation, which greatly facilitates the analysis (cf. Horn and Johnson 1991). For the reader's convenience the basic properties of the tensor product are reviewed in appendix A.

The instantaneous covariance matrix of the state $\boldsymbol{\psi}$ is $\mathbf{C} = \boldsymbol{\psi}\boldsymbol{\psi}^\dagger$, where $\dagger$ denotes hermitian transposition. The covariance matrix $\mathbf{C}$ can be associated with the n^2 vector $\mathbf{c} = \boldsymbol{\psi}^* \otimes \boldsymbol{\psi}$, where $*$ denotes complex conjugation and $\otimes$ denotes the tensor product, with the convention that to each $n \times n$ matrix $\mathbf{M}$, with elements m_{ij}, we associate the n^2 vector, $\mathbf{m}$, with elements

$$\mathbf{m} \equiv [m_{11}, \ldots, m_{n1}, m_{12}, \ldots, m_{n2}, \ldots, m_{1n}, \ldots, m_{nn}]^\mathrm{T}. \qquad (5)$$

Reversing this process, the elements of an n^2 vector are formed into an $n \times n$ matrix by breaking the vector into n equal parts and arranging these as the n columns of the matrix. With this convention, matrix equations can be cast as vector equations. For example, $\mathbf{AC}$ is written using tensor products as $(\mathbf{I} \otimes \mathbf{A})\mathbf{c}$, where $\mathbf{c}$ is the vector associated with matrix $\mathbf{C}$ as described above, $\mathbf{I}$ is the identity matrix, and the tensor product $\mathbf{I} \otimes \mathbf{A}$ is an $n^2 \times n^2$ matrix, the dimension being appropriate for transforming the n^2 vector $\mathbf{c}$. Similarly, the matrix product

$\mathbf{CA}$ is written $(\mathbf{A}^\mathrm{T} \otimes \mathbf{I})\mathbf{c}$, where T denotes unconjugated transposition.

Consider a state vector, $\boldsymbol{\psi}$, which for each realization of the forcing, $f(t)$, satisfies the linear time-dependent equation

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}(t)\boldsymbol{\psi} + f(t), \qquad (6)$$

in which the linear operator, $\mathbf{A}(t)$, is a certain (not random) matrix function of time, and $\boldsymbol{\psi}$ is an n-dimensional column vector.

We see from (6) using the properties of the tensor product that each realization of the vector covariance, $\mathbf{c} = \boldsymbol{\psi}^* \otimes \boldsymbol{\psi}$, satisfies the equation

$$\frac{d(\boldsymbol{\psi}^* \otimes \boldsymbol{\psi})}{dt} = \frac{d\boldsymbol{\psi}^*}{dt} \otimes \boldsymbol{\psi} + \boldsymbol{\psi}^* \otimes \frac{d\boldsymbol{\psi}}{dt}$$
$$= \mathbf{I} \otimes \mathbf{A}(t)\boldsymbol{\psi}^* \otimes \boldsymbol{\psi} + \mathbf{A}(t)^* \otimes \mathbf{I}\boldsymbol{\psi}^* \otimes \boldsymbol{\psi}$$
$$+ f^* \otimes \boldsymbol{\psi} + \boldsymbol{\psi}^* \otimes f. \qquad (7)$$

Consider forming the ensemble average of the vector covariance over forcing realizations, with the ensemble average indicated by $[\mathbf{c}]$. Notice that the notation $[\cdot]$ denotes ensembling over forcing realizations rather than over operator realizations, which is indicated by $\langle \cdot \rangle$. For delta-correlated stochastic forcing, $[f(t)f^\dagger(s)] = \mathbf{Q}\delta(t - s)$ with forcing structure matrix $\mathbf{Q} \equiv [ff^\dagger]$, it can be shown[1] that the rhs term, representing correlation of the state and the forcing in (7), has ensemble average

$$[f^* \otimes \boldsymbol{\psi} + \boldsymbol{\psi}^* \otimes f] = [f^* \otimes f] = q, \qquad (8)$$

where q is the covariance vector corresponding to covariance matrix $\mathbf{Q}$, so that we obtain from (7) the covariance evolution equation,

$$\frac{d[\mathbf{c}]}{dt} = \mathbf{I} \otimes \mathbf{A}(t)[\mathbf{c}] + \mathbf{A}(t)^* \otimes \mathbf{I}[\mathbf{c}] + q, \qquad (9)$$

which is the tensor form of the familiar Lyapunov equation (Farrell and Ioannou 1996, hereafter FI96),

$$\frac{d[\mathbf{C}]}{dt} = \mathbf{A}(t)[\mathbf{C}] + [\mathbf{C}]\mathbf{A}(t)^\dagger + \mathbf{Q}. \qquad (10)$$

The tensor covariance equation (9) makes apparent that the evolution operator of the covariance matrix is the certain Lyapunov superoperator

$$L_c = \mathbf{I} \otimes \mathbf{A}(t) + \mathbf{A}(t)^* \otimes \mathbf{I}. \qquad (11)$$

The prefix "super" is introduced as a reminder that this operator acts on matrices associated with the state, albeit

---

[1] If $\mathbf{P}(t, s)$ is the propagator of $\mathbf{A}(t)$, then $\psi(t) = \int_0^t \mathbf{P}(t, s)f(s) \, ds$. The fact that $\mathbf{P}(t, t) = \mathbf{I}$ and $[f(t)^* \otimes f(s)] = \delta(t - s)q$ implies that $[f^* \otimes \boldsymbol{\psi}] = \int_0^t [f^*(t) \otimes \mathbf{P}(t, s)f(s)] \, ds = (1/2)q$, where $\mathbf{q}$ is the vector covariance associated with $\mathbf{Q} = [ff^\dagger]$. The factor of 1/2 in the above expression comes from the conventional property of the delta function: $\int_0^t \delta(t - s) \, ds = 1/2$. Similarly, $[\boldsymbol{\psi}^* \otimes f] = (1/2)q$ and (9) and (10) follow.

matrices in vector form, and not on state vectors themselves. The superoperator has dimension $n^2 \times n^2$ if the state vector has dimension n. We use the term covariance to refer to the vector $\mathbf{c}$ and to the equivalent matrix $\mathbf{C}$, as appropriate.

Consider evolution of the covariance $\mathbf{c}(0) = \boldsymbol{\psi}(0)^* \otimes \boldsymbol{\psi}(0)$ associated with a sure (not random) initial state $\boldsymbol{\psi}(0)$ and for simplicity take the certain (not random) operator $\mathbf{A}$ to be autonomous. At time t, the covariance is

$$\mathbf{c}(t) = \exp(L_c t)\,\mathbf{c}(0). \qquad (12)$$

If the initial conditions are uncertain with ensemble average initial covariance $[\mathbf{c}(0)]$, then the ensemble average covariance at time t is

$$[\mathbf{c}(t)] = \exp(L_c t)[\mathbf{c}(0)]. \qquad (13)$$

If $\mathbf{A}$ is time independent and stable, the ensemble average covariance approaches asymptotically:

$$[\mathbf{c}(\infty)] = -(\mathbf{I} \otimes \mathbf{A} + \mathbf{A}^* \otimes \mathbf{I})^{-1}\,q. \qquad (14)$$

In matrix notation, (14) is the familiar Lyapunov equation (cf. FI96),

$$\mathbf{A}[\mathbf{C}_\infty] + [\mathbf{C}_\infty]\mathbf{A}^\dagger = -\mathbf{Q}. \qquad (15)$$

Pure and mixed states and the interpretation of EOF analysis

When the state of the system, $\boldsymbol{\psi}$, is known exactly the system is said to be in a pure state. In this case the associated covariance $\mathbf{c} = \boldsymbol{\psi}^* \otimes \boldsymbol{\psi}$ has rank one. Conversely, covariances of rank one are associated with systems in a pure state, and this state can be determined within a phase by singular value decomposition (SVD) of the covariance matrix. Because physical states are real, the arbitrariness in the phase of the state is resolved, and rank one covariances can be inverted for a pure structure, within a sign; and association of a covariance with a state is unambiguous if the system is in a pure state.

However, covariances with rank exceeding one can not be inverted for structure and we say that the state of the system is mixed and that the rank of the covariance identifies the number of states involved in producing the covariance. When the covariance of the system is mixed there does not exist an unambiguous preparation of the system with the given covariance and a large set of structures $\boldsymbol{\psi}$ could equally well have been responsible for producing the covariance.

The familiar empirical orthogonal function (EOF) analysis into orthonormal structures is obtained from singular value decomposition of the covariance matrix.[2] It is tempting to regard the resulting EOF basis as identifying the states that produced the covariance or at least to regard the EOFs and the singular values of the co-

variance matrix as having special significance among the possible ensembles of states that could have produced the covariance. Although this is valid for our examples, it is unfortunately not the case in general; for example, consider the 2×2 covariance with EOFs $\mathbf{e}_1$ with eigenvalue $\boldsymbol{\alpha}_1$ and its orthogonal $\mathbf{e}_2$ with eigenvalue $\boldsymbol{\alpha}_2$ and with $\boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2$. The resulting covariance matrix is

$$\mathbf{C} = \boldsymbol{\alpha}_1 \mathbf{e}_1 \mathbf{e}_1^\dagger + \boldsymbol{\alpha}_2 \mathbf{e}_2 \mathbf{e}_2^\dagger, \qquad (16)$$

or in tensor product notation,

$$\mathbf{c} = \boldsymbol{\alpha}_1 \mathbf{e}_1^* \otimes \mathbf{e}_1 + \boldsymbol{\alpha}_2 \mathbf{e}_2^* \otimes \mathbf{e}_2. \qquad (17)$$

Consider now the quite different states:

$$\mathbf{e}_+ = \sqrt{\boldsymbol{\alpha}_1}\mathbf{e}_1 + \sqrt{\boldsymbol{\alpha}_2}\mathbf{e}_2 \qquad (18\mathrm{a})$$

$$\mathbf{e}_- = \sqrt{\boldsymbol{\alpha}_1}\mathbf{e}_1 - \sqrt{\boldsymbol{\alpha}_2}\mathbf{e}_2. \qquad (18\mathrm{b})$$

It can be verified that an ensemble in which $\mathbf{e}_+$ and $\mathbf{e}_-$ appear with equal probability produces the same covariance matrix $\mathbf{C}$,

$$\mathbf{C} = \frac{1}{2}\mathbf{e}_+\mathbf{e}_+^\dagger + \frac{1}{2}\mathbf{e}_-\mathbf{e}_-^\dagger = \boldsymbol{\alpha}_1\mathbf{e}_1\mathbf{e}_1^\dagger + \boldsymbol{\alpha}_2\mathbf{e}_2\mathbf{e}_2^\dagger, \quad (19)$$

demonstrating the ambiguity inherent in associating states with covariances. Note, however, that the assumed orthogonality of the original states $\mathbf{e}_1$ and $\mathbf{e}_2$ has not been preserved by the transformation forming $\mathbf{e}_+$ and $\mathbf{e}_-$, as is immediately verified by forming the dot product $\mathbf{e}_+^\dagger\mathbf{e}_- = \boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 \neq 0$.

In general, EOFs are one basis out of an infinity of possible bases that give rise to the same covariance matrix. The EOF's are special in that they form the only[3] orthogonal basis, and also in that the EOFs form the optimal basis for $\mathbf{C}$ in the least squares sense.[4] However, there is generally no a priori reason based solely on the covariance to suppose that the EOF basis in fact produced a given covariance matrix.[5] Accepting this ambiguity in identifying the states producing a given $\mathbf{C}$, one may inquire if there exist restrictions on the class of candidate bases for a given covariance.

We can construct an infinity of nonorthogonal states that produce the same covariance matrix $\mathbf{C}$ by generalizing the simple example given earlier. Consider[6] $\mathbf{e}^i$ to be the ith EOF with associated singular value $\boldsymbol{\alpha}_i$ and define the orthogonal but not orthonormal set $\boldsymbol{\psi}^i = \sqrt{\boldsymbol{\alpha}_i}\mathbf{e}^i$ (no summation). Then the covariance $\mathbf{C}$ can be written in vector form as

$$\mathbf{c} = \sum_i \boldsymbol{\psi}^{i*} \otimes \boldsymbol{\psi}^i. \qquad (20)$$

Consider now a unitary matrix $\mathbf{U}$ ($\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$) that associates the set of states $\boldsymbol{\psi}^i$ with a new set of states

---

[2] Singular values and vectors of the necessarily hermitian covariance matrices can be obtained by eigenanalysis.

[3] Unless the covariance is the identity.

[4] This is due to the Eckart–Young–Mirsky property of singular value decomposition (cf. Van Huffel and Vandewalle 1991).

[5] For an interesting example of this see Ambaum et al. (2001).

[6] Superscripts will be used here to distinguish specific vectors of the base from their components.

by $\boldsymbol{\phi}^i = \Sigma_j \, U_{ij}\boldsymbol{\psi}^j$ (note that the sum acts on the the states themselves and not their components, so that $\boldsymbol{\phi}^{i\dagger} = \Sigma_j \, U_{ij}^*\boldsymbol{\psi}_j^\dagger$). Then the states $\boldsymbol{\psi}^i$ and $\boldsymbol{\phi}^i$ form the same covariance $\mathbf{C}$. Indeed,

$$\sum_i \boldsymbol{\phi}^{i*} \otimes \boldsymbol{\phi}^i = \sum_i \left(\sum_j \sum_k U_{ij}^* U_{ik} \boldsymbol{\psi}^{j*} \otimes \boldsymbol{\psi}^k \right)$$

$$= \sum_j \sum_k \boldsymbol{\psi}^{j*} \otimes \boldsymbol{\psi}^k \left(\sum_i U_{ij}^* U_{ik} \right)$$

$$= \sum_j \sum_k \boldsymbol{\psi}^{j*} \otimes \boldsymbol{\psi}^k \delta_{jk} = \mathbf{c}. \qquad (21)$$

This unitary matrix of coefficients is the maximum allowed freedom in the choice of the ensembles (Schrödinger 1936; Hughston et al. 1993). Also, it is easily seen that $\boldsymbol{\psi}^i$ form the only orthogonal basis because the states $\boldsymbol{\phi}^i$ cannot be orthogonal as

$$\boldsymbol{\phi}^{i\dagger}\boldsymbol{\phi}^j = \sum_{kl} U_{ik}^* U_{jl}(\boldsymbol{\psi}^{k\dagger}\boldsymbol{\psi}^l) = \sum_k \boldsymbol{\alpha}_k U_{ik}^* U_{jk} \neq 0, \quad (22)$$

unless the $\boldsymbol{\psi}^i$ are orthonormal and $\mathbf{C}$ is the identity.

We conclude that there is a unique orthogonal basis for $\mathbf{C}$ that is the EOF's, and another nonorthogonal candidate basis for every possible unitary transformation. While in general there is no a priori reason based on the covariance alone to conclude that the orthogonal basis produced a given $\mathbf{C}$, if the additional assumption is made of Gaussianity of the pdf of states, based, for example, on observation or on the system being linear and forced by a spatially and temporally white noise process, then the EOF basis is required for representation of the Gaussian pdf in EOF coordinates. On the other hand, we may have observational evidence for systematic occurrence of preferred states associated with one of the possible nonorthogonal bases; however, this would imply a non-Gaussian pdf for the states.

Advancing the covariance matrix under certain dynamics does not alter the initial rank of the covariance. This follows from the uniqueness of solutions of certain differential equations. To each sure initial covariance $c(0) = \psi(0)^* \otimes \psi(0)$ produced from the pure state $\psi(0)$ corresponds at a time later, t, the covariance $c(t) = \psi(t)^* \otimes \psi(t)$, which is also rank one. Similarly, if the initial covariance has rank $r \leq n$, it can be expanded in its EOF basis:

$$c(0) = \sum_{i=1}^r \lambda_i e_i(0)^* \otimes e_i(0), \qquad (23)$$

where $e_i(0)$ are the orthogonal eigenfunctions of $c(0)$ and all $\lambda_i > 0$. At time t, $c(t)$ becomes

$$c(t) = \sum_{i=1}^r \lambda_i e_i(t)^* \otimes e_i(t), \qquad (24)$$

where all the $e_i(t) = \mathbf{P}(t)e_i(0)$ are distinct, because of the uniqueness of the solutions, and as a result the rank of the covariance is preserved under certain dynamics.

It may prove useful for the development that follows to isolate the property of the certain superoperator (11),

$$L_c = \mathbf{I} \otimes \mathbf{A}(t) + \mathbf{A}(t)^* \otimes \mathbf{I}, \qquad (25)$$

that leads to the invariance of rank. Consider again the rank one initial covariance $c(0) = \psi(0)^* \otimes \psi(0)$, which at t is advanced to $c(t) = P(t)c(0)$, where $P(t)$ is the propagator associated with the certain superoperator L_c. Because $c(t) = \psi(t)^* \otimes \psi(t)$ we can express the propagator of the superoperator $P(t)$ in terms of the propagator $\mathbf{P}(t)$, which advances $\psi(0)$ as

$$P(t) = \mathbf{P}(t)^* \otimes \mathbf{P}(t). \qquad (26)$$

This is true because

$$c(t) = \psi(t)^* \otimes \psi(t) = (\mathbf{P}(t)\psi(0))^* \otimes (\mathbf{P}(t)\psi(0))$$

$$= (\mathbf{P}^*(t) \otimes \mathbf{P}(t))(\psi(0)^* \otimes \psi(0))$$

$$= P(t)c(0) \qquad (27)$$

The property of the certain superoperator that makes possible the splitting (26) of the propagator of the covariance dynamics as the tensor product of the propagators of the dynamics $\mathbf{P}(t)$ is the commutation[7] of $\mathbf{I} \otimes \mathbf{A}$ and $\mathbf{A}^* \otimes \mathbf{I}$. Under uncertain dynamics the superoperator cannot be split into commuting parts and as a result the propagator $P(t)$ is not the tensor product of the propagators as in (26) with the consequence that the rank of the covariance is not preserved. As a result under uncertain dynamics, if initially we start with a pure state, at a later time the state of the system will become mixed.

3. Covariance dynamics of uncertain systems

Consider an uncertain mean flow with ensemble mean operator $\mathbf{A}$, and fluctuation operator $\epsilon\eta(t)\mathbf{B}$, with $\eta(t)$ a Gaussian zero mean, unit variance process with autocorrelation time t_c. According to (9) with $\mathbf{A}(t) = \mathbf{A} + \epsilon\eta(t)\mathbf{B}$ for each realization of the uncertain flow the covariance formed after averaging over the realizations of the (additive) forcing evolves as

$$\frac{dc}{dt} = A\mathbf{c} + \epsilon\eta(t)B\mathbf{c} + q, \qquad (28)$$

in which the symbol $[\cdot]$, denoting the ensemble average over (additive) forcing realizations, has been suppressed. The mean and fluctuation superoperators in (28) are

$$A = \mathbf{I} \otimes \mathbf{A} + \mathbf{A}^* \otimes \mathbf{I},$$

$$B = \mathbf{I} \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{I}. \qquad (29)$$

The ensemble average evolution equation for the covariance, $\mathbf{c}$ over realizations of η, indicated by $\langle\cdot\rangle$, and valid for evolution from $t = 0$, is obtained by an argument similar to that leading to (2) with result

---

[7] Because $(\mathbf{I} \otimes \mathbf{A})(\mathbf{A}^* \otimes \mathbf{I}) = \mathbf{A}^* \otimes \mathbf{A} = (\mathbf{A}^* \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{A})$.

$$\frac{d\langle\mathbf{c}\rangle}{dt} = \{A + \epsilon^2 BD(t)\}\langle\mathbf{c}\rangle, \qquad (30)$$

where

$$D(t) = \int^{t_0} \mathbf{e}^{As}B\mathbf{e}^{-As}\mathbf{e}^{-\nu s} \, ds. \qquad (31)$$

Repeated application of the tensor product property [Eq. (A2)] gives

$$BD(t) = (\mathbf{I} \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{I})\{\mathbf{D}^*(t) \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{D}(t)\}$$
$$= \mathbf{D}^*(t) \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{D}(t) + \mathbf{I} \otimes \mathbf{BD}(t)$$
$$+ \mathbf{B}^*\mathbf{D}^*(t) \otimes \mathbf{I}, \qquad (32)$$

where

$$\mathbf{D}(t) = \int_0^t \mathbf{e}^{\mathbf{A}s}\mathbf{B}\mathbf{e}^{-\mathbf{A}s}\mathbf{e}^{-\nu s} \, ds, \qquad (33)$$

and the ensemble average covariance evolution equation can be written

$$\frac{d\langle\mathbf{c}\rangle}{dt} = L\langle\mathbf{c}\rangle, \qquad (34)$$

in which appears the generalized Lyapunov superoperator L:

$$L = \mathbf{I} \otimes \mathbf{A} + \mathbf{A}^* \otimes \mathbf{I} + \epsilon^2\{\mathbf{I} \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{I}\}$$
$$\times \{\mathbf{I} \otimes \mathbf{D}(t) + \mathbf{D}(t)^* \otimes \mathbf{I}\}. \qquad (35)$$

As was previously demonstrated in the argument leading to (4), for short autocorrelation times (35) reduces to the familiar equivalent white noise superoperator:

$$L_w = \mathbf{I} \otimes \mathbf{A} + \mathbf{A}^* \otimes \mathbf{I} + \frac{\epsilon^2}{\nu}(\mathbf{I} \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{I})^2$$
$$= \mathbf{I} \otimes \left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right) + \left(\mathbf{A}^* + \frac{\epsilon^2}{\nu}\mathbf{B}^{*2}\right) \otimes \mathbf{I}$$
$$+ \frac{2\epsilon^2}{\nu}\mathbf{B}^* \otimes \mathbf{B}. \qquad (36)$$

In matrix notation the ensemble covariance dynamical equation corresponding to (34), which is valid for any fluctuation autocorrelation time, $t_c = 1/\nu$, and for evolution from $t = 0$ is

$$\frac{d\langle\mathbf{C}\rangle}{dt} = \{\mathbf{A} + \epsilon^2\mathbf{BD}(t)\}\langle\mathbf{C}\rangle + \langle\mathbf{C}\rangle\{\mathbf{A}^\dagger + \epsilon^2\mathbf{D}^\dagger(t)\mathbf{B}^\dagger\}$$
$$+ \epsilon^2\{\mathbf{B}\langle\mathbf{C}\rangle\mathbf{D}^\dagger(t) + \mathbf{D}(t)\langle\mathbf{C}\rangle\mathbf{B}^\dagger\}. \qquad (37)$$

The short autocorrelation limit of (37) is (Arnold 1992: Sardeshmukh et al. 2001)

$$\frac{d\langle\mathbf{C}\rangle}{dt} = \left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right)\langle\mathbf{C}\rangle + \langle\mathbf{C}\rangle\left(\mathbf{A}^\dagger + \frac{\epsilon^2}{\nu}\mathbf{B}^{2\dagger}\right)$$
$$+ \frac{2\epsilon^2}{\nu}\mathbf{B}\langle\mathbf{C}\rangle\mathbf{B}^\dagger. \qquad (38)$$

In the following the $\langle\cdot\rangle$ symbol indicating ensemble average over operator realizations is also suppressed, so that when $\mathbf{C}$ appears it is understood to be the ensemble average covariance with the ensemble average taken over the forcing realizations and/or operator realizations as appropriate.

a. Steady-state covariance maintained by stochastic forcing

Under the ergodic assumption the asymptotic ensemble average covariance, when it exists, equals the covariance obtained by taking a long time average of a single realization of the flow. It exists when the uncertain superoperator L is asymptotically decaying in which case the forced covariance is

$$\mathbf{c}_\infty = \lim_{t\to\infty}\left\{\int_0^t P(t, s) \, ds\right\}q, \qquad (39)$$

where $P(t, s)$ is the propagator associated with the uncertain ensemble mean covariance dynamics.

In the short autocorrelation time limit if the generalized Lyapunov operator, L_w, is asymptotically stable and the ensemble mean operator, $\mathbf{A}$, is time independent, then (39) converges to

$$\mathbf{c}_\infty = -L_w q. \qquad (40)$$

In equivalent matrix notation the equilibrium covariance satisfies the generalized Lyapunov equation for steady state $\mathbf{C}$ (Arnold 1992):

$$\left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right)\mathbf{C} + \mathbf{C}\left(\mathbf{A}^\dagger + \frac{\epsilon^2}{\nu}\mathbf{B}^{2\dagger}\right) + \frac{2\epsilon^2}{\nu}\mathbf{BCB}^\dagger = -\mathbf{Q}.$$
$$(41)$$

Note that the familiar Lyapunov equation (15) has been modified in two ways: the mean operator has been augmented by a term proportional to $\mathbf{B}^2$, and the term $\mathbf{BCB}^\dagger$ appears.

b. Examples of covariance dynamics

The streamfunction for the barotropic flow with fluctuating wind example introduced in Part I is governed by

$$\frac{d\boldsymbol{\psi}}{dt} = -ik(U_0 + \epsilon\eta(t))\boldsymbol{\psi} - r\boldsymbol{\psi}. \qquad (42)$$

Consider a single wave with wavenumber, k, in which case the mean operator is $\mathbf{A} = -ikU_0 - r$, the fluctuation operator is $\mathbf{B} = -ik$, and $\mathbf{B}$ commutes with $\mathbf{A}$. In this case both superoperators $B = \mathbf{I} \otimes \mathbf{B} + \mathbf{B}^* \otimes \mathbf{I}$ and $D = \mathbf{I} \otimes \mathbf{D}(t) + \mathbf{D}^*(t) \otimes \mathbf{I}$ vanish in (35), so the mean covariance evolves as in the certain problem, and the asymptotic covariance maintained by the fluctuating operator is identical to that maintained by the mean operator. From a physical perspective the reason is that

wind fluctuations in this example change the phase of the waves but do not affect the wave structure or growth. As a result each realization has the same square amplitude and the moment growth rates are identical. From the point of view of the covariance equation, phase averaging in the ensemble mean equation accounted for by the $\mathbf{B}^2$ term in (37) or (38) is exactly canceled by the $\mathbf{BCB}^\dagger$ term.

The effect of the fluctuations is clarified by considering the equation for the variance, which is trace $(\mathbf{C})$ = trace$(\psi\psi^\dagger)$ = $\psi^\dagger\psi$:

$$\frac{d\langle\psi^\dagger\psi\rangle}{dt} = \psi^\dagger[\mathbf{A} + \mathbf{A}^\dagger + \epsilon\eta(t)(\mathbf{B} + \mathbf{B}^\dagger)]\psi. \quad (43)$$

If the operator $\mathbf{B} + \mathbf{B}^\dagger$ vanishes identically, as is the case for barotropic wind fluctuation, then the evolution of $\psi\psi^\dagger$ is not influenced by the fluctuations. Consider the ensemble average of (43) obtained by taking the trace of (37):

$$\frac{d\langle\psi^\dagger\psi\rangle}{dt}$$

$$= \text{trace}\{[\mathbf{A} + \mathbf{A}^\dagger + \epsilon^2((\mathbf{B} + \mathbf{B}^\dagger)\mathbf{D} + \mathbf{D}^\dagger(\mathbf{B} + \mathbf{B}^\dagger))]\langle\mathbf{C}\rangle\},$$
$$(44)$$

which has the short autocorrelation limit

$$\frac{d\langle\psi^\dagger\psi\rangle}{dt}$$

$$= \text{trace}\left[\left(\mathbf{A} + \mathbf{A}^\dagger + \frac{\epsilon^2}{\nu}[(\mathbf{B} + \mathbf{B}^\dagger)\mathbf{B} + \mathbf{B}^\dagger(\mathbf{B} + \mathbf{B}^\dagger)]\right)\langle\mathbf{C}\rangle\right].$$
$$(45)$$

It is important to notice that the influence of operator fluctuations on the ensemble mean variance depends on the symmetrized operator $\mathbf{B} + \mathbf{B}^\dagger$ rather than on $\mathbf{B}$ itself, indicating that the fluctuations contribute to second-moment perturbation growth through the hermitian part of the fluctuation operator. Through this term fluctuations of sufficient amplitude generically render the covariance equation asymptotically unstable. For example, if fluctuation in the coefficient of Rayleigh friction were accounted for by $\mathbf{B} = -1$, then the ensemble average covariance equation in the short autocorrelation time limit becomes

$$\frac{d\langle\mathbf{C}\rangle}{dt} = \left(-ikU_0 - r + \frac{\epsilon^2}{\nu}\right)\langle\mathbf{C}\rangle$$

$$+ \langle\mathbf{C}\rangle\left(ikU_0 - r + \frac{\epsilon^2}{\nu}\right) + \frac{2\epsilon^2}{\nu}\langle\mathbf{C}\rangle$$

$$= \left(-2r + \frac{4\epsilon^2}{\nu}\right)\langle\mathbf{C}\rangle, \quad (46)$$

making the second moment unstable for fluctuations of

rms magnitude $\epsilon > (r\nu/2)^{1/2}$, with second-moment exponent $\lambda_2 = -r + 2\epsilon^2/\nu$.

Consider two waves, with wavenumbers k_1 and k_2, and fluctuation operator

$$\mathbf{B} = \begin{pmatrix} -ik_1 & 0 \\ 0 & -ik_2 \end{pmatrix}. \quad (47)$$

In this example the fluctuation superoperators B and B^2 do not vanish, but the variance remains the same as that maintained by the mean operator, $\mathbf{A}$, because $\mathbf{B} + \mathbf{B}^\dagger = 0$ and the second-moment exponent is equal to the growth rate of the mean operator. However, the fluctuations may result in EOFs differing from those obtained in the absence of fluctuations.

c. Further properties of the Lyapunov superoperator

We wish to further analyze second-moment stability and gather tools needed to obtain optimal excitation structures for both deterministic and stochastic forcing of uncertain systems. First, we must collect some properties of the Lyapunov superoperator.

A general matrix could be advanced in time with any of the following: the certain Lyapunov superoperator L_c given by (11); the generalized Lyapunov superoperator, L, given by (35); or its white noise form, L_w, given by (36). The matrices of interest, however, are physically realizable covariance matrices that have the properties of being hermitian and positive definite. Care must be taken in the analysis that only physically realizable solutions to (37) or (38) are retained. Specifically, the individual eigenfunctions of the Lyapunov superoperator, L, are not necessarily either hermitian or positive definite, and physically realizable covariances must be composed of superpositions of eigenfunctions that do have these properties. For example, consider the optimal perturbation problem: determine the initial covariance matrix leading to maximum ensemble mean energy growth at a chosen time. This optimization is complicated by the fact that it must be done in the set of positive definite hermitian matrices, a convex subset of the set of complex matrices that does not form a subspace and as a result is not spanned by a linear basis.[8]

We require some basic properties of the superoperator L. First, L preserves the hermiticity of $\mathbf{C}$. This is easily seen by noticing that the hermitian transpose of a covariance matrix, $\mathbf{C}(t)$, obeys the same dynamical equa-

---

[8] That the positive definite (hermitian) matrices do not form a subspace is clear on noting that negative multiples of a member of the set is not in the set. However, it follows immediately that given two positive definite covariances $\mathbf{C}_1$ and $\mathbf{C}_2$ we can always construct another covariance as a convex linear combination of the two. Indeed $\mathbf{C} = \lambda\mathbf{C}_1 + (1 - \lambda)\mathbf{C}_2$, with $1 \geq \lambda \geq 0$, is hermitian, and because for any state ψ, $\psi^\dagger\mathbf{C}\psi = \lambda(\psi^\dagger\mathbf{C}_1\psi) + (1 - \lambda)(\psi^\dagger\mathbf{C}_2\psi) > 0$, $\mathbf{C}$ is also positive definite given that $\mathbf{C}_{1,2}$ are. (A subset of a vector space is said to be convex if the set contains the straight line segment connecting any two points in the set.)

tion as does the difference $\mathbf{C}_d(t) = \mathbf{C}(t) - \mathbf{C}^\dagger(t)$, so that if initially $\mathbf{C}$ is hermitian, that is, $\mathbf{C}_d(0) = 0$, then by necessity $\mathbf{C}_d(t) = 0$ for all times. Similarly, L preserves the antihermiticity of an antihermitian matrix. Second, L (or L_c) preserves the positivity of $\mathbf{C}$. This property is easy to derive if we step back to the physical origin of the ensemble average equations. Consider a positive definite initial covariance matrix $\mathbf{C}(0)$. Decompose $\mathbf{C}(0)$ by SVD into its EOFs in the form

$$\mathbf{C}(0) = \sum_i \lambda_i \mathbf{e}_i^* \otimes \mathbf{e}_i, \qquad (48)$$

where λ_i can be identified with the positive eigenvalues of $\mathbf{C}(0)$ and $\mathbf{e}_i$ with the corresponding eigenvectors. Because the sum of any positive multiple of positive definite matrices is positive definite, it is sufficient to prove that each $\mathbf{e}_i^* \otimes \mathbf{e}_i$ evolves to a positive definite matrix. But this must be so, for if $\mathbf{P}(t)$ denotes the propagator associated with any realization of the flow, the covariance at time t is $\mathbf{C}(t) = \langle \mathbf{P}(t)^* \mathbf{e}_i^* \otimes \mathbf{P}(t) \mathbf{e}_i \rangle$, which is positive definite.

Third, the eigenvalue of L with maximum real part is necessarily real and as a result L satisfies the principle of exchange of stabilities (Drazin and Reid 1981). For if the maximally growing eigenvalue had an imaginary part a positive definite initial covariance matrix would, with time, acquire negative diagonal entries, which is not possible.

A basic difference between the certain superoperator L_c and its generalized counterparts L is that while the former preserves the rank of the covariance matrix, the latter do not. Consider a sure initial state with a (necessarily) rank-one covariance. Each realization of the operator fluctuations yields a pure state, but under fluctuations the ensemble-averaged state becomes mixed, with covariance rank greater than one (see Fig. 1). This property of increase of the rank of the covariance can be traced to the $\mathbf{BCB}^\dagger$ term or its finite t_c generalization in the ensemble average covariance equations (37)–(38).

d. Determining second-moment stability boundaries for an uncertain system

The mechanism underlying second-moment destabilization is transparent in the case of commuting $\mathbf{A}$ and $\mathbf{B}$. Although restrictive, commuting $\mathbf{A}$ and $\mathbf{B}$ has example in the meteorological context: in the midlatitudes a time modulation of the jet strength is commonly observed and if the mean jet is idealized as a velocity profile linearly increasing with height and the jet fluctuations are assumed to be of the same form, then the operator $\mathbf{A} + \epsilon\mathbf{B}\eta(t)$ governing evolution of perturbations commutes with the mean operator, $\mathbf{A}$. That $\mathbf{A}$ and $\mathbf{B}$ commute assures that the fluctuations do not change the sample (Lyapunov) stability of $\mathbf{A}$ (Farrell and Ioannou 1999). However, perturbation energy, regarded as an ensemble average, is a second-moment quantity and may grow without bound.
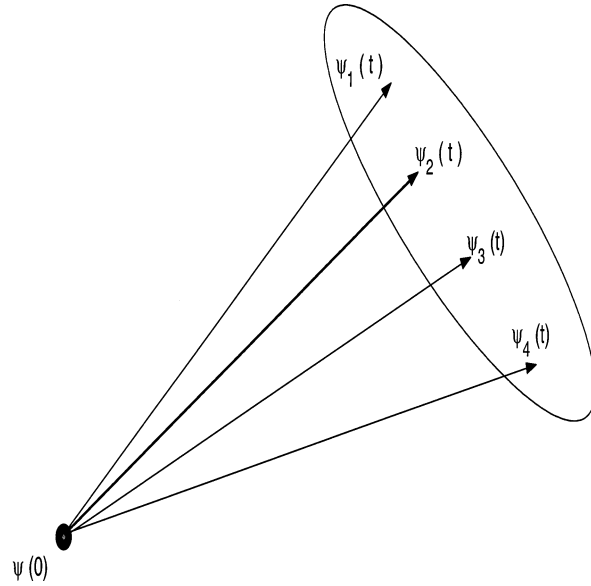

FIG. 1. Schematic evolution of a sure initial condition $\psi(0)$ in an uncertain system. After time t the evolved states $\psi(t)$ lie in the region shown. Initially the covariance matrix $\mathbf{C}(0) = \psi(0)\psi^\dagger(0)$ is rank one, but at time t the covariance matrix is of rank greater than one. For example, if the final states were $\psi_i(t)(i = 1, \ldots, 4)$ with equal probability, the covariance at time $t[\mathbf{C}(t) = 1/4 \sum_{i=1}^4 \psi_i(t)\psi_i^\dagger(t)]$ would be rank four, representing a mixed state. By contrast, in certain systems the rank of the covariance matrix is invariant and a pure state evolves to a pure state.

For commuting $\mathbf{A}$ and $\mathbf{B}$ the generalized Lyapunov superoperator for all autocorrelation times is given by the asymptotic form of the operator, which is identical to the short autocorrelation time superoperator (36):

$$L_w = \mathbf{I} \otimes \left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right) + \left(\mathbf{A}^* + \frac{\epsilon^2}{\nu}\mathbf{B}^{*2}\right) \otimes \mathbf{I}$$
$$+ \frac{2\epsilon^2}{\nu}\mathbf{B}^* \otimes \mathbf{B}. \qquad (49)$$

Eigenanalysis of this operator is immediate; for if ϕ_i are the eigenfunctions of $\mathbf{A}$ and $\mathbf{B}$ with eigenvalues λ_i and μ_i, respectively, the eigenfunctions of L_w are $\phi_i^* \otimes \phi_j$ with eigenvalues

$$\lambda_j + \lambda_i^* + \frac{\epsilon^2}{\nu}(\mu_j + \mu_i^*)^2. \qquad (50)$$

It follows that the maximally growing eigenfunction $\mathbf{c}_{max}$ is composed of the same eigenfunction from each of $\mathbf{A}$ and $\mathbf{B}$ and so has this common structure, that is, $\mathbf{c}_{max} = \phi_\alpha^* \otimes \phi_\alpha$. The greatest second-moment growth rate according to (50)[9] is

---

[9] That the maximum real part of (50) occurs for $i = j = \alpha$ for some α follows because for any two numbers a and b it can be shown by simple algebraic manipulations that $\Re\{a + b^* + (c + d^*)^2\} \le \max\{2\Re(a) + 4\Re(c)^2, 2\Re(b) + 4\Re(d)^2\}$. Recall also that the second-moment growth rate is equal to the maximum eigenvalue of (49) divided by two.
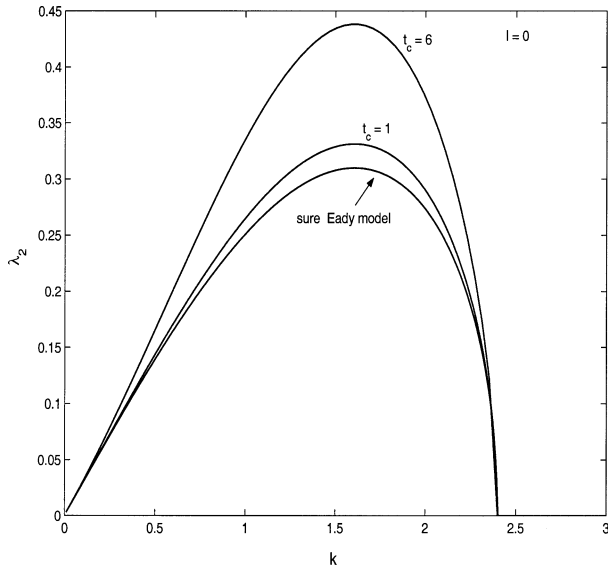
FIG. 2. Second-moment exponent as a function of wavenumber in the Eady model with shear fluctuation amplitude $\epsilon = 1/3$ and autocorrelation times $t_c = 6$ and $t_c = 1$. For comparison the bottom curve is the exponent in the absence of fluctuations. The meridional wavenumber is $l = 0$.

$$\lambda_2 = \Re(\lambda_\alpha) + \frac{2\epsilon^2}{\nu}[\Re(\mu_\alpha)]^2. \qquad (51)$$

This demonstrates the generally valid property that the eigenvalue of L with greatest real part is itself real, and that for commuting $\mathbf{A}$ and $\mathbf{B}$ a stable mean operator $\mathbf{A}$ with $\Re(\lambda_\alpha) < 0$ becomes second-moment unstable for

$$\frac{\epsilon^2}{\nu} \geq -\frac{\Re(\lambda_\alpha)}{2[\Re(\mu_\alpha)]^2}. \qquad (52)$$

This shows that any stable mean operator $\mathbf{A}$ is made second-moment unstable by sufficiently large fluctuations with structure matrix $\mathbf{B}$ that commutes with $\mathbf{A}$, requiring only that $\mathbf{B}$ have nonvanishing real eigenvalue. The structure most second-moment unstable is not necessarily the most unstable eigenfunction of $\mathbf{A}$ but rather the eigenfunction of $\mathbf{A}$ for which $\Re(\lambda_\alpha) + 2\epsilon^2/\nu[\Re(\mu_\alpha)]^2$ is maximized.

As an example consider the second-moment stability of the inviscid Eady model with modulations of the shear described in Part I. The second-moment exponent rate as a function of zonal wavenumber, k, for the case of $l = 0$, is shown in Fig. 2 for Gaussian shear fluctuations of rms magnitude $\epsilon = 1/3$ and autocorrelation times $t_c = 1$ and $t_c = 6$. For comparison the growth rate of perturbations on the mean wind profile is also shown. The increase of the second-moment exponent occurs only in the unstable region of the operator. This is because for waves shorter than the short-wave cutoff the perturbation operator has an imaginary spectrum and from (50) commuting fluctuations cannot increase the growth rate of the certain operator. In the unstable region

the fluctuating shear varies the growth rate, which produces an asymptotic second-moment exponent exceeding that of the mean operator. Because this is a commuting example, the growth rate of the second moment is the same for all autocorrelation times as that predicted by the equivalent white noise operator. The growth rates increase with autocorrelation time in accord with the increase of the equivalent white noise variance $\epsilon^2 t_c$.

When $\mathbf{A}$ and $\mathbf{B}$ do not commute the second-moment growth rates cannot be directly related to the growth rates of the constituent operators $\mathbf{A}$ and $\mathbf{B}$, and the asymptotic stability of the Lyapunov superoperator depends on the autocorrelation time of the fluctuations [cf. (33)]. Analysis of such cases requires explicit calculations of the stability of the generalized superoperator but in realistic problems this calculation is intractable because the Lyapunov superoperator for a system of dimension n has dimension n^4. We seek a method to efficiently estimate the critical fluctuation magnitude necessary to produce second-moment instability as well as the structure of the marginally stable covariance. Such a method is described below and demonstrated for the short autocorrelation form of the superoperator. With small adjustments it can be adapted to calculate the stability of the superoperator for any autocorrelation time provided that $\lim_{t\to\infty} \mathbf{D}(t) = \mathbf{D}_\infty$ exists so that the superoperator becomes asymptotically autonomous.

Denote by L_0 the linear superoperator associated with the certain operator $\mathbf{A}$,

$$L_0\mathbf{C} = \mathbf{AC} + \mathbf{CA}^\dagger, \qquad (53)$$

and assume that $\mathbf{A}$ is asymptotically stable. Denote by L_1 the linear superoperator associated with the uncertain fluctuation operator $\mathbf{B}$,

$$L_1\mathbf{C} = \sum_i \left[\frac{1}{2}\mathbf{B}_i^2\mathbf{C} + \frac{1}{2}\mathbf{CB}_i^{2\dagger} + \mathbf{B}_i\mathbf{CB}_i^\dagger\right], \qquad (54)$$

and compose from L_0 and L_1 the generalized Lyapunov superoperator

$$L_w = L_0 + \alpha^2 L_1, \qquad (55)$$

where $\alpha^2 \equiv 2\epsilon^2/\nu$. Because the eigenvalue with maximum real part of (55) is by necessity real, the generalized Lyapunov superoperator L at marginal stability must have a zero eigenvalue by the principle of exchange of stability. Exploiting this fact the noise level, α_c^2, necessary for marginal stability can be determined by eigenalysis of the superoperator $L_u = -L_0^{-1}L_1$. For if λ_{max} is the maximum eigenvalue of L_u (necessarily real and positive), then the critical α_c is given by $\alpha_c^2 = 1/\lambda_{max}$. This follows because at marginal stability the eigenvalue with maximum real part of the superoperator,

$$L_w = \alpha^2 L_0\left(\frac{I}{\alpha^2} - L_u\right), \qquad (56)$$

is zero, and consequently at marginal stability the ei-

genvalue of L_u with maximum real part is $\lambda_{\max} = 1/\alpha^2$. The eigenvalues, λ, of L_u and its eigenfunctions, $\mathbf{C}$, can be determined by solving the equation

$$\lambda(\mathbf{AC} + \mathbf{CA}^\dagger)$$

$$= -\sum_i \left[\frac{1}{2} \mathbf{B}_i^2 \mathbf{C} + \frac{1}{2} \mathbf{CB}_i^{2\dagger} + \mathbf{B}_i \mathbf{CB}_i^\dagger \right]. \quad (57)$$

The eigenfunctions of this equation consist of the positive definite covariances (the positivity is assured by the stability of $\mathbf{A}$) with the property that the stochastic forcing of the mean operator [proportional to the lhs of Eq. (57)] required for statistical equilibrium is equal to the stochastic forcing produced by the equilibrium covariance under the influence of disturbance by fluctuations with structure $\mathbf{B}$ [given by the rhs of Eq. (57)]. The marginally unstable eigenfunction can be found by solving this eigenproblem (it can be readily solved by the power method) with the advantage that this calculation does not require knowing the critical value of the fluctuation magnitude, $\epsilon = (\nu/2\lambda_{\max})^{1/2}$.

e. Solving the generalized Lyapunov equation for the statistically steady covariance in an uncertain system

For fluctuation levels, ϵ, for which the uncertain Lyapunov operator is stable we can determine the stationary response covariance, $\mathbf{C}$, maintained by additive stochastic forcing with forcing covariance $\mathbf{Q}$. This stationary response covariance satisfies the generalized Lyapunov equation (cf. 41):

$$\left(\mathbf{A} + \frac{\epsilon^2}{\nu} \sum_i \mathbf{B}_i^2 \right) \mathbf{C} + \mathbf{C} \left(\mathbf{A} + \frac{\epsilon^2}{\nu} \sum_i \mathbf{B}_i^2 \right)^\dagger$$

$$+ \frac{2\epsilon^2}{\nu} \sum_i \mathbf{B}_i \mathbf{CB}_i^\dagger + \mathbf{Q} = 0. \quad (58)$$

This equation can be written using the definitions of L_0 and L_1 introduced earlier as

$$L_0 \mathbf{C} + \alpha^2 L_1 \mathbf{C} + \mathbf{Q} = 0. \quad (59)$$

We proceed to solve (59) by expanding $\mathbf{C}$ in the perturbation series:

$$\mathbf{C} = \mathbf{C}_0 + \alpha^2 \mathbf{C}_1 + \alpha^4 \mathbf{C}_2 + \cdots. \quad (60)$$

The terms in this series can be evaluated recursively by solving the Lyapunov equations:

$$L_0 \mathbf{C}_0 + \mathbf{Q} = 0, \qquad L_0 \mathbf{C}_1 + L_1 \mathbf{C}_0 = 0, \qquad \ldots,$$

$$L_0 \mathbf{C}_n + L_1 \mathbf{C}_{n-1} = 0. \quad (61)$$

The solution of each of these Lyapunov equations is assured to yield physically realizable covariances, that is, hermitian and positive definite, because $\mathbf{A}$ is assumed stable. Therefore, the stationary covariance obtained by summing the series, if the series converges, is also hermitian and positive definite. This stationary covariance is
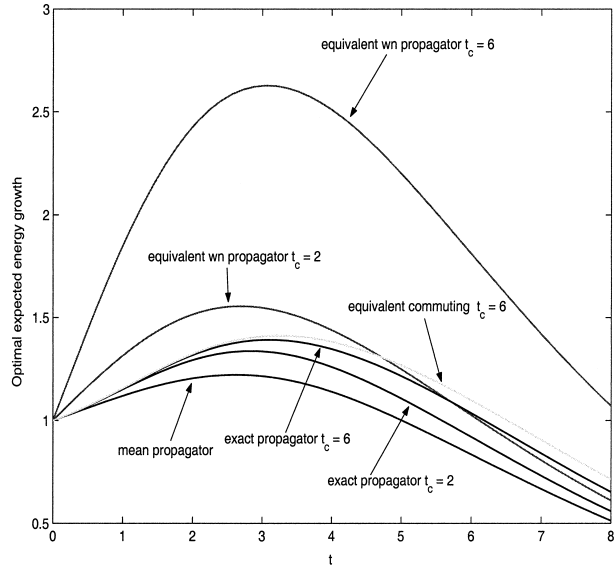

FIG. 3. Expected optimal energy growth as a function of optimizing time for the uncertain Eady model with velocity fluctuations of amplitude $\epsilon = 1/3$ and autocorrelation times $t_c = 2$ and $t_c = 6$. Shown are optimal growth given by the equivalent white noise propagator and by the exact propagator. For comparison the optimal growth from the mean propagator with no fluctuations is also shown. The fluctuations generally increase perturbation growth and the equivalent white noise approximation overestimates the growth. The fluctuating wind has vertical structure $u(z) = z^2$. The wavenumbers are $k = 3$, $l = 3$. The coefficient of linear friction is $r = 0.3$.

$$\mathbf{C} = (I + \alpha^2 L_u + \alpha^4 L_u^2 + \alpha^6 L_u^3 + \cdots)\mathbf{C}_0$$

$$= (I - \alpha^2 L_u)^{-1}\mathbf{C}_0, \quad (62)$$

where, as previously, $L_u = -L_0^{-1} L_1$. The series is assured to converge if the variance of the fluctuations α satisfies $\alpha^2 \lambda_{\max} < 1$, where $\lambda_{\max}$ is the maximal eigenvalue of L_u, which is real and positive because this system obeys the principle of exchange of stability. This expansion confirms that the critical fluctuation magnitude leading to marginal stability is $\alpha_c^2 = 1/\lambda_{\max}$ and that for $\alpha < \alpha_c$ a stationary state can be obtained under stochastic forcing.

Although the analysis above exploited the assumption of delta-correlated fluctuations, the methods presented can be used to obtain similar results for more general forms of fluctuation.

4. Optimal perturbations and stochastic optimal forcing for uncertain systems

In order to obtain optimal perturbations under uncertain dynamics we must first specify a norm to measure the covariances. A natural norm is the variance

$$\|\mathbf{C}\| = \text{trace}(\mathbf{C}). \quad (63)$$

While the trace of a matrix does not define a norm for general matrices, it does define a norm for physically realizable positive definite hermitian matrices for which
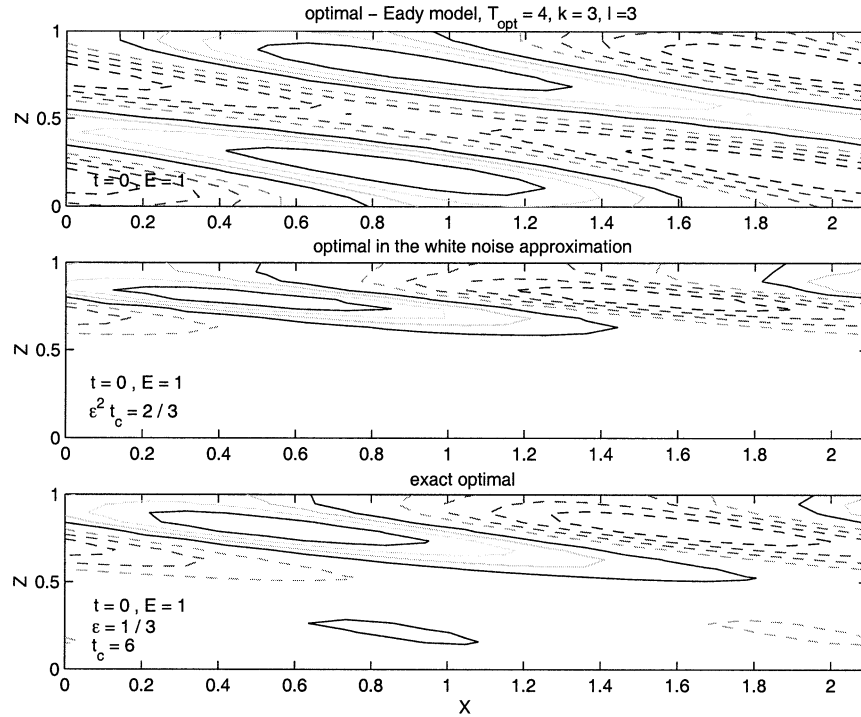
FIG. 4. Structure of the optimal perturbation that leads to greatest expected energy at $t = 4$ in the uncertain Eady model. The amplitude of the fluctuations is $\epsilon = 1/3$ and the autocorrelation time is $t_c = 6$; other parameters are as in Fig. 3: (top) the optimal of the mean operator, which produces energy growth 1.14; (middle) the optimal of the equivalent white noise dynamics, which produces energy growth 3.9; (bottom) the optimal of the exact dynamics, which produces energy growth 1.4.

it is always true that trace($\mathbf{C}$) ≥ 0 with equality only when $\mathbf{C}$ is identically zero. This norm is chosen in part because for covariances produced by pure states it is identical with the square of the Euclidean norm of the state, which in energy coordinates is the perturbation energy. Because any sum of positive definite covariances is also positive definite, this norm is linear when it acts on sums of covariances, that is, $\|\mathbf{C} + \mathbf{D}\| = \|\mathbf{C}\| + \|\mathbf{D}\|$, while for covariance differences this norm satisfies the triangle inequality, $\|\mathbf{C} - \mathbf{D}\| \leq \|\mathbf{C}\| + \|\mathbf{D}\|$.

Recall that an n^2 vector $\mathbf{c}$ corresponds to the $n \times n$ covariance matrix $\mathbf{C}$. The trace norm on $\mathbf{C}$ is evaluated on the vector covariance c by setting $\|\mathbf{c}\| = T\mathbf{c}$, where the trace superoperator T is defined as the $1 \times n^2$ matrix $T_{1i} = \delta_{i,(k-1)n+k}$ for $k = 1, \ldots, n$. Clearly, the trace norm of the vector c is distinct from the Euclidean norm. It is important to note that this linear norm on the covariances is not related to an inner product, and as a result orthogonality between covariances is not defined.

A norm that defines an inner product on the covariances with orthogonality properties is the Frobenius norm. The Euclidean norm on vector covariances, $\|\mathbf{c}\|_F = \sqrt{\mathbf{c}^\dagger \mathbf{c}}$, corresponds to the Frobenius norm on covariance matrices,

$$\|\mathbf{C}\|_F = \sqrt{\text{trace}(\mathbf{C}^\dagger \mathbf{C})}. \tag{64}$$

For covariance matrices of rank one, the associated vec-

tor covariance is $\mathbf{c} = \boldsymbol{\psi}^* \otimes \boldsymbol{\psi}$ for a pure state $\boldsymbol{\psi}$, and in such cases the Frobenius norm of $\mathbf{C}$ is equal to the trace norm of $\mathbf{C}$, and the dot product between two rank-one covariance matrices is equal to the square of the magnitude of the dot product of the corresponding state vectors.

In the previous sections we have determined the superoperator that advances an initial covariance $\mathbf{c}(0)$ to the covariance t units of time later $[\mathbf{c}(t) = P(t)\mathbf{c}(0)]$ and also the superoperator that produces the steady-state covariance $\mathbf{c}(\infty)$ from the stochastic forcing covariance $q(\mathbf{c}(\infty)) = \{\lim_{t\to\infty} \int_0^t P(t, s) \, ds\}q)$, whenever this is defined. In all cases we have a mapping from an initial covariance c^i to a final covariance $\mathbf{c}^f (\mathbf{c}^f = M\mathbf{c}^i)$, and the optimization problem is to determine the initial c^i of unit magnitude that produces the $\mathbf{c}^f$ of largest magnitude. If M is the propagator associated with the Lyapunov superoperator, we interpret the optimal as the initial covariance $\mathbf{c}^i_{\text{opt}}$ that produces the largest final covariance $\mathbf{c}^f$. If M is the mapping from forcing structure matrix q to the equilibrium-maintained covariance $\mathbf{c}(\infty)$, then the optimal q_{opt} identifies the forcing structure matrix producing the largest maintained variance.

If covariances are measured in the Frobenius norm, $\|\mathbf{C}\|_F$, then the optimization problem is solved by sin-
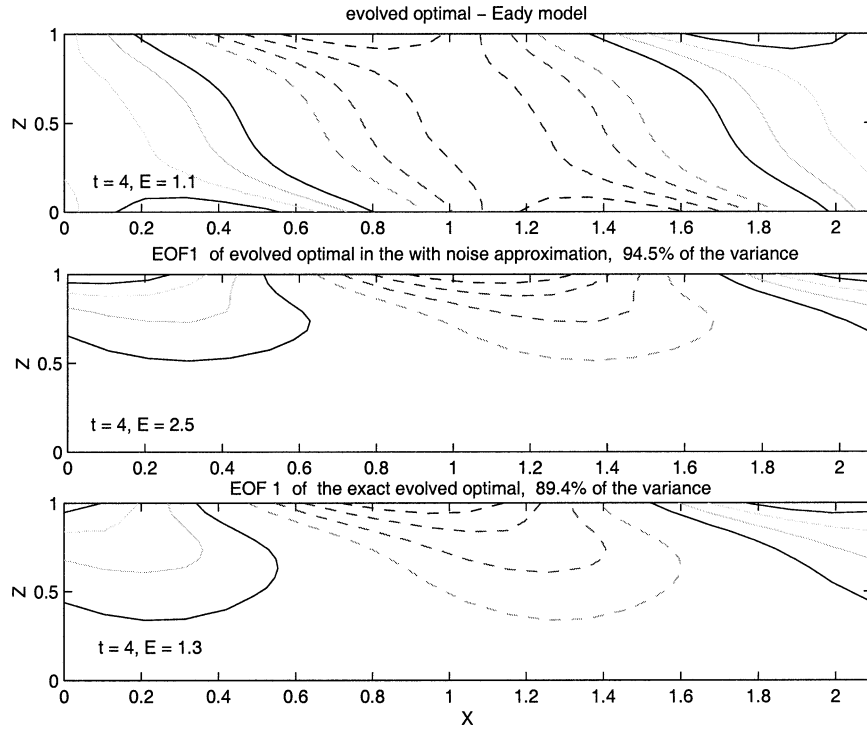
Fig. 5. Structure of the first EOF evolved from the sure optimal initial state shown in Fig. 4: (top) the sure evolved optimal of the mean dynamics; (middle) the first EOF of the covariance obtained using the equivalent white noise dynamics; the first EOF accounts for 96% of the evolved optimal covariance; (bottom) the first EOF of the covariance obtained using the exact dynamics. Note that despite the exaggerated growth factor obtained using the equivalent white noise dynamics, the structure of the evolved covariance is well approximated.

gular value decomposition of the mapping from initial to final covariances in which the mapping is factored as $M = \mathbf{U}\Sigma\mathbf{V}^{\dagger}$. If the diagonal elements of Σ, σ_i, are ordered in descending magnitude, then the first column of $\mathbf{V}$ is the optimal covariance $\mathbf{c}_{\text{opt}}^{i}$ (or q_{opt} for the stochastic optimal) in the Frobenius norm and the optimal amplification is the first singular value:

$$\sigma_1 \equiv \|M\|_F = \max\frac{\|M\mathbf{c}^i\|_F}{\|\mathbf{c}^i\|_F}. \tag{65}$$

Note that this optimal covariance is not necessarily rank one so that the covariance need not be associated with a pure state and the optimal growth need not be realizable by a single integration of the model. Such a mixed optimal covariance can be physically prepared only by separate initializations of the system with the pure states that give rise to this covariance and subsequently forming the ensemble covariance of these simulations. Although obtaining the optimal covariance in the Frobenius norm is straightforward, operationally this norm is less physically relevant as a measure of the covariance than is the trace because it maximizes the square sum of all the elements of the covariance matrix, which is not commonly associated with a physical quantity, while the trace of the covariance matrix is the square sum of

the diagonal elements and can be identified with the physical quantity variance.

It is useful to consider SVD of the composite operator formed by the mapping M and the trace superoperator T (TM, rather than of the mapping M alone). Because TM projects n^2-dimensional covariances to the reals it has only one nonzero singular value; we denote this nonzero singular value by σ, and the associated singular vector by v, so that $TM = \sigma v^{\dagger}$. This SVD reveals that the action of TM is just taking the inner product with v. The covariance matrix corresponding to v, $\mathbf{V}$, almost succeeds in identifying the optimal initial covariance, as it is among initial covariances with unit Frobenius norm ($\|\mathbf{C}^i\|_F = 1$) the one producing maximum final variance max ($\|\mathbf{C}^f\|$). However, it does not succeed in identifying the positive definite hermitian covariance of unit variance ($\|\mathbf{C}^i\| = 1$) leading to maximum final variance. Remarkably, this optimal covariance is the rank-one covariance formed from the first EOF of $\mathbf{V}$ and the growth is $\sigma\alpha_1$ with a_1 the first singular value of $\mathbf{V}$. This result, which does not obtain unless $\mathbf{V}$ is both hermitian and positive definite, is proven in appendix B. In appendix C an alternative proof by construction of the optimal covariance is given that has the advantage of being suitable for computation of the optimals in large systems.
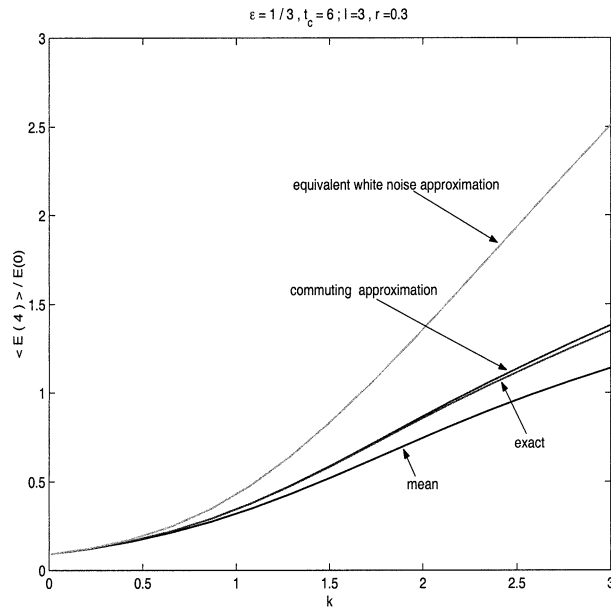
FIG. 6. Optimal expected energy growth as a function of wavenumber, k, for the uncertain Eady model with fluctuation amplitude $\epsilon = 1/3$ and autocorrelation time $t_c = 6$. Shown are the optimal growth at $t = 4$ obtained from the equivalent white noise propagator, the exact propagator, the propagator of Eq. (66), which would have been exact if $\mathbf{A}$ and $\mathbf{B}$ commuted, and the optimal growth from the mean propagator. The fluctuating wind has vertical structure $u(z) = z^2$. The wavenumbers are $k = 3$, $l = 3$. The coefficient of linear friction is $r = 0.3$.

In certain dynamics optimal perturbations can be ordered according to growth, or stochastic optimals according to contribution to variance, but such an ordering cannot be performed with covariances, as orthogonality is not defined for the trace norm because the trace norm is not associated with an inner product. However, we can form an orthogonal set of optimal perturbations arising from pure state covariances and in this way identify a set of optimal perturbations for uncertain dynamics. The second pure state initial condition, orthogonal to the first, that produces the second largest growth is the second EOF of $\mathbf{V}$ and its growth is $\sigma\alpha_2$. Proceeding in this way we obtain an orthonormal basis ordered according to variance amplification, analogous to that obtained for deterministic propagators by SVD of the deterministic propagator. However, note that we do not obtain in this way a second optimal covariance matrix, but only the second optimal pure initial state, because as previously mentioned there is no meaning to a second optimal covariance as the set of positive definite hermitian matrices do not form an inner product space in the trace norm.

Consider the statistically steady ensemble average covariance of a stochastically forced system, which, as discussed earlier, is equal under the assumption of ergodicity to the time average covariance of a single realization of the system as would be accumulated by a time-averaging instrument. We have seen that the struc-

ture of stochastic forcing producing the largest maintained variance is a pure state given by the first EOF of TM with $M = \lim_{t\to\infty} \int_0^t P(t, s)\, ds$.

Consider the problem of determining ensemble forecast perturbations when the tangent linear forecast error system is uncertain, perhaps because of parameter uncertainty (Palmer 2001). One method for choosing ensemble members is to take optimal perturbations (Molteni et al. 1996; Ehrendorfer and Tribbia 1997). In the presence of uncertainty we have shown that the optimal producing greatest expected growth is a covariance of rank one. This implies that there is a single optimal perturbation for the uncertain forecast system that produces the greatest expected error growth; subsequent pure state optimals could be chosen to complement the basis set of optimal perturbations that compose the ensemble.

5. Second-moment growth and stochastically maintained variance: Optimals and stochastic optimals in the uncertain Eady problem

Consider the Eady problem presented in Part I, but for simplicity with wind fluctuations confined to the form $u(z) = z^2$ and rms amplitude $\epsilon = 1/3$. Take autocorrelation times $t_c = 2$ and $t_c = 6$ in order to assess the role of finite correlations and the range of validity of the equivalent white noise approximation (36) to the exact uncertain Lyapunov superoperator (35). For simplicity take Rayleigh damping with damping timescale chosen to stabilize the model so that it reaches a statistically steady state under stochastic forcing.

Consider first determining the perturbation producing the largest expected variance (trace) of $\mathbf{C}(t)$ over all matrices $\mathbf{C}(0)$ with unit initial variance. The map that connects $\mathbf{C}(t)$ and $\mathbf{C}(0)$ is the uncertain propagator $P(t)$ and the optimal initial covariance matrix we seek is the rank-one initial covariance formed by the first EOF of the covariance associated with right singular vector, $\mathbf{V}$, of $TP(t)$, where T is the trace superoperator. The variance growth over t achieved by this optimal perturbation has magnitude $\sigma\alpha_1$, where σ is the singular value of $TP(t)$ and α_1 is the largest eigenvalue of the covariance associated with the singular vector $\mathbf{V}$. The remaining set of mutually orthogonal optimal perturbations each produces variance growth of $\sigma\alpha_i$, where α_i are the remaining eigenvalues of $\mathbf{V}$. While the optimal initial covariance matrix is of rank one, at later times the evolved optimal covariance matrix becomes mixed and its structure can be described using EOF analysis.

Optimal variance growth as a function of optimizing time for the Eady model is shown in Fig. 3. The optimal growth obtained using the exact covariance dynamics [Eq. (37)] is compared with that obtained using the equivalent white noise dynamics [Eq. (38)] that is formally valid for sufficiently short autocorrelation times, t_c. A further comparison with the optimal growth attained by the mean zonal flow Eady operator in the
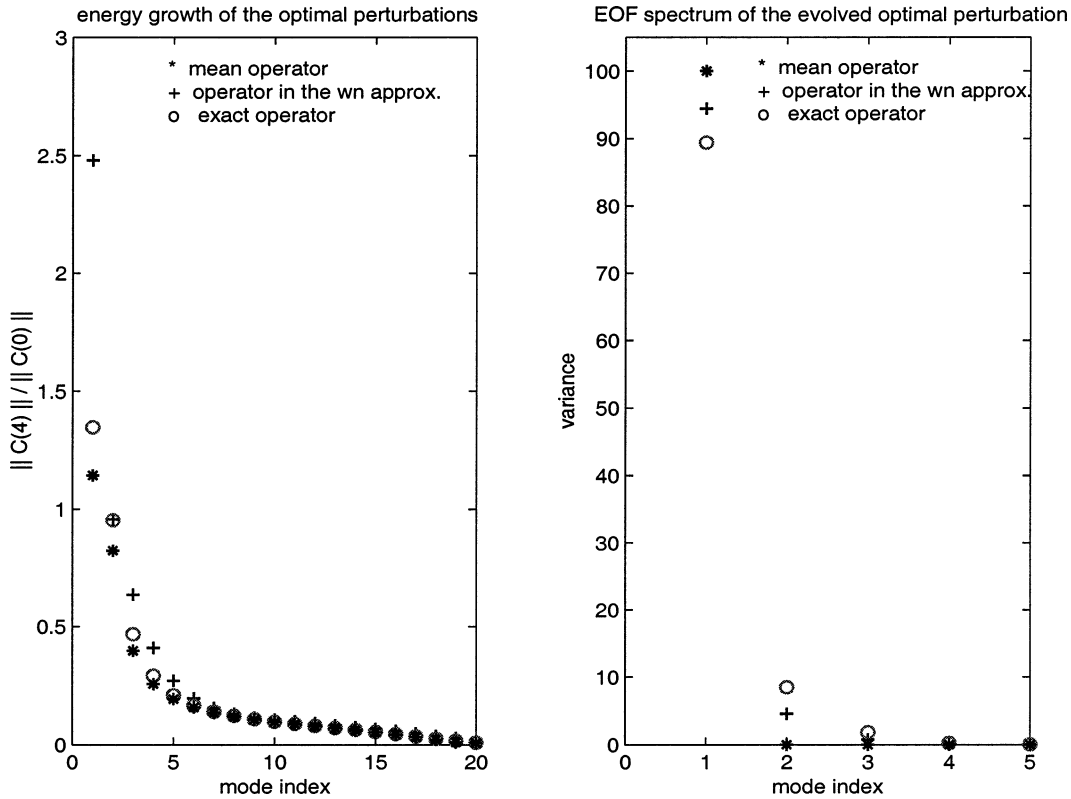
FIG. 7. (left) Expected energy growth achieved by the optimal perturbations in four time units. Shown are the growth achieved by optimal perturbations obtained using the exact ensemble mean square dynamics (circles), the growth achieved by the optimal perturbations obtained using the equivalent white noise dynamics (crosses), and the growth achieved by the optimals of the mean dynamics in the absence of fluctuations (stars). (right) EOF decomposition of the covariance at $t = 4$ arising from evolution of the rank-one initial covariance of the top optimal perturbation. Shown are the variance percentage accounted for by the EOFs of the covariance evolved by the exact dynamics (circles), the variance accounted for by the EOFs of the equivalent white noise dynamics (crosses), and the variance accounted for by the EOFs of the mean dynamics (stars). The covariance evolved with the mean dynamics remains rank one and a single EOF accounts for 100% of its variance. The evolved covariance under uncertain dynamics is mixed and spanned by approximately two states. The amplitude of the fluctuations is $\epsilon = 1/3$ and the autocorrelation time is $t_c = 10$; the model and the other parameters are as in Fig. 3.

absence of fluctuations is also shown. Note that fluctuations increase the expected variance growth. While the equivalent white noise approximation overestimates the growth potential, it identifies the correct structure both for the optimal perturbation and the evolved optimal covariance. This is shown for the $t = 4$ optimal and the first EOF of the evolved optimal covariance at the optimizing time in Figs. 4 and 5, respectively. Note that optimals in the fluctuating Eady model are concentrated near the upper boundary where fluctuations of the shear are largest.

The inaccuracy in the optimal growth estimate made by the white noise approximation is not due primarily to overestimation of the fluctuation operator magnitude, $\epsilon^2 B^2/\nu$, but rather to assuming that this value is attained from the start. We know from the discussion of the ensemble mean dynamics [cf. Eq. (3)] that the ensemble mean correction requires a time of order t_c to attain its asymptotic form. A more accurate approximation of optimal growth for short times is obtained if the ensemble

mean fluctuation correction is allowed to build up over time according to

$$\frac{dc}{dt} = \left\{ A + \frac{\epsilon^2}{\nu}(1 - \mathbf{e}^{-\nu t})B^2 \right\} c, \qquad (66)$$

which would be exact if A and B commuted. Remarkably, the optimal growth obtained from the propagator associated with the equivalently commuting (66) accurately tracks the optimal growth obtained from the propagator of exact superoperator (35) (cf. Fig. 3). The accuracy of the growth calculated from (66) can also be seen in Fig. 6, which shows the expected optimal energy growth attained in four time units as a function of wavenumber.

While in an uncertain system we cannot determine a set of orthogonal covariances ordered according to growth potential, we can determine a set of orthogonal optimal perturbations ordered according to growth potential. The growth resulting from these optimal per-
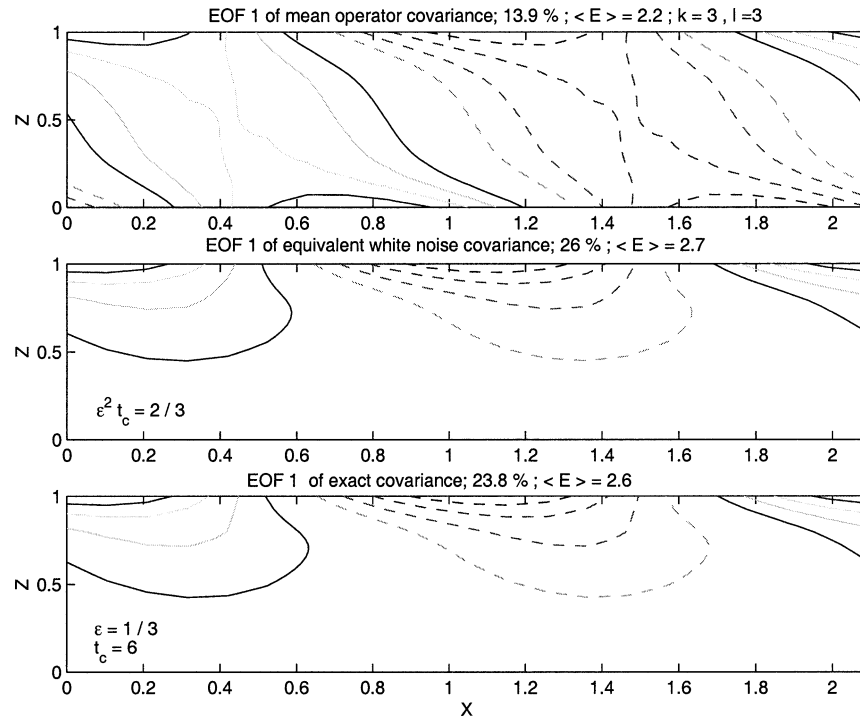
FIG. 8. Structure of the first EOF of maintained variance in the Eady model. (top) The first EOF produced by temporally and spatially white noise additive forcing of the mean operator; the maintained energy is 2.2, and the first EOF accounts for 13.9% of the variance. (middle) The first EOF produced in the equivalent white noise approximation by temporally and spatially white noise forcing of the operator associated wind fluctuating about the mean; the fluctuating wind is $u(z) = z^2$, the rms amplitude of the fluctuations is $\epsilon = 1/3$, and the autocorrelation time is $t_c = 6$; the maintained energy is 2.7, and the first EOF accounts for 26% of the variance. (bottom) The exact first EOF for the operator associated with the same fluctuating wind as in the middle panel but with the assumption that the fluctuations are Gaussian with Kubo number $K = 2$; the maintained energy is 2.6, and the first EOF accounts for 23.8% of the variance; the wavenumbers are $k = 3$, $l = 3$; the coefficient of linear friction is $r = 0.3$.
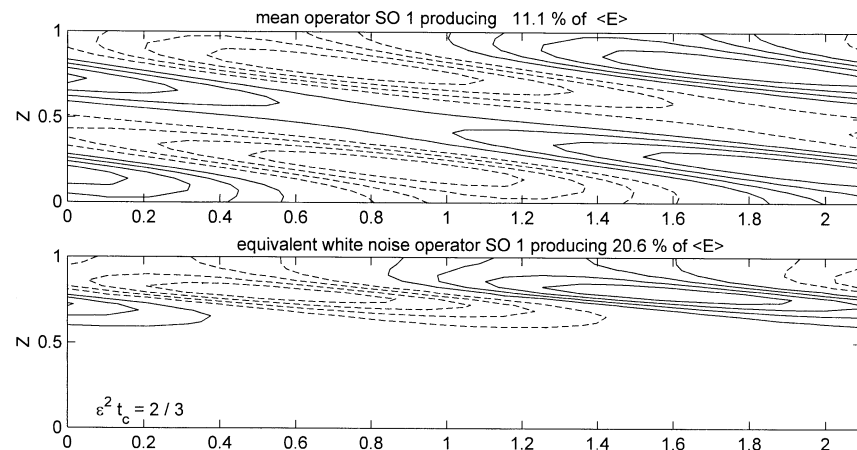

FIG. 9. (top) Structure of the first stochastic optimal, which is responsible for producing 11.1% of the total variance when the mean operator of the Eady model is stochastically forced with temporally white additive noise with the spatial structure of the stochastic optimal. (bottom) The structure of the first stochastic optimal in the fluctuating Eady model in the equivalent white noise approximation; the fluctuating wind is $u(z) = z^2$, the rms amplitude of the fluctuations is $\epsilon = 1/3$, and the autocorrelation time is $t_c = 6$; this stochastic optimal is responsible for producing 20.6% of the total variance. The wavenumbers are $k = 3$, $l = 3$; the coefficient of linear friction is $r = 0.3$.
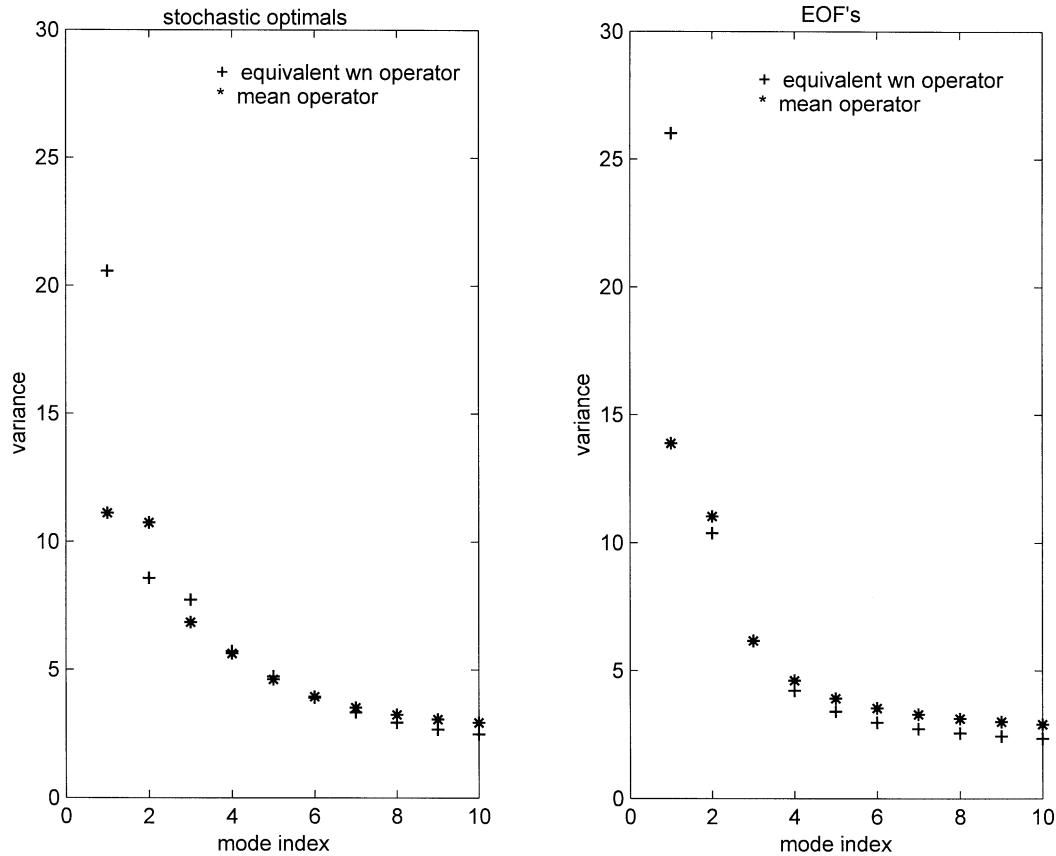
FIG. 10. (left) Variance maintained by the first 10 stochastic optimals of the uncertain Eady model: the variance maintained by the equivalent white noise approximation (crosses), and for reference the variance maintained by the mean operator without fluctuations (stars). (right) The percentage of the variance of the uncertain Eady model arising from the first 10 EOFs with stochastic forcing white in space and time: variance explained by the equivalent white noise approximation (crosses), variance explained by the mean operator (stars); the fluctuating wind is $u(z) = z^2$, the rms amplitude of the fluctuations is $\epsilon = 1/3$ and the autocorrelation time is $t_c = 10$; the wavenumbers are $k = 3$, $l = 3$; the coefficient of linear friction is $r = 0.1$.

turbations is shown for the $t = 4$ optimization in the left panel of Fig. 7.

While the optimal covariance has rank one, and therefore can be identified within a phase with a pure optimal perturbation, the evolved covariance is mixed by the uncertain dynamics. The rank of the covariance is approximately equal to the number of EOFs of the covariance with appreciable variance. In the right panel of Fig. 7 the EOF spectrum of the evolved optimal covariance for the $t = 4$ optimal is shown at the optimizing time. Under uncertain dynamics the sure initial covariance becomes mixed with approximately rank two.

We turn now to the steady-state variance maintained in the uncertain Eady problem under spatially and temporally white forcing. Although in this example operator fluctuations only marginally increase the variance, the variance structure shifts markedly to the upper level (Fig. 8). The origin of upper-level short waves in the atmosphere lacks persuasive theoretical explanation and we are pursuing this mechanism of stochastic destabilization.

We seek the optimal forcing covariance matrix with the property that when the associated structure is forced white in time the maximum variance is maintained; this structure is called the stochastic optimal. Finding the stochastic optimal requires, according to (40), obtaining the optimal for the map $M = -L_w^{-1}$, where L_w is the asymptotic superoperator of the covariance dynamics for short autocorrelation times. The structure of the first stochastic optimal for the exact operator and for the equivalent white noise approximation to this operator is shown in Fig. 9. The stochastic optimals in this uncertain Eady model have largest amplitude at upper levels where shear fluctuations are largest. The variance maintained by the stochastic optimals is shown in Fig. 10 (left panel), and the decomposition of the maintained covariance matrix into its EOF components is shown Fig. 10 (right panel).

While in this example the stochastic optimals were obtained by singular value decomposition of L^{-1}, this is computationally expensive and there is an alternative method for finding these optimals by eigenanalysis of

the generalized back Lyapunov equation in direct analogy with the method of solution used for certain dynamics. This method and the back Lyapunov equation for uncertain dynamics are described in appendix C.

6. Discussion and conclusions

In Part I of this paper we examined ensemble mean stability of uncertain systems from the perspective of generalized stability theory (GST) and in Part II we addressed second-moment stability properties of uncertain systems from this perspective. We first expressed the equations governing covariance dynamics of certain operators in tensor product notation and then extended these to uncertain operators. We obtained stability boundaries for second-moment growth in uncertain systems in terms of the amplitude and structure of operator fluctuations. A physical implication of these results is that for a stochastically forced linear system to have finite maintained variance its parameter fluctuations must lie within these bounds.

Analysis of the stochastically forced uncertain system allows extension of previous results on the variance, fluxes, and structures in statistically steady turbulent flows (cf. Farrell and Ioannou 1994, 1995, 1998) to take account of uncertainty in the operator.

The time mean covariance of a stochastically forced uncertain system can be obtained from second-moment analysis if an ergodic assumption connecting the ensemble and time means of stable forced systems is made. From this perspective second-moment ensemble mean stability can be interpreted as a necessary condition for the existence of a bounded forced state in an uncertain system. This is because second-moment unstable systems, even if sample stable, support unbounded excursions in variance that prevent establishing a finite forced variance regime.

Optimal perturbations play a major role in GST and for uncertain systems we obtain a method for finding the initial condition leading to greatest expected variance growth at any chosen time as well as the stochastic forcing structures that maintain the greatest variance. Remarkably, one pure initial condition is found constructively that maximizes expected growth in an uncertain system and an analogous pure forcing structure is found that maximizes the expected variance.

One application of this work is a method for taking account of forecast error system uncertainty in choosing forecast ensembles by substituting uncertain optimals for certain optimals in the ensemble. Another application is to the problem of explaining the origin of upper-level short waves in the midlatitude jet, which are dynamically similar to the upper-level structures in the fluctuating jet example (Fig. 8).

APPENDIX A

Basic Properties of the Kronecker (Tensor) Product

The tensor or Kronecker product of the $k \times l$ matrix $\mathbf{A}$ with the $m \times n$ matrix $\mathbf{B}$ is the $km \times ln$ matrix $\mathbf{A} \otimes \mathbf{B}$ defined as follows:

$$\mathbf{A} \otimes \mathbf{B} \equiv \begin{Bmatrix} A_{11}B & \cdots & A_{1l}B \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ A_{k1}B & \cdots & A_{kl}B \end{Bmatrix}. \quad (A1)$$

Two properties of the tensor product will be used repeatedly:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}) \quad \text{and} \quad (A2)$$

$$(\mathbf{A} \otimes \mathbf{B})^{\dagger} = \mathbf{A}^{\dagger} \otimes \mathbf{B}^{\dagger}. \quad (A3)$$

APPENDIX B

Construction of the Optimal Covariance

Consider singular value decomposition of the composition of the trace superoperator T with the mapping M: $TM = \mathbf{U}\Sigma\mathbf{V}^{\dagger}$, in which $\mathbf{V}$ is a single column. We will show that the first singular vector $\mathbf{V}$ is the optimal initial covariance that results in maximum variance after mapping by M. This shows that the optimal is physically realizable and necessarily rank one.

First, we establish that $\mathbf{V}$ is hermitian. If $\mathbf{V}$ were not hermitian, it could be decomposed into its hermitian and antihermitian parts. Because the antihermitian part has zero trace under the mapping M, a nonhermitian $\mathbf{V}$ would produce the same variance as its hermitian part but would have larger Frobenius norm. Therefore, the optimal $\mathbf{V}$ is hermitian.

Second, we establish that $\mathbf{V}$ must be positive definite. Consider the eigendecomposition of $\mathbf{V}$ in the orthonormal basis of its eigenvectors e_i ordered descending in the magnitude of their real eigenvalue α_i. In vector form this eigendecomposition can be written as

$$\mathbf{v} = \sum_i \boldsymbol{\alpha}_i \mathbf{e}_i^* \otimes \mathbf{e}_i; \quad (B1)$$

if $\mathbf{V}$ were not positive definite, some of the $\boldsymbol{\alpha}_i$ would be negative, and v could then be written as the difference of two positive definite matrices $\mathbf{V} = \mathbf{V}_+ - \mathbf{V}_-$, by partitioning the summation over the positive and negative eigenvalues. The action of the linear dynamics does not affect this partitioning, because each positive definite hermitian matrix is mapped to a positive definite hermitian matrix. But then we reach a contradiction: $\mathbf{V}$ is the covariance that produces the largest growth; if it were not positive definite, then a higher energy would be attained by either $\mathbf{V}_+$ or $\mathbf{V}_-$ (because the diagonal

elements of a positive definite matrix are by necessity positive). We conclude that the optimal initial covariance that maximizes the final covariance is by necessity hermitian and positive definite.

Third, we show that the optimal covariance is necessarily a rank-one covariance, that is, a covariance produced by a sure initial condition, and that the optimal is the leading EOF of **V**. Moreover, the EOFs can be shown to order all pure state covariances according to the variance that each produces in the final state.

That the optimal must be a pure state is clear from an argument based on superposition.[B1] A general unit variance realizable initial covariance $\mathbf{C}^i(\|\mathbf{C}^i\| = 1)$ can be decomposed into its orthonormal EOFs, f_i, in the vector form $\mathbf{c}^i = \Sigma_k \gamma_k \mathbf{c}_k$, where $\mathbf{c}_k$, is the rank-one vector covariance associated with each EOF, that is, $\mathbf{c}_k = f_k^* \otimes f_k$. Because trace($\mathbf{C}_k$) = 1, it follows that $\Sigma_k \gamma_k = 1$. The final covariance $\mathbf{c}^f$ will be $\mathbf{c}^f = \Sigma_k \gamma_k M \mathbf{c}_k$ and its total variance will be the linear sum of the corresponding trace ($M\mathbf{c}_k$). Out of the covariances formed by the EOFs, let $\mathbf{c}_\alpha = f_\alpha^* \otimes f_\alpha$ be the initial covariance of unit variance that leads to the largest final variance, trace ($M\mathbf{c}_\alpha$); then because $\Sigma_i \gamma_i = 1$, a higher growth is obtained by $\mathbf{c}_\alpha = f_\alpha^* \otimes f_\alpha$ rather than $\mathbf{C}^i$ itself.

Having established that the optimal initial condition is rank one, consider a general rank one realizable initial covariance produced by the sure state $\boldsymbol{\psi}^i = \Sigma_m \beta_m e_m$, where e_m are the orthonormal EOFs of **V**, and β_m undetermined but such that $\Sigma_m |\beta_m|^2 = 1$, so that the initial covariance,

$$\mathbf{c}^i = \boldsymbol{\psi}^{i*} \otimes \boldsymbol{\psi}^i = \sum_m \sum_l \beta_m^* \beta_l \mathbf{e}_m^* \otimes \mathbf{e}_l, \quad \text{(B2)}$$

that is generated from it has unit trace. Consider the action of TM on $\mathbf{c}^i$. Using (B1) and the identity $(\mathbf{A} \otimes \mathbf{B})^\dagger = \mathbf{A}^\dagger \otimes \mathbf{B}^\dagger$, the mapping TM can be written as

$$TM = \sigma \sum_{k=1}^n \boldsymbol{\alpha}_k \mathbf{e}_k^{*\dagger} \otimes \mathbf{e}_k^\dagger, \quad \text{(B3)}$$

because the EOF coefficients $\boldsymbol{\alpha}_k$ are real (and positive). Note that because of the orthonormality of the eigenvectors, e_k, we have for any $k \neq l$ that the vectors $e_k^* \otimes e_l$ are perpendicular to v in the Euclidean vector inner product [cf. Eq. (B1)]. We thus obtain

---

There is an alternative way to see that the optimal covariance is by necessity rank one. As discussed earlier the positive definite covariances form a convex subset, that is, any convex combination of $\mathbf{C}_1$ and $\mathbf{C}_2$, $\mathbf{C} = \lambda \mathbf{C}_1 + (1 - \lambda)\mathbf{C}_2$ with $1 \geq \lambda \geq 0$, produces a positive definite hermitian covariance. The pure states are the "vertices" of this subset in the sense that they cannot be written as convex combinations of covariances. The measure we have introduced is a linear measure on the covariances and consequently the optimum is attained at a vertex, that is, for a pure state covariance (cf. Farrell and Ioannou 2000).

$$TM\mathbf{c}^i = \sigma\left(\sum_k \boldsymbol{\alpha}_k \mathbf{e}_k^{*\dagger} \otimes \mathbf{e}_k^\dagger\right) \sum_m \sum_l \beta_m^* \beta_l \mathbf{e}_m^* \otimes \mathbf{e}_l$$

$$= \sum_k \boldsymbol{\alpha}_k |\beta_k^2| \leq \boldsymbol{\alpha}_1, \quad \text{(B4)}$$

because $\Sigma_k |\beta_k|^2 = 1$ and for all k, $\boldsymbol{\alpha}_k > 0$, and the maximum growth is attained by choosing $\beta_k = 0$ for $k > 1$ and $\beta_1 = 1$. Therefore, it is attained by the first EOF of **V** and the variance growth is $\sigma\boldsymbol{\alpha}_1$.

APPENDIX C

The Back Lyapunov Equation for Uncertain Dynamics and Construction of the Optimals

Consider the system

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}(t)\boldsymbol{\psi} + F\xi(t), \quad \text{(C1)}$$

in which the asymptotically stable certain operator $\mathbf{A}(t)$ is stochastically forced by a temporally white scalar noise process $\xi(t)$. The stochastic optimal is the forcing vector F of unit Euclidean norm that produces the greatest mean variance. First, notice that because $\boldsymbol{\psi}(t) = \int_0^t \mathbf{P}(t, s)F\xi(s)\,ds$ where $\mathbf{P}(t, s)$ is the propagator, the variance at time t is

$$\langle \boldsymbol{\psi}^\dagger \boldsymbol{\psi} \rangle = F^\dagger \left\{ \int_0^t \mathbf{P}^\dagger(t, s)\mathbf{P}(t, s)\,ds \right\} F. \quad \text{(C2)}$$

It is apparent from this expression that the eigenvector of the hermitian matrix

$$\mathbf{S}(t) = \int_0^t \mathbf{P}^\dagger(t, s)\mathbf{P}(t, s)\,ds \quad \text{(C3)}$$

with largest eigenvalue is the forcing structure that produces the greatest variance at time t. The other eigenvectors of $\mathbf{S}(t)$ complete the set of mutually orthogonal forcings that can be ordered according to the variance that each produces at time t.

The stochastic optimal is obtained by eigenanalysis of $\mathbf{S}_\infty = \lim_{t \to \infty} \mathbf{S}(t)$. This limit can be easily obtained by noting that $\mathbf{S}(t)$ satisfies the equation

$$\frac{d\mathbf{S}}{dt} = \mathbf{I} + \mathbf{A}^\dagger(t)\mathbf{S} + \mathbf{S}\mathbf{A}(t). \quad \text{(C4)}$$

If **A** is autonomous and asymptotically stable, the steady state $\mathbf{S}_\infty$ satisfies the back Lyapunov equation

$$\mathbf{A}^\dagger \mathbf{S}_\infty + \mathbf{S}_\infty \mathbf{A} = -\mathbf{I}. \quad \text{(C5)}$$

From this equation, $\mathbf{S}_\infty$ can be easily determined and the stochastic optimals obtained.

The advantage of (C5) is that it is often computationally easier to obtain the optimals of an operator rather than the optimals of a superoperator. We wish to determine the appropriate back Lyapunov equations for

an uncertain system. Because each realization satisfies the back Lyapunov equation

$$\frac{d\mathbf{S}}{dt} = \mathbf{I} + \{\mathbf{A}^{\dagger} + \epsilon\eta(t)\mathbf{B}\}\mathbf{S} + \mathbf{S}\{\mathbf{A} + \epsilon\eta(t)\mathbf{B}\}, \quad (C6)$$

we can obtain the ensemble average equation by the same method that was used to obtain the ensemble average covariance in section 3. The resulting steady-state back Lyapunov equation for the noncommuting case is in the short autocorrelation limit:

$$\left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right)^{\dagger}\mathbf{S}_{\infty} + \mathbf{S}_{\infty}\left(\mathbf{A} + \frac{\epsilon^2}{\nu}\mathbf{B}^2\right) + \frac{2\epsilon^2}{\nu}\mathbf{B}^{\dagger}\mathbf{S}_{\infty}\mathbf{B} = -\mathbf{I}.$$
$$(C7)$$

The steady state $\mathbf{S}_{\infty}$ in high-dimension systems can be obtained using (C7) through the series procedure described in section 3e. Eigenanalysis of $\mathbf{S}_{\infty}$ then provides the stochastic optimals. This development shows constructively that the stochastic optimals are rank-one covariances.

Analogously and as an alternative to the procedure used in section 4 for obtaining the optimal initial condition that leads to the greatest expected perturbation growth at any time t, we can proceed in the following manner. At time t the perturbation square amplitude for each realization of the fluctuations is

$$\boldsymbol{\psi}^{\dagger}\boldsymbol{\psi} = \boldsymbol{\psi}_0^{\dagger}\{\mathbf{P}^{\dagger}(t, 0)\mathbf{P}(t, 0)\}\boldsymbol{\psi}_0. \quad (C8)$$

It is apparent from this expression that the eigenvector of the hermitian matrix

$$\mathbf{H}(t) = \mathbf{P}^{\dagger}(t, 0)\mathbf{P}(t, 0) \quad (C9)$$

with largest eigenvalue is the initial condition that leads to the greatest amplitude at time t. The other eigenvectors of $\mathbf{H}(t)$ complete the set of mutually orthogonal initial conditions ordered according to their growth at time t. To obtain the expected optimal perturbation we first form the ensemble average of equation

$$\frac{d\mathbf{H}}{dt} = \mathbf{A}^{\dagger}(t)\mathbf{H} + \mathbf{H}\mathbf{A}(t), \quad (C10)$$

which is satisfied by each realization. For example, the ensemble mean equation in the white noise approximation is

$$\frac{d\langle\mathbf{H}\rangle}{dt} = \left\{\mathbf{I} \otimes \mathbf{A}^{\dagger} + \mathbf{A}^{\mathrm{T}} \otimes \mathbf{I}\right.$$
$$\left. + \frac{\epsilon^2}{\nu}(\mathbf{I} \otimes \mathbf{B}^{\dagger} + \mathbf{B}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}} \otimes \mathbf{I})^2\right\}\langle\mathbf{H}\rangle, \quad (C11)$$

which determines $\langle\mathbf{H}\rangle$. Eigenanalysis of $\langle\mathbf{H}\rangle$ determines the pure optimal initial state that leads to the largest square amplitude growth at time t.

REFERENCES

Ambaum, M. H. P., B. J. Hoskins, and D. B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? *J. Climate,* **14,** 3495–3507.

Arnold, L., 1992: *Stochastic Differential Equations: Theory and Applications.* Krieger, 228 pp.

Drazin, P. G., and W. H. Reid, 1981: *Hydrodynamic Stability.* Cambridge University Press, 527 pp.

Ehrendorfer, M., and J. J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.,* **54,** 286–313.

Farrell, B. F., and P. J. Ioannou, 1994: A theory for the statistical equilibrium energy and heat flux produced by transient baroclinic waves. *J. Atmos. Sci.,* **51,** 2685–2698.

——, and ——, 1995: Stochastic dynamics of the midlatitude atmospheric jet. *J. Atmos. Sci.,* **52,** 1642–1656.

——, and ——, 1996: Generalized stability. Part I: Autonomous operators. *J. Atmos. Sci.,* **53,** 2025–2041.

——, and ——, 1998: Perturbation structure and spectra in turbulent channel flow. *Theor. Comput. Fluid Dyn.,* **11,** 215–227.

——, and ——, 1999: Perturbation growth and structure in time-dependent flows. *J. Atmos. Sci.,* **56,** 3622–3639.

——, and ——, 2000: Perturbation dynamics in atmospheric chemistry. *J. Geophys. Res.,* **105** (D7), 9303–9320.

——, and ——, 2002: Perturbation growth and structure in uncertain flows. Part I. *J. Atmos. Sci.,* **59,** 2629–2646.

Horn, R. A., and C. R. Johnson, 1991: *Topics in Matrix Analysis.* Cambridge University Press, 607 pp.

Hughston, L. P., R. Jozsa, and W. K. Wooters, 1993: A complete classification of quantum ensembles having a given density matrix. *Phys. Lett.,* **183A,** 14–18.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.,* **127,** 279–304.

Sardeshmukh, P., C. Penland, and M. Newman, 2001: Rossby waves in a stochastically fluctuating medium. *Progress in Probability.* P. Imkeller and J. S. von Storch, Eds., Vol. 49, Birkhauser Verlag, 369–384.

Schrödinger, E., 1936: Probability relations between separated systems. *Proc. Cambridge Philos. Soc.,* **32,** 446–452.

Van Huffel, S., and J. Vandewalle, 1991: *The Total Least Squares Problem: Computational Aspects and Analysis.* SIAM, 300 pp.

Van Kampen, N. G., 1992: *Stochastic Processes in Physics and Chemistry.* Elsevier, 465 pp.