Distributed Forcing of Forecast and Assimilation Error Systems

BRIAN F. FARRELL

Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts

PETROS J. IOANNOU

Department of Physics, University of Athens, Athens, Greece

ABSTRACT

Temporally distributed deterministic and stochastic excitation of the tangent linear forecast system governing forecast error growth and the tangent linear observer system governing assimilation error growth is examined. The method used is to determine the optimal set of distributed deterministic and stochastic forcings of the forecast and observer systems over a chosen time interval. Distributed forcing of an unstable system addresses the effect of model error on forecast error in the presumably unstable forecast error system. Distributed forcing of a stable system addresses the effect on the assimilation of model error in the presumably stable data assimilation system viewed as a stable observer. In this study, model error refers both to extrinsic physical error forcing, such as that which arises from unresolved cumulus activity, and to intrinsic error sources arising from imperfections in the numerical model and in the physical parameterizations.

1. Introduction

A comprehensive understanding of the dynamics of linear systems is fundamental to physical science and, while linear systems arise in a wide variety of contexts, in this work we have in mind the particular examples of the dynamics of perturbations to the tangent linear forecast and data assimilation error systems. Among important problems in dynamical system theory is understanding the response of a linear system to forcing distributed in time. In the context of linear forecast error systems, temporally distributed forcing is used to account for the effect of model error on forecast predictability; while in the case of a data assimilation system it addresses the effect of model error on state identification for model initialization.

Because the forecast system is nonnormal as well as time dependent, its study requires the methods of generalized stability theory (GST; Farrell and Ioannou 1996a,b). These methods have heretofore been applied inter alia to the problem of cyclogenesis (Farrell 1982, 1989), the problem of transition to turbulence of boundary layer and laboratory channel flows (Farrell 1988a; Butler and Farrell 1992; Reddy and Henningson

1993), the maintenance of the turbulent state in laboratory shear flows (Farrell and Ioannou 1993b; Bamieh and Dahleh 2001; Jovanović and Bamieh 2001), the control of boundary layer and channel shear flow turbulence (Bewley and Liu 1998; Farrell and Ioannou 1998; Kim 2003; Högberg et al. 2003a,b), and in numerical weather prediction to determine the impact of uncertainties in the initial state on the forecast (Farrell 1988b; Lacarra and Talagrand 1988; Farrell 1990; Molteni and Palmer 1993; Buizza and Palmer 1995). However, as the initialization of forecast models is improved with the advent of new data sources and the introduction of variational assimilation methods, the medium range forecast will become increasingly affected by uncertainties resulting from incomplete physical parameterizations and numerical approximations in the forecast model, and by the necessarily misrepresented subgrid-scale chaotic processes such as cumulus convection that act as temporally stochastic distributed forcing of the forecast error system. These influences, referred to collectively as model error, conventionally appear as an external forcing in the forecast error system. Improving the understanding of model error and specifically identifying forcings that lead to the greatest forecast errors is centrally important in predictability studies. In analogy with the optimal perturbations that lead to the greatest forecast error in the case of initial condition error, these continuous error sources will be called optimal distributed forcings.

Corresponding author address: Dr. Brian F. Farrell, Harvard University, Pierce Hall 107V, Oxford St. Mail Area H0162, Cambridge, MA 02138.
E-mail: farrell@deas.harvard.edu

Two distinct dynamical representations for model error will be considered. The first is deterministic distributed forcing and the second is stochastic distributed forcing. The underlying theory for the deterministic case is based on analysis of the dynamical system as a mapping from the space of input forcings to the space of output states (Dullerud and Paganini 2000). This theory enabled Glover (1984) to obtain the optimal truncation of a dynamical system by using a singular value decomposition of the map between past forcings and future responses of the system (cf. Zhou and Doyle 1998). The singular values of this map, called Hankel singular values, have recently been used to truncate the linearized fluid equations and construct a reduced-order Kalman filter for use in meteorological assimilation (Farrell and Ioannou 2001a,b).

The link between distributed forcing and the growth of forecast errors is direct, but of equal interest is the role of distributed forcing in limiting the accuracy with which a system state can be identified. In order to address this problem we consider the data assimilation system viewed as a Luenberger observer (Luenberger 1971). Such an observer is necessarily asymptotically stable in contrast to the forward forecast problem, which is generally asymptotically unstable.

Recently D'Andrea and Vautard (2000) obtained approximate optimal temporally distributed deterministic forcings of the forecast error system (which they refer to as forcing singular vectors) and Barkmeijer et al. (2003) obtained the optimal temporally distributed deterministic forcing of the forecast error system over fixed spatial structures. In this paper, we extend the study of distributed forcings of the forecast error system by finding efficient methods for obtaining the exact solution of the optimal temporally distributed deterministic and stochastic forcing of the forecast and assimilation systems.

First we introduce the norms used to measure the forcing and the error state, and review the method used to obtain the optimal deterministic forcing that produces the greatest error state norm at a given time. The optimal problem for stochastic forcing and for impulsive forcing is then solved. Finally, examples are shown for optimal deterministic, stochastic, and impulsive forcing in forecast systems and observer systems.

2. Optimal deterministic forcing of linear systems

The problem of obtaining the forcing of a time-varying system over the time interval $[0, T]$ that leads to the greatest state norm at time T has been solved in the context of control theory (see, i.e., Dullerud and Paganini 2000). Because of the novelty and power of the methods used to solve this problem, and their relation to methods used for obtaining optimal truncation of meteorological systems (cf. Farrell and Ioannou 2001a,b), it is useful to review this solution.

We seek the deterministic forcing $f(t)$ of unit norm on $t \in [0, T]$ producing the greatest state norm at time T that is, it maximizes the square norm of the state $\|\psi(T)\|^2$, assuming the state is initially zero, $\psi(0) = 0$, and that ψ obeys the linear equation:

$$\frac{d\psi}{dt} = \mathbf{A}(t)\psi + \mathbf{f}(t). \tag{1}$$

The forcing $\mathbf{f}(t)$ is taken to be in $\mathcal{L}_2([0, T])$, that is, a square integrable vector function on the interval $[0, T]$ with an inner product:

$$(\mathbf{f}, \mathbf{g})_{\mathcal{L}_2} = \int_0^T \mathbf{f}^H(t)\mathbf{g}(t)dt. \tag{2}$$

The square norm in this space is

$$\|\mathbf{f}\|_{\mathcal{L}_2}^2 = \int_0^T \mathbf{f}^H(t)\mathbf{f}(t)dt. \tag{3}$$

The state ψ is in C^N, that is, an N dimensional complex vector. We choose the Euclidean inner product for the state vectors:

$$(\psi_1, \psi_2) = \psi_1^H\psi_2, \tag{4}$$

where H denotes the hermitian transpose. The square of the norm in this space is

$$\|\psi\|^2 = \psi^H\psi. \tag{5}$$

We have employed the subscript $\mathcal{L}_2$ to denote the inner product of square integrable functions in order to distinguish it from the inner product of vectors.

With these norms we seek the forcing, $\mathbf{f}$, maximizing the forcing normalized final state at time T:

$$R_d = \frac{\|\psi(T)\|^2}{\|\mathbf{f}\|_{\mathcal{L}_2}^2}. \tag{6}$$

The state ψ at time T is obtained for any forcing, $\mathbf{f}$, by the linear map ϕ:

$$\psi(T) = \phi\mathbf{f} = \int_0^T \mathbf{\Phi}(T, t)\mathbf{f}(t)dt. \tag{7}$$

In (7) $\mathbf{\Phi}(T, t)$ is the propagator associated with the linear dynamical system (1).

In order to obtain the optimal distributed forcing, some auxiliary results are needed. First the adjoint of the mapping ϕ is required. The adjoint of ϕ is denoted ϕ^* and it maps a state $\psi_1(T)$ to a forcing $\mathbf{g}(t)$ defined over the interval $[0, T]$, with the property that

$$[\phi^*\psi_1(T), \mathbf{f}(t)]_{\mathcal{L}_2} \equiv [\mathbf{g}(t), \mathbf{f}(t)]_{\mathcal{L}_2} \equiv \int_0^T \mathbf{g}^H(s)\mathbf{f}(s)ds$$

$$= [\psi_1(T), \psi(T)]. \tag{8}$$

This adjoint map ϕ^* is (cf. Farrell and Ioannou 2001a):

$$\phi^*\psi_1(T) = \mathbf{\Phi}^H(T, t)\psi_1(T). \tag{9}$$

Composition of the mapping ϕ and its adjoint ϕ^* maps system states to system states. This is a rank N mapping that can be seen using (7) and (9) to have matrix representation:

$$\mathbf{C} = \phi\phi^* = \phi\mathbf{\Phi}^{\mathrm{H}}(T, t) = \int_0^T \mathbf{\Phi}(T, t)\mathbf{\Phi}^{\mathrm{H}}(T, t)dt, \quad (10)$$

which is also the finite-time state covariance matrix under the assumption of temporally white noise forcing with unit covariance $\mathbf{I}$. The covariance matrix $\mathbf{C}$ in (10) is hermitian, and is assumed to be full rank. The covariance can also be obtained at time T by integrating from $t = 0$ to $t = T$ the Lyapunov equation:

$$\frac{d\mathbf{C}}{dt} = \mathbf{A}(t)\mathbf{C} + \mathbf{C}\mathbf{A}^{\mathrm{H}}(t) + \mathbf{I}, \quad (11)$$

with the initial condition $\mathbf{C}(0) = 0$ (cf. Farrell and Ioannou 1996a).

Now we define the projection operator:

$$P \equiv \phi^*\mathbf{C}^{-1}\phi. \quad (12)$$

Operator P maps forcings to forcings. This linear map is an orthogonal projector on the space of functions because it is hermitian and satisfies the property:

$$P^2 = P. \quad (13)$$

Indeed,

$$P^2 = \phi^*\mathbf{C}^{-1}\phi\phi^*\mathbf{C}^{-1}\phi = \phi^*\mathbf{C}^{-1}\phi = P. \quad (14)$$

Also note the property:

$$\phi P = \phi. \quad (15)$$

Indeed,

$$\phi P = \phi\phi^*\mathbf{C}^{-1}\phi = \mathbf{C}\mathbf{C}^{-1}\phi = \phi. \quad (16)$$

Consider a forcing $\mathbf{f}$, its projection $P\mathbf{f}$, and its orthogonal complement $(I - P)\mathbf{f}$. Then (15) establishes that forcing $(I - P)\mathbf{f}$ makes no contribution to the state at $\psi(T)$.

Assume that at time T, the system is in state ψ, and consider all the forcings $\mathbf{f}$ that could have produced this state, that is, all $\mathbf{f}$ with the property $\phi\mathbf{f} = \psi$. That there exist forcings for which any chosen state ψ can be reached at time T requires that the system be controllable, but this is ensured by the full rank of $\mathbf{C}$, which in the control literature is referred to as the controllability Grammian (Brockett 1970).

All the forcings, $\mathbf{f}$, that produce state ψ have the same projection $P\mathbf{f}$. To show this, note that by assumption:

$$\phi\mathbf{f} = \psi, \quad (17)$$

and by (15) the projected forcing $P\mathbf{f}$ also produces the state ψ so

$$\phi P\mathbf{f} = \psi. \quad (18)$$

Multiplying (18) by $\phi^*\mathbf{C}^{-1}$ gives

$$\phi^*\mathbf{C}^{-1}\phi P\mathbf{f} = \phi^*\mathbf{C}^{-1}\psi. \quad (19)$$

But the left side of (19) is $P^2\mathbf{f} = P\mathbf{f}$. Hence all $\mathbf{f}$ that produce state ψ have the same projection:

$$P\mathbf{f} = \phi^*\mathbf{C}^{-1}\psi. \quad (20)$$

It is now easy to see that the forcing $P\mathbf{f}$ is the minimal forcing producing the state ψ. To show this, we decompose the forcing into its orthogonal complements:

$$\mathbf{f} = P\mathbf{f} + (I - P)\mathbf{f}. \quad (21)$$

Because

$$P(I - P) = 0, \quad (22)$$

we have

$$\|\mathbf{f}\|_{\mathcal{L}_2}^2 = \|P\mathbf{f} + (I - P)\mathbf{f}\|_{\mathcal{L}_2}^2 = \|P\mathbf{f}\|_{\mathcal{L}_2}^2 + \|(I - P)\mathbf{f}\|_{\mathcal{L}_2}^2, \quad (23)$$

and it follows that

$$\|\mathbf{f}\|_{\mathcal{L}_2}^2 \geq \|P\mathbf{f}\|_{\mathcal{L}_2}^2 = \|\phi^*\mathbf{C}^{-1}\psi\|_{\mathcal{L}_2}^2 = \psi^{\mathrm{H}}\mathbf{C}^{-1}\psi, \quad (24)$$

which proves that the forcing

$$\mathbf{f}(t) = \phi^*\mathbf{C}^{-1}\psi, \quad (25)$$

is the minimal forcing to produce state ψ. It is easy to verify that this forcing produces ψ:

$$\phi(\phi^*\mathbf{C}^{-1}\psi) = (\phi\phi^*\mathbf{C}^{-1})\psi = \mathbf{C}\mathbf{C}^{-1}\psi = \psi. \quad (26)$$

Using (24) the maximization (6) over functions, $\mathbf{f}$, can be transformed to a maximization over states, $\psi(T)$:

$$\max_{\mathbf{f}(t)}\frac{\|\psi(T)\|^2}{\|\mathbf{f}(t)\|_{\mathcal{L}_2}^2} = \max_{\psi(T)}\frac{\|\psi(T)\|^2}{\psi(T)^{\mathrm{H}}\mathbf{C}^{-1}\psi(T)}. \quad (27)$$

This equation shows that the optimal over forcings (lhs) is equal to an optimal over states (rhs). Note that the form of the second optimization is reminiscent of the covariance or Mahalanobis metric often used in predictability analysis (Palmer et al. 1998) suggesting the interpretation of optimals weighted by the Mahalanobis metric as structures that are most easily forced.

Quotient (27) is maximized for unit forcing by the state:

$$\psi_{\mathrm{opt}}(T) = \sqrt{\lambda_1}\boldsymbol{v}_1, \quad (28)$$

where λ_1 is the maximum singular value of $\mathbf{C}$ and $\boldsymbol{v}_1$ is the corresponding singular vector of $\mathbf{C}$ [$\boldsymbol{v}_1$ is conventionally called the top empirical orthogonal function (EOF) of $\mathbf{C}$]. The optimal unit forcing is obtained from (9) and (25):

$$\mathbf{f}_{\mathrm{opt}}(t) = \phi^*\mathbf{C}^{-1}\psi_{\mathrm{opt}}(T) = \mathbf{\Phi}^{\mathrm{H}}(T, t)\mathbf{C}^{-1}\sqrt{\lambda_1}\boldsymbol{v}_1$$

$$= \frac{1}{\sqrt{\lambda_1}}\mathbf{\Phi}^{\mathrm{H}}(T, t)\boldsymbol{v}_1, \quad (29)$$

and the optimal value of the quotient (27) is clearly λ_1. It is also clear from (29) that the optimal forcing structure at optimizing time, $\mathbf{f}_{opt}(T)$, is identical to the optimal state $\boldsymbol{v}_1$. If the system is asymptotically stable, λ_1 remains finite as $T \rightarrow \infty$, and if the system is autonomous as well as stable, λ_1 asymptotes to a constant; while if the system is asymptotically unstable, λ_1 grows exponentially as $T \rightarrow \infty$. In the limit of small optimizing time $\mathbf{C} = O(T)$ and therefore (27) approaches zero linearly in time, a result that can be traced to the quadratic weighting of the forcing by $\|\cdot\|_{L_2}^2$, which has the effect of penalizing concentration of forcing in time.

If $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are two distinct singular vectors of $\mathbf{C}$, the corresponding forcings $\mathbf{f}_{1,2}(t) = \phi^* \mathbf{C}^{-1} \boldsymbol{v}_{1,2}$ are orthogonal because

$$\int_0^T \mathbf{f}_1^H(t) \mathbf{f}_2(t) \, dt = \boldsymbol{v}_1^H \mathbf{C}^{-1} \phi \phi^* \mathbf{C}^{-1} \boldsymbol{v}_2 = \boldsymbol{v}_1^H \mathbf{C}^{-1} \boldsymbol{v}_2$$

$$= \frac{1}{\lambda_2} \boldsymbol{v}_1^H \boldsymbol{v}_2 = 0. \quad (30)$$

The optimal forcing and the associated state of the system can be obtained simultaneously by integrating the following coupled forward and adjoint systems backward over the finite interval from time $t = T$ to the initial time $t = 0$:

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

$$\frac{d\mathbf{f}}{dt} = -\mathbf{A}^H(t)\mathbf{f}, \quad (31)$$

with $\boldsymbol{\psi}(T) = \sqrt{\lambda_1} \boldsymbol{v}_1$ and $\mathbf{f}(T) = \boldsymbol{v}_1 / \sqrt{\lambda_1}$. The initial state $\boldsymbol{\psi}_{opt}(0) = 0$ is recovered as a consistency check.

Consider forcings, $\mathbf{f}$, with $\|\mathbf{f}\|_{L_2}^2 = 1$. The states these produce will lie on an ellipsoid with axes of length $\sqrt{\lambda_i}$ in the directions of the orthonormal $\boldsymbol{v}_i$, where $\boldsymbol{v}_i$ are the singular vectors of $\mathbf{C}$ with corresponding singular values λ_i. The corresponding forcings

$$\mathbf{f}_i(t) = \frac{1}{\sqrt{\lambda_i}} \boldsymbol{\Phi}^H(T, t)\boldsymbol{v}_i, \quad (32)$$

which produce the extremal states $\boldsymbol{\psi}_i(T) = \sqrt{\lambda_i}\boldsymbol{v}_i$ are an orthonormal finite dimensional set of functions that span the space of minimal forcings. These forcings induce a characterization of $\boldsymbol{\psi}(T)$ and its excitation by distributed forcing analogous to the characterization induced by the evolved optimals and the optimal initial conditions in the initial condition excitation problem. For this reason we extend the characterization of optimality to all the forcings that render the Rayleigh quotient (6) extremal. A distinction between optimal distributed forcings and optimal initial conditions is that the set of optimal distributed forcings, $\mathbf{f}_i$, is the minimal spanning set of forcings that produce nonzero states at $t = T$, but the set of optimal forcings, $\mathbf{f}_i$, do not span the whole space $\mathcal{L}_2([0, T])$. However, the optimal forcings

do split the whole of $\mathcal{L}_2([0, T])$ into two orthogonal components: the range of P, $P\mathcal{L}_2([0, T])$, comprises linear combinations of the orthogonal set of $\mathbf{f}_i$ which span the forcings that contribute to the state at T, and the orthogonal space of functions in the kernel of P, $(I - P)\mathcal{L}_2([0, T])$, which do not contribute to the state at T.

An arbitrary forcing $\mathbf{f}$ may be projected on its minimal component $P\mathbf{f}$ with the property that only this component of the forcing produces a nonzero state at the optimizing time T.

Specifically, an arbitrary state $\boldsymbol{\psi}(T)$ can be expanded in the basis of the orthonormal eigenvectors, $\boldsymbol{v}_i$, of the finite-time covariance matrix $\mathbf{C}(T)$:

$$\boldsymbol{\psi}(T) = \sum_{i=1}^N \alpha_i \boldsymbol{v}_i, \quad (33)$$

where

$$\alpha_i = \boldsymbol{v}_i^H \boldsymbol{\psi}(T). \quad (34)$$

Then the minimal forcing,

$$\mathbf{f} = \sum_{i=1}^N \frac{\alpha_i}{\sqrt{\lambda_i}} \mathbf{f}_i, \quad (35)$$

is the only relevant forcing for producing this state. Further, because λ_i generally fall rapidly with i, most states are difficult to reach; for instance by (35), states $\boldsymbol{\psi}(T)$ that project on $\boldsymbol{v}_i$ with small values of λ_i require large-amplitude forcing. It follows that calculation of the first few eigenvectors, $\boldsymbol{v}_i$, of the finite-state covariance $\mathbf{C}(T)$ and the optimal forcings, $\mathbf{f}_i$, that produce them, suffices to characterize the potential influence of model error on the forecast and assimilation systems. In this way the decomposition of forcings into contributing and noncontributing spaces can be made independent of the state dimension N, when N is large, by taking the spanning set for the contributing space to be restricted to the forcings with substantial λ_i.

3. Optimal constant spatial structure for stochastic forcing of a linear system

In the previous section we obtained the optimal deterministic forcing for producing a specified state and the maximum state norm at a given time. However, situations arise in which the forcing is not known precisely but can be characterized statistically. A common characteristic of statistically specified, physically derived forcing is that the forcing is stochastic with a short correlation time when compared to the characteristic time scales of the linear operator. Given such a stochastic time dependence we seek the constant spatial forcing structure producing the greatest expected state square norm at a chosen time. This situation is formulated using the following stochastic equation:

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}(t)\boldsymbol{\psi} + \mathbf{f}\eta(t), \quad (36)$$

where $\mathbf{f}$ is a column vector specifying the spatial structure and $\eta(t)$ a temporally white noise scalar function of unit variance satisfying $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_2 - t_1)$.

The optimal spatial structure, $\mathbf{f}_{\text{opt}}$, of the stochastic forcing is to be found under the assumption that its temporal behavior is a white noise process, $\eta(t)$. This contrasts with the deterministic optimal forcing problem studied above, in which the optimal response was found over both the spatial and temporal structure of the forcing.

The problem of obtaining the fixed spatial structure of temporally white forcing that excites the greatest expected square state norm in the asymptotic limit $t \rightarrow \infty$ has been addressed previously (Farrell and Ioannou 1993a,b, 1996a; Kleeman and Moore 1997). This structure is the stochastic optimal. The problem of obtaining the analogous finite-time stochastic optimal under both certain and uncertain dynamics was studied by Farrell and Ioannou (2002a,b). We will now briefly review how the stochastic optimal is determined.

The ensemble expected square norm of the state at time T is

$$\langle \boldsymbol{\psi}^{\text{H}}(T)\boldsymbol{\psi}(T) \rangle =$$

$$\left\langle \int_0^T ds_1 \int_0^T ds_2 \, \eta^{\text{H}}(s_1) \mathbf{f}^{\text{H}} \boldsymbol{\Phi}^{\text{H}}(T, s_1) \boldsymbol{\Phi}(T, s_2) \mathbf{f} \eta(s_2) \right\rangle \tag{37}$$

$$= \mathbf{f}^{\text{H}} \mathbf{B}(T) \mathbf{f}, \tag{38}$$

where the matrix $\mathbf{B}(T)$ is

$$\mathbf{B}(T) = \int_0^T \boldsymbol{\Phi}^{\text{H}}(T, s) \boldsymbol{\Phi}(T, s) \, ds. \tag{39}$$

The matrix $\mathbf{B}(T)$ can be obtained by integrating from $t = 0$ to $t = T$ the equation,

$$\frac{d\mathbf{B}}{dt} = \mathbf{A}^{\text{H}}(t)\mathbf{B} + \mathbf{B}\mathbf{A}(t) + \mathbf{I}, \tag{40}$$

with the initial condition $\mathbf{B}(0) = 0$.

It is immediate from (38) that the singular vectors of the hermitian matrix $\mathbf{B}$ are the finite-time stochastic optimals. The finite-time stochastic optimal associated with the largest singular value of $\mathbf{B}$ is the unit norm spatial forcing structure that, when excited white in time over the interval $[0, T]$, produces the greatest expected square norm of the state at time T. This unit norm spatial structure results in unit variance forcing in (36). This largest expected square state response, when normalized by the forcing covariance, is

$$\frac{\langle \boldsymbol{\psi}^{\text{H}}(T)\boldsymbol{\psi}(T) \rangle_{\text{opt}}}{\mathbf{f}_{\text{opt}}^{\text{H}} \mathbf{f}_{\text{opt}}} = \frac{\mathbf{f}_{\text{opt}}^{\text{H}} \mathbf{B}(T) \mathbf{f}_{\text{opt}}}{\mathbf{f}_{\text{opt}}^{\text{H}} \mathbf{f}_{\text{opt}}} = \mu_1, \tag{41}$$

where μ_1 is the largest singular value of $\mathbf{B}(T)$ and $\mathbf{f}_{\text{opt}}$ the corresponding singular vector. The forcing normalizations in (41) are not the same as in the deterministic case. Deterministic forcings are normalized by the forcing's integrated square amplitude as in (3). Such a normalization is not appropriate for white noise forcing because white noise realizations are unbounded in square amplitude. The normalization in (41) assures a unit variance forcing in (36), which is appropriate for white noise forcing.

4. Optimal impulsive forcing on an interval

Consider an impulsive forcing applied at a point in time, t_0:

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}(t)\boldsymbol{\psi} + \mathbf{f}\delta(t - t_0), \tag{42}$$

where $\mathbf{f}$ is the spatial structure of the forcing. It is assumed, as previously, that the state is initially zero, $\boldsymbol{\psi}(0) = 0$, and that the forcing is applied at a point in the interval so that, $0 \leq t_0 \leq T$, where T is a final time.

We seek the time t_0 and the forcing spatial structure $\mathbf{f}$ of unit norm, corresponding to unit impulse, resulting in the greatest state norm at T, that is, we seek to maximize the quotient:

$$R_p = \frac{\boldsymbol{\psi}^{\text{H}}(T)\boldsymbol{\psi}(T)}{\mathbf{f}^{\text{H}}\mathbf{f}}. \tag{43}$$

The state at time T is

$$\boldsymbol{\psi}(T) = \boldsymbol{\Phi}(T, t_0)\mathbf{f}. \tag{44}$$

Consequently, we seek the spatial structure that maximizes

$$R_p = \frac{\mathbf{f}^{\text{H}}\boldsymbol{\Phi}^{\text{H}}(T, t_0)\boldsymbol{\Phi}(T, t_0)\mathbf{f}}{\mathbf{f}^{\text{H}}\mathbf{f}}. \tag{45}$$

The maximum value of R_p is the square of the largest singular value σ_1^2, of the propagator $\boldsymbol{\Phi}(T, t_0)$, which is also the square of the norm of $\boldsymbol{\Phi}(T, t_0)$, denoted $\|\boldsymbol{\Phi}(T, t_0)\|_2^2$. The forcing applied at t_0 producing the greatest state norm at T is the top singular vector of the propagator $\boldsymbol{\Phi}(T, t_0)$, which is the deterministic global optimal initial condition corresponding to optimizing time $T - t_0$. Therefore, the optimal point forcing has the structure of the global optimal initial condition over $t \in [0, T]$ and is applied at the global optimal time in the past. If we denote by t_{go} the time t for which $\|\boldsymbol{\Phi}(T, t)\|_2$ is maximum, that is, the time at which the global optimal growth is achieved, then the forcing is applied at $t_0 = t_{\text{go}}$. It is clear that the optimal point forcing is identical to the optimal initial condition at t_0.

Given an arbitrary final state $\boldsymbol{\psi}(T)$ the optimal impulsive forcing to produce $\boldsymbol{\psi}(T)$ is obtained by finding

$$\min_{t \in [0,T]} \|\boldsymbol{\Phi}(t, T)\boldsymbol{\psi}(T)\|, \tag{46}$$

in which $\boldsymbol{\Phi}(t, T) = \boldsymbol{\Phi}^{-1}(T, t)$, and taking t_0 as the time of this minimum and $\mathbf{f} = \boldsymbol{\Phi}(t, T)\boldsymbol{\psi}(T)\delta(t - t_0)$ as the impulsive forcing that optimally excites state $\boldsymbol{\psi}(T)$.

A related problem is finding the minimum constant amplitude time-varying structure over an interval that produces a given state. The solution is to force the normalized function obtained from the backward propagator:

$$\mathbf{f}(t) = \boldsymbol{\Phi}(t, T)\boldsymbol{\psi}(T) \Big/ \int_0^T \frac{\|\boldsymbol{\Phi}(t, T)\boldsymbol{\psi}(T)\|}{\|\boldsymbol{\Phi}(s, T)\boldsymbol{\psi}(T)\|} \, ds. \quad (47)$$

5. Examples

a. A scalar system example

Consider the one-dimensional and necessarily normal system:

$$\frac{d\psi}{dt} = \alpha\psi + f(t). \quad (48)$$

From (28) and (29) and the fact that a one-dimensional system is self-adjoint, it follows that the deterministic unit norm forcing maximizing the state at optimizing time T is

$$f_{\mathrm{opt}}(t) = \sqrt{\frac{2\alpha}{e^{2\alpha T} - 1}} \, e^{\alpha(T-t)}, \quad (49)$$

and the corresponding maximum state at optimizing time T is

$$\psi_{\mathrm{opt}}(T) = \sqrt{\frac{e^{2\alpha T} - 1}{2\alpha}}. \quad (50)$$

The optimal forcing for this example is shown in Fig. 1a for a representative stable case ($\alpha = -1$) and in Fig. 1b for a similar unstable case ($\alpha = 1$); both for $T = 5$. Note that the forcing in the stable case is concentrated near the optimizing time at the end of the interval, because forcing at earlier times would excite the decaying state and, as a result, be suboptimal. In the unstable case the forcing is concentrated at the start of the interval to take advantage of the growth. This is a general property of optimal forcings. The square norm of the state at the optimizing time T as a function of T is shown for the stable case in Fig. 1c (case labeled deterministic). For small T the square norm of the state is $|\psi_{\mathrm{opt}}(T)|^2 \approx T$ but for large T it asymptotes to $1/2\alpha$. The optimal forcing at optimizing time T has norm $f_{\mathrm{opt}}(T) = 1/\sqrt{\lambda}$ while the optimal state norm is $\psi_{\mathrm{opt}} = \sqrt{\lambda}$ with an optimal finite-time variance:

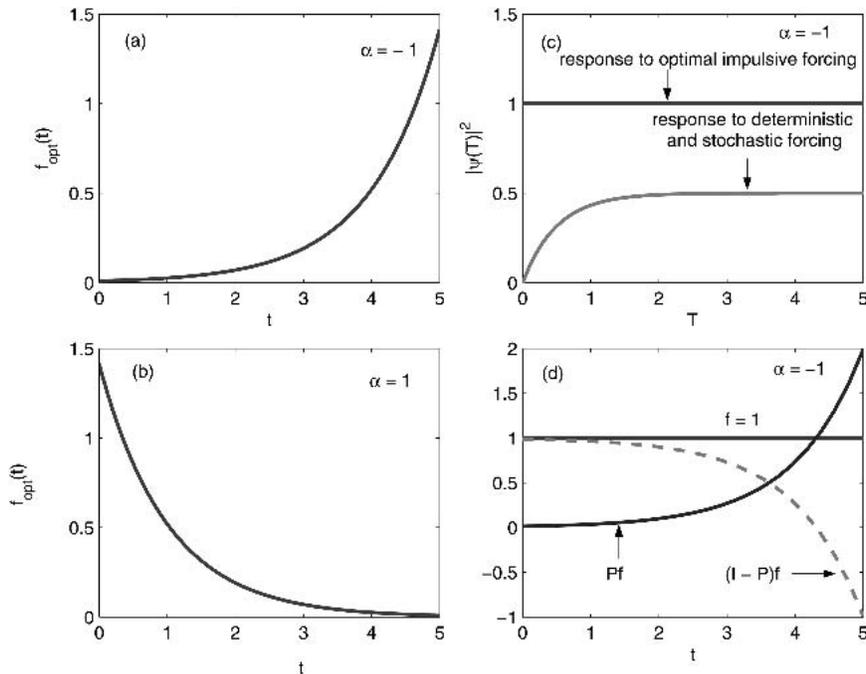$$\lambda = \frac{e^{2\alpha T} - 1}{2\alpha}. \quad (51)$$

FIG. 1. For the one-dimensional system (48): (a) unit norm optimal forcing as a function of time for the stable case with $\alpha = -1$ and $T = 5$, (b) unit norm optimal forcing for the unstable case with $\alpha = 1$ and $T = 5$, (c) square of the final state for the optimal deterministic, stochastic, and impulsive forcings at optimizing time T as a function of the optimizing time for the stable case with $\alpha = -1$. The optimal deterministic and the optimal stochastic responses coincide. (d) A decomposition of $f(t) = 1$ into the component contributing, Pf, and the component not contributing, $f - Pf$, to the state at $t = 5$; also for $\alpha = -1$.

As in this example, it is generally the case that in stable normal systems the optimal forcing extends over a time interval of the order of the decay time of the system while in nonnormal systems the optimal forcing assumes the temporal extent of the system's global optimal as will be shown. The expected square norm of the state at T for unit covariance stochastic forcing is easily seen to be

$$\langle |\boldsymbol{\psi}(T)|^2\rangle_{\text{opt}} = \frac{e^{2\alpha T} - 1}{2\alpha}, \tag{52}$$

which is identical to the optimal deterministic state square norm. This is shown in Fig. 1 for $a = -1$. Also shown in Fig. 1 is the norm of the state for optimal impulsive forcing, which is identical to the optimal obtained from any unit initial condition on the interval. Because in the case of a stable normal system the amplitude of the states decays monotonically after they are excited, the optimal point for impulsive forcing is at the optimizing time T itself and consequently the optimal response is equal to 1 for all choices of T.

Assume that in this simple scalar system we are given a forcing $f(t)$. To find the part of $f(t)$ that influences the state at T, we need only to calculate the projection of $f(t)$ on the space of optimal forcing that according to (12) is

$$Pf = e^{\alpha(T-t)} \frac{2\alpha}{e^{2\alpha T} - 1} \int_0^T e^{\alpha(T-s)} f(s)\, ds. \tag{53}$$

The part of $f(t)$ that does not influence the final state is $f - Pf$. For example if the forcing were constant, $f(t) = 1$, only the projection,

$$Pf = 2 \frac{e^{\alpha T} - 1}{e^{2\alpha T} - 1} e^{\alpha(T-t)}, \tag{54}$$

produces a contribution to the state at T, while $f - Pf$ produces no contribution. This decomposition of $f(t) = 1$ into contributing and noncontributing components is shown in Fig. 1.

b. The Reynolds system example

Consider the two-dimensional highly nonnormal Reynolds system that models streak formation in boundary layers:

$$\frac{d\boldsymbol{\psi}}{dt} = \mathbf{A}\boldsymbol{\psi} + \mathbf{f}(t), \tag{55}$$

with stable matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} -1 & 10 \\ 0 & -2 \end{pmatrix}. \tag{56}$$

The square norm of the state at time T produced by optimal deterministic forcing over the interval $t \in [0, T]$ is shown as a function of T in Fig. 2. The response is much higher than that found for a normal system with the same eigenvalues because of the nonnormality of the system [e.g., (56)]. The variance equation for the above system is
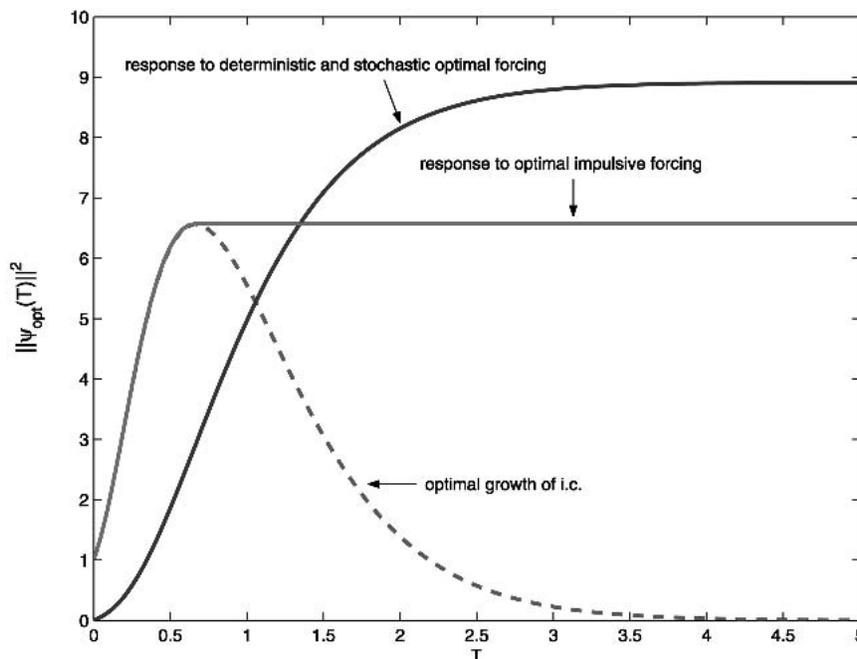

FIG. 2. For the Reynolds system example: the optimal square norm of the state $\|\boldsymbol{\psi}_{\text{opt}}(T)\|^2$ at optimizing time for deterministic, stochastic, and impulsive forcing as a function of optimizing time T. For comparison, the optimal growth for unit norm initial conditions, $\|e^{\mathbf{A}T}\|^2$, is also shown. Note that the optimal stochastic response coincides with the optimal deterministic response in this two-dimensional example.

$$\frac{d\|\boldsymbol{\psi}\|^2}{dt} = \boldsymbol{\psi}^{\mathrm{H}}(\mathbf{A}^{\mathrm{H}} + \mathbf{A})\boldsymbol{\psi} + (\mathbf{f}\boldsymbol{\psi}^{\mathrm{H}} + \boldsymbol{\psi}\mathbf{f}^{\mathrm{H}}). \quad (57)$$

The first term $\boldsymbol{\psi}^{\mathrm{H}}(\mathbf{A}^{\mathrm{H}} + \mathbf{A})\boldsymbol{\psi}$ represents the variance tendency from the background mediated by the nonnormality of $\mathbf{A}$ while $(\mathbf{f}\boldsymbol{\psi}^{\mathrm{H}} + \boldsymbol{\psi}\mathbf{f}^{\mathrm{H}})$ is the variance tendency directly accumulated from the forcing. These two terms are plotted in Fig. 3c as a function of time for the forcing that leads to maximum state norm at time $T = 5$ in system (55). Note that the variance obtained directly from forcing is far less than the variance obtained from nonnormal processes.

Because this is a two-dimensional example, the state at $t = T$ produced by the optimal stochastic and deterministic forcing are identical.[1] The optimal norm of the state produced by impulsive forcing and, for comparison, the optimal initial value growth $\|e^{\mathbf{A}T}\|^2$ are shown as a function of T in Fig. 2. Note that the norm of the

_____

[1] The reason is that in two-dimensional systems the eigenvalues of the covariance matrix $\mathbf{C}$ and the stochastic optimal matrix $\mathbf{B}$ are equal because the $\mathbf{B}$ and $\mathbf{C}$ matrices possess two invariants: they have the same trace and also satisfy the relation $\mathfrak{R}[\mathrm{trace}(\mathbf{AC})] = \mathfrak{R}[\mathrm{trace}(\mathbf{AB})]$, which can be easily verified from their respective Lyapunov equations ($\mathfrak{R}$ denotes the real part). The eigenvalues of $\mathbf{C}$ and $\mathbf{B}$ are, in general, different for nonnormal systems with dimension greater than two.

state produced by optimal impulsive forcing follows the optimal growth curve until T reaches the global optimal time t_{go} (which occurs for this case at $t_{\mathrm{go}} = 0.67$), then for $T > t_{\mathrm{go}}$ it maintains this constant value. The optimal impulsive forcing is applied at time $t = T - t_{\mathrm{go}}$ and has the structure of the global optimal initial condition.

The structure of the forcing as a function of time is shown in Fig. 3a and the corresponding evolution of the state norm is shown in Fig. 3b. The forcing is seen to peak t_{go} units of times prior to the optimizing time T.

6. Modeling data assimilation as a stable observer system

Consider assimilating data taken from the exact solution of a nonlinear system, $\mathbf{x}_t$. The forecast error $\mathbf{e}_f = \mathbf{x}_f - \mathbf{x}_t$ obeys the following equation:

$$\frac{d\mathbf{e}_f}{dt} = \mathbf{A}\mathbf{e}_f + \mathbf{f}(t), \quad (58)$$

in which model error is represented by $\mathbf{f}(t)$ and $\mathbf{A}$ is the tangent linear system about $\mathbf{x}_t$. Forecast error typically grows exponentially when averaged over a sufficiently long time due to the Lyapunov instability that inherently results from the time dependence of $\mathbf{A}$ (Arnold 1998; Farrell and Ioannou 1999). However, a Lyapunov
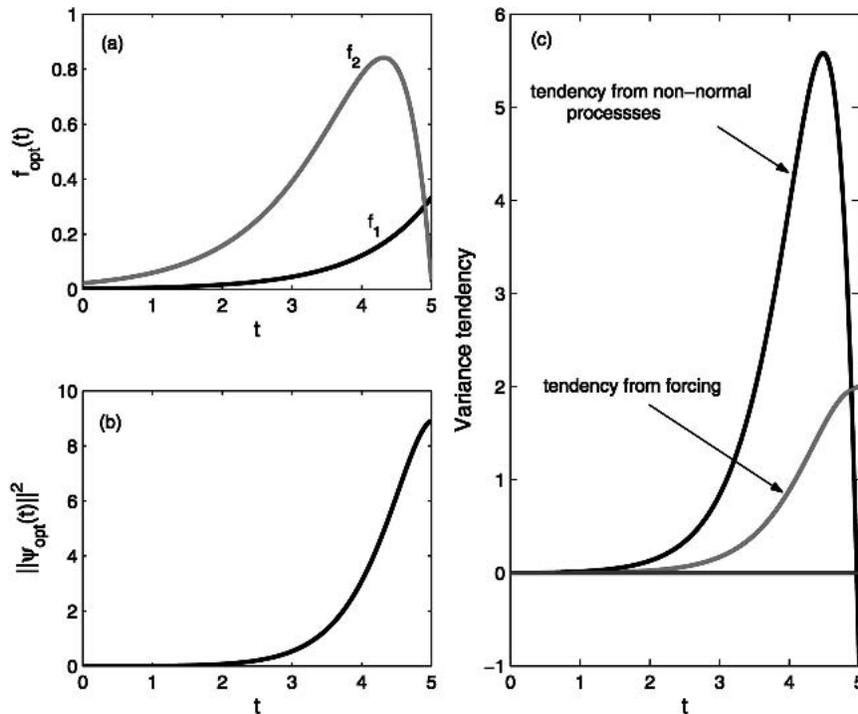

FIG. 3. For the Reynolds system example: (a) time development of the components of the optimal forcing $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$ that optimize the state energy at $T = 5$; (b) corresponding evolution of the square norm of the state; (c) variance tendency due to nonnormal processes $\boldsymbol{\psi}_{\mathrm{opt}}^{\mathrm{H}}(t)$ $(\mathbf{A} + \mathbf{A}^{\mathrm{H}})\boldsymbol{\psi}_{\mathrm{opt}}(t)$ and to forcing $\mathbf{f}_{\mathrm{opt}}^{\mathrm{H}}(t)\boldsymbol{\psi}_{\mathrm{opt}}(t) + \boldsymbol{\psi}_{\mathrm{opt}}^{\mathrm{H}}(t)\mathbf{f}_{\mathrm{opt}}(t)$ as a function of time. The variance tendency due to nonnormal processes exceeds variance tendency due to forcing.

stable system related to (58) can be constructed to model the process of state estimation for forecast initialization, a process referred to as data assimilation. This can be done because a data assimilation system is equivalent to a system in which the matrix $\mathbf{A}$ in (58) is modified by the introduction of observations so that the solution to the modified system, $\mathbf{x}_a$, converges to the system state, $\mathbf{x}_t$, in a statistical sense (Luenberger 1971). The effect of model and observational error on data assimilation can be analyzed quite generally using this representation of a data assimilation system as a Luenberger observer.

Assume that observations, $\mathbf{y}_{ob}$ of the state are taken at various locations and introduced into the model. Let the observations be related to the true state of the atmosphere, $\mathbf{x}_p$, as

$$\mathbf{y}_{ob} = \mathbf{H}\mathbf{x}_t + \mathbf{R}^{1/2}\mathbf{w}, \qquad (59)$$

where $\mathbf{H}$ is an observation matrix, $\mathbf{R}$ is the observational error covariance, and $\mathbf{w}$ is a vector of white noise processes. Given an observational platform, $\mathbf{y}_{ob}$ is available while the truth $\mathbf{x}_t$ is not available.

Assimilate these observations to obtain an analysis, $\mathbf{x}_a$, with analysis error $\mathbf{e}_a = \mathbf{x}_a - \mathbf{x}_t$ satisfying the Luenberger observer system:

$$\frac{d\mathbf{e}_a}{dt} = \mathbf{A}\mathbf{e}_a + \mathbf{K}(\mathbf{y}_{ob} - \mathbf{H}\mathbf{x}_a) + \mathbf{f}(t)$$

$$= (\mathbf{A} - \mathbf{K}\mathbf{H})\mathbf{e}_a + \mathbf{K}\mathbf{R}^{1/2}\mathbf{w} + \mathbf{f}(t). \qquad (60)$$

The above equation defines the assimilation system model. Ideally, the gain, $\mathbf{K}$, is chosen to minimize the analysis error variance trace ($\langle \mathbf{e}_a \mathbf{e}_a^{\mathrm{H}} \rangle$).

Unlike the forecast error system, a Luenberger observer system is asymptotically stable. Any gain, $\mathbf{K}$, that stablizes $\mathbf{A} - \mathbf{K}\mathbf{H}$ results in an observer with bounded error, this error being forced by a combination of model error $\mathbf{f}(t)$ and observational error $\mathbf{R}$ (cf. 60). It is immediately evident that in the absence of model and observation error the analysis will converge to truth. Good gains, $\mathbf{K}$, do not just stabilize the operator but simultaneously reduce the nonnormality of the tangent linear operator so that the minimum of trace ($\langle \mathbf{e}_a \mathbf{e}_a^{\mathrm{H}} \rangle$) is maintained by the combination of observational and model error. The optimal gain under the assumption of Gaussian error statistics is the Kalman gain (Kalman 1964; Zhou and Doyle 1998). A suboptimal observer, such as might be associated with an operational forecast system, can be modeled as a stabilizing $\mathbf{K}$ not equal to the Kalman gain. Just as analysis of the tangent linear forecast system reveals how initialization error and distrib-
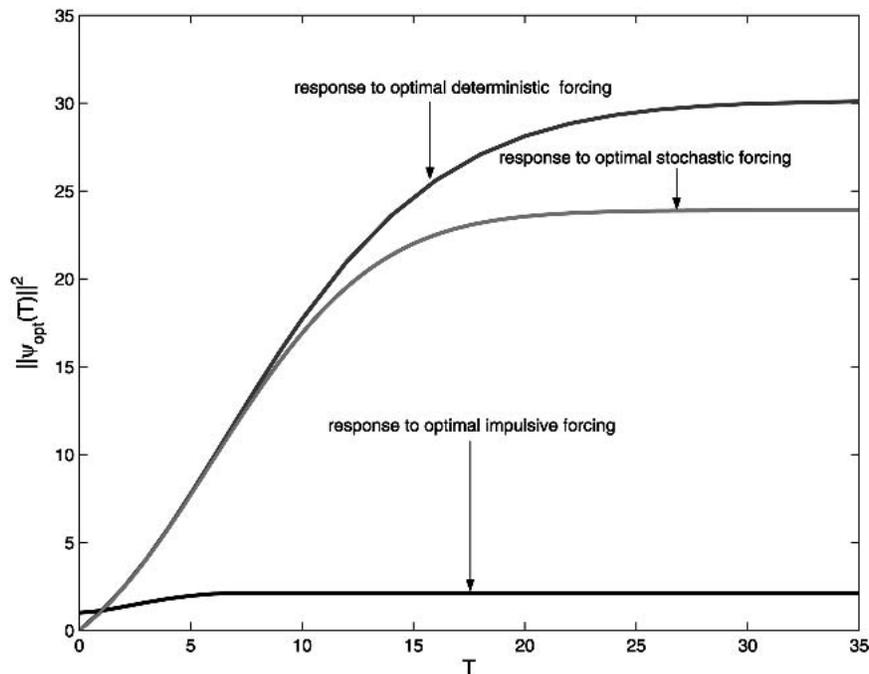

FIG. 4. For the observer-stabilized Eady model: the square norm of the optimal state at optimizing time T for unit norm deterministic distributed forcing; the optimal expected square norm of the state for statistically steady stochastic forcing; and the optimal square norm of the state for impulsive forcing. The zonal wavenumber is $k = 1.5$ and the Rayleigh damping coefficient is $r = 0.15$. The Eady model is stabilized by the asymptotic Kalman gain, obtained under the perfect model assumption and with unit observation error variance. The unobserved system is unstable with dimensional growth rate $1/2.3 \; \mathrm{day}^{-1}$. The decay rate of perturbations in the observer system is $1/2.6 \; \mathrm{day}^{-1}$.

uted forcing contribute to forecast error, so does analysis of the observer system reveal how model and observation error contribute to analysis error.

We are now equipped to study the effect of distributed error on data assimilation and will begin by seeking the deterministic and stochastic forcings leading to the greatest deviations of the estimate $\mathbf{x}_a$ from the true state $\mathbf{x}_t$ as measured by the norm of $\mathbf{e}$.

7. The Eady assimilation and forecast system model example

Consider as an example error system the Eady model (Eady 1949) with a zonal wavenumber chosen so that the model is unstable. Introduce continuous observation of the error streamfunction at the channel center and construct a stable Kalman observer under the assumption of zero model and unit observation error. Under these assumptions, a single continuous observation at the channel center suffices to make this optimal observer asymptotically stable.

The nondimensional linearized equation for streamfunction perturbations ψ in the Eady model is

$$\frac{\partial D^2 \psi}{\partial t} = -ik\,z\,D^2\psi - r\,D^2\psi, \tag{61}$$

in which the perturbation is assumed to be of form $\psi(z, t)e^{ikx}$, where k is the zonal wavenumber. The per-

turbation potential vorticity is $D^2\psi$, with $D^2 \equiv \partial^2/\partial z^2 - k^2$ and the perturbation potential vorticity damping rate is r. This model is Boussinesq with constant stratification on an f plane and has periodic boundary conditions in the zonal, x, direction and a solid lid at height $z = H$. The zonal flow is $U(z) = z$. Horizontal scales are nondimensionalized by $L = 1000$ km; vertical scales by $H = fL/N = 10$ km; velocity by $U_0 = 30$ m s; and time by $T = L/U_0$, so that a time unit is approximately 9.25 h. Midlatitude values are chosen for the Brunt–Väisälä frequency, $N = 10^{-2}$ s^{-1}, and Coriolis parameter, $f = 10^{-4}$ s^{-1}. The coefficient of Rayleigh damping has value $r = 0.15$ corresponding to a decay rate of $1/2.5$ day^{-1}.

Conservation of potential temperature at the ground, $z = 0$, and tropopause, $z = 1$, provides the following boundary conditions:

$$\frac{\partial^2 \psi}{\partial t \partial z} = ik\psi - r\frac{\partial \psi}{\partial z} \quad \text{at } z = 0, \tag{62}$$

$$\frac{\partial^2 \phi}{\partial t \partial z} = -ik\frac{\partial \psi}{\partial z} + ik\psi - r\frac{\partial \psi}{\partial z} \quad \text{at } z = 1. \tag{63}$$

The model is discretized using $N = 21$ points in the vertical and expressed in the matrix form:

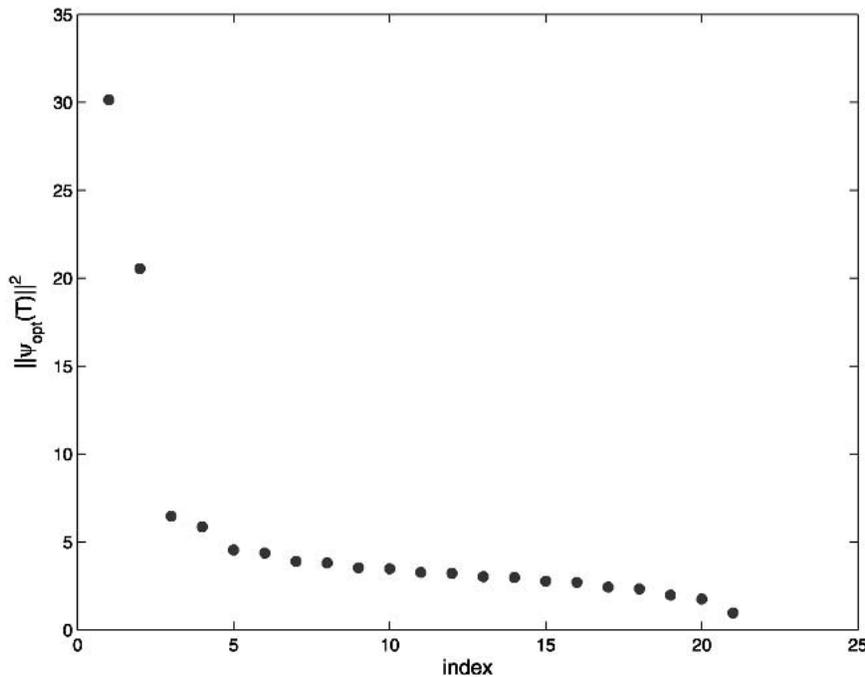$$\frac{d\phi}{dt} = \mathbf{A}\phi. \tag{64}$$

FIG. 5. Optimal deterministic square state norm at $T = 40$ in the observer-stabilized Eady model accounted for by each of the 21 orthogonal forcing structures. There are two dominant responses to deterministic forcing, suggesting the generally valid result that a few forcings account for most assimilation error. The parameters are as described in Fig. 4.

In (64) generalized velocity coordinates $\boldsymbol{\phi} = \mathbf{M}^{1/2}\boldsymbol{\psi}$ were used so that the Euclidian square norm of the system state corresponds to error energy. The error energy is $\boldsymbol{\phi}^{\mathrm{H}}\boldsymbol{\phi} = \boldsymbol{\psi}^{\mathrm{H}}\mathbf{M}\boldsymbol{\psi}$, and for a vertical discretization increment δz the metric $\mathbf{M}$ is

$$\mathbf{M} = -\frac{\delta z}{4} D^2. \tag{65}$$

The assimilation system in generalized velocity variables is governed by the operator $\mathbf{A}_a = \mathbf{A} - \mathbf{KH}$, where $\mathbf{K}$ is the Kalman gain for observation at the channel center given by $\mathbf{K} = \mathbf{XH}^{\mathrm{H}}\mathbf{R}^{-1}$, where $\mathbf{X}$ satisfies the continuous-time algebraic Ricatti equation under the perfect model assumption (cf. Zhou and Doyle 1998):

$$\mathbf{AX} + \mathbf{XA}^{\mathrm{H}} - \mathbf{XH}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{HX} = 0, \tag{66}$$

and $\mathbf{H} = \mathbf{PM}^{-1/2}$ is the observation matrix, in which, for our example, $\mathbf{P}$ is a row vector with one nonzero element corresponding to the grid point at the channel center, and $\mathbf{R}$ is the observation error covariance matrix (which is a scalar in our example). In contrast, the unstable forecast system is governed by the operator $\mathbf{A}$. We seek in both the stable assimilation and the unstable forecast system the optimal deterministic and stochastic forcing leading to the greatest error at time T.

a. The observer-stabilized Eady model

Let $\mathbf{A}_a$ be the matrix of the observer-stabilized Eady model for zonal wavenumber $k = 1.5$. While the unobserved Eady model at this wavenumber is unstable with dimensional growth rate $1/2.5$ day^{-1}, the observer-stabilized model has a minimum dimensional decay rate $1/2.6$ day^{-1}. The square norm of the state at time T produced by optimal distributed forcing over the interval $t \in [0, T]$ is shown as a function of T in Fig. 4. The square norm of the state at T produced by each of the 21 orthonormal optimal forcings is shown in Fig. 5. Two EOFs and their associated forcing structures account for most of the perturbation dynamics suggesting the generally valid conclusion that a few forcings account
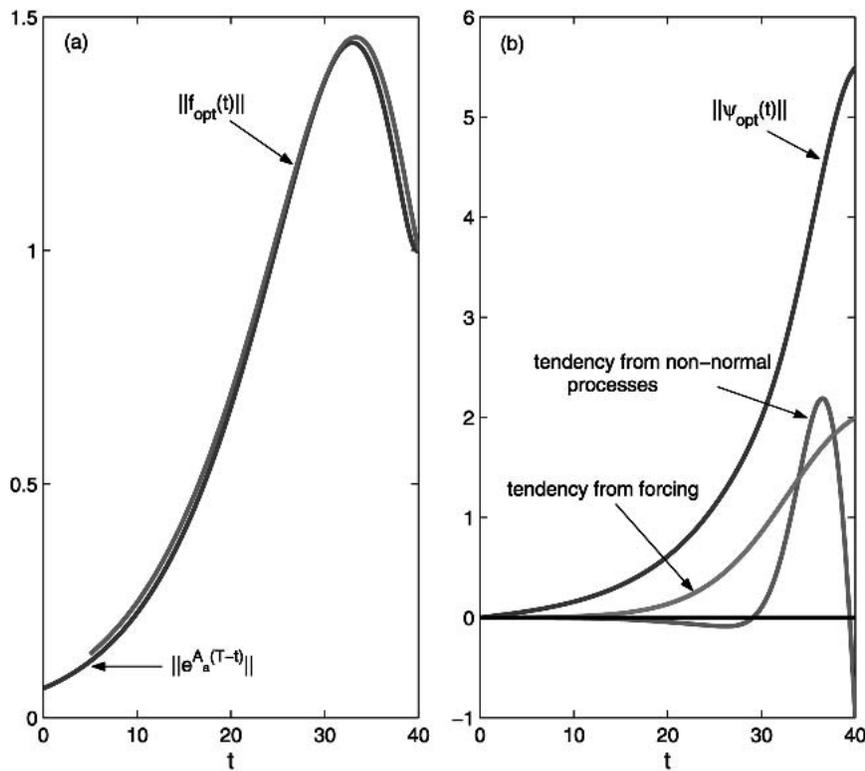

FIG. 6. For the observer-stabilized Eady model: (a) time development of the norm of the deterministic forcing that optimizes the energy of the state at $T = 40$. The norm of the retarded propagator $\|e^{\mathbf{A}_a(T-t)}\|$ giving the optimal growth attainable from an initial state is also shown. Note that the optimal deterministic forcing is concentrated in the interval of the supreme of the maximum norm of the retarded propagator. (b) The corresponding time development of the norm of the state $\|\boldsymbol{\psi}_{\mathrm{opt}}(t)\|$, the variance tendency due to nonnormal processes $\boldsymbol{\psi}_{\mathrm{opt}}^{\mathrm{H}}(t)(\mathbf{A} + \mathbf{A}^{\mathrm{H}})\boldsymbol{\psi}_{\mathrm{opt}}(t)$, and the variance tendency due to forcing $\mathbf{f}_{\mathrm{opt}}^{\mathrm{H}}(t)\boldsymbol{\psi}_{\mathrm{opt}}(t) + \boldsymbol{\psi}_{\mathrm{opt}}^{\mathrm{H}}(t)\mathbf{f}_{\mathrm{opt}}(t)$. The forcing is normalized so that $\|\mathbf{f}_{\mathrm{opt}}(T)\|_{L_2} = 1$. The parameters are the same as in Fig. 4.

for most of the error and the majority of the forcings make relatively little contribution.

The norm of the distributed forcing that maximizes the state norm at $T = 40$ is shown as a function of time in Fig. 6a along with the optimal growth $\|e^{\mathbf{A}_a(T-t)}\|$. The norm of the forcing follows the optimal growth curve and is maximized at $T - t_{go}$, where t_{go} is the time for which the maximum of the norm of the propagator $e^{\mathbf{A}_a t}$ occurs (in this case the maximum occurs at $t_{go} = 7$). The corresponding growth of the error state energy, the energy tendency by nonnormal processes, and the energy tendency produced by forcing are shown in Fig. 6b.

The evolution in structure of the first deterministic optimal is shown in the left panels of Fig. 7. The corresponding evolution in structure of the state is shown in the right panels. At optimizing time both the state and the optimal forcing assume the structure of the first eigenvector of the finite-time covariance matrix at time T, $\mathbf{C}(T)$.

The error norm produced by optimal stochastic forcing, and the response to optimal impulsive forcing are shown as a function of T in Fig. 4. The norm of the error due to impulsive forcing follows the deterministic initial state optimal growth curve reaching at time $t_{go} = 7$ the global optimal norm after which it assumes the constant value of the global optimal growth. For optimizing times $T > t_{go}$ the optimal impulsive forcing occurs at

time $t = T - t_{go}$ and has the structure of the global optimal initial condition. This structure and that of the stochastic optimal are shown in Fig. 8. If the spatial structure of the white noise forcing is taken to be that of the stochastic optimal shown in Fig. 8a then the resulting state covariance at optimizing time is primarily composed of the first three eigenvectors of $\mathbf{C}(T)$ shown in Fig. 9.

b. The unstable Eady forecast model

The error in the unstable Eady model at $k = 1.5$ grows exponentially. The error norm for optimal distributed, stochastic, and impulsive forcing are shown in Fig. 10. The optimal stochastic forcing and the optimal deterministic forcing produce almost identical error at time T because the system is unstable and excitation of a single structure, the adjoint of the mode, is responsible for nearly all the error at time T. The time development of the norm of the optimal deterministic forcing for optimizing time $T = 20$ is shown in Fig. 11. The forcing is concentrated at the beginning of the interval in order to take advantage of the instability. The structure of the forcing and the error at various times are shown in Fig. 12. At the optimizing time the structure both of the error and the forcing is that of the fastest growing mode, at earlier times the forcing has the structure of the adjoint of this mode. The stochastic optimal
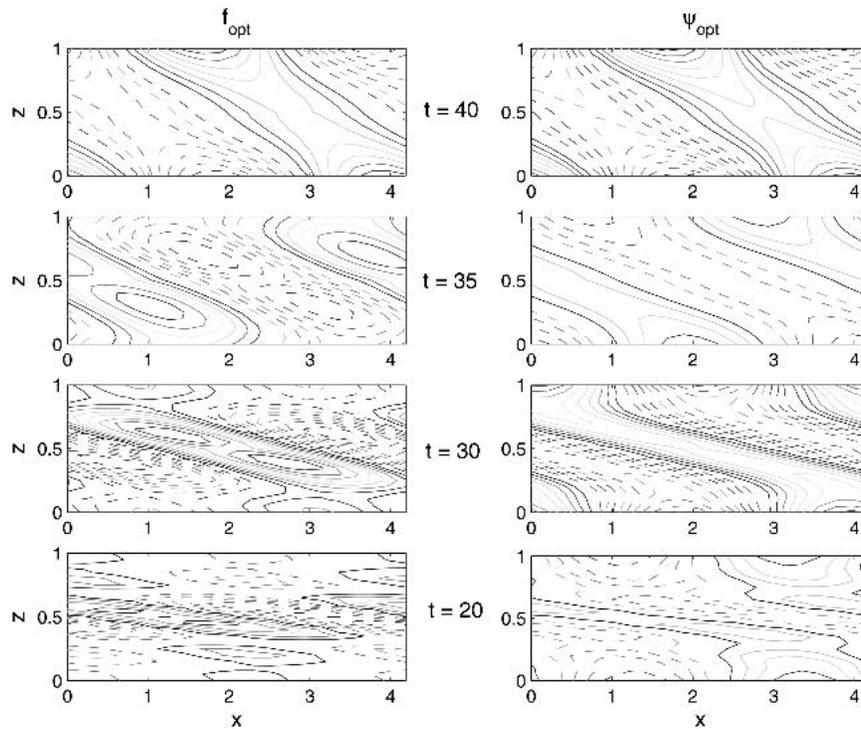

FIG. 7. For the observer-stabilized Eady model: (left) temporal evolution of the structure of the deterministic optimal forcing Re $[\mathbf{f}(z, t)e^{ikx}]$ for $T = 40$. The optimal forcing is shown from the bottom to the top at times $t = 20, 30, 35, 40$. (right) The corresponding structure of the optimal state Re $[\boldsymbol{\psi}(z, t)e^{ikx}]$. The parameters are as described in Fig. 4.
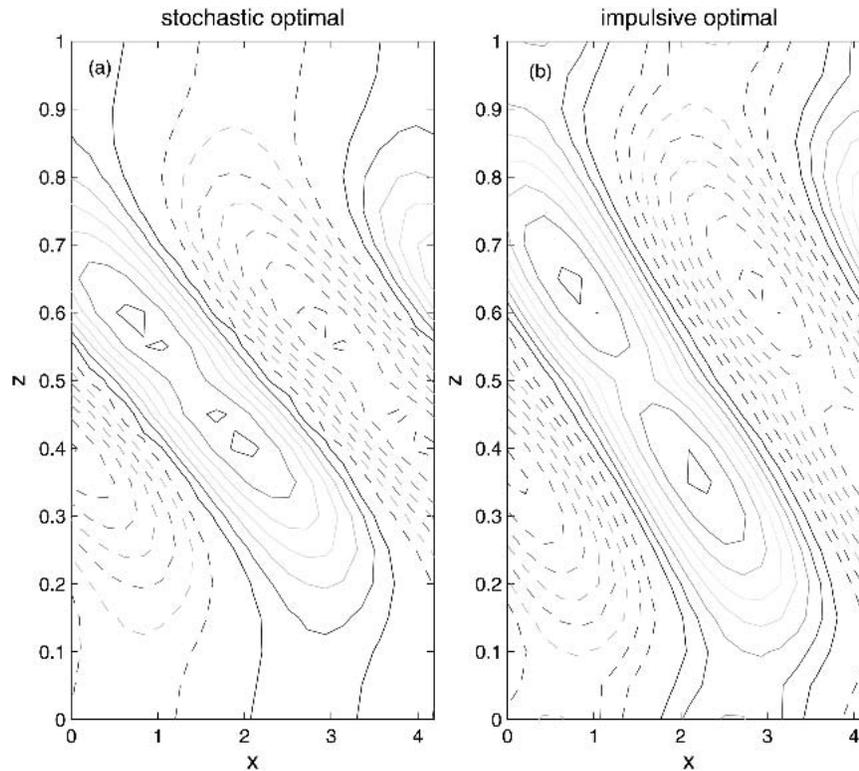
FIG. 8. For the observer-stabilized Eady model: (a) structure of the stochastic optimal that produces the greatest expected square error at $T = 40$, (b) structure of the optimal impulsive forcing that produces the greatest expected square error at $T = 40$. The optimal impulsive forcing is concentrated 7 units of time before the optimization time and has the structure of the optimal initial condition associated with the global optimal growth. The parameters are as described in Fig. 4.

and the optimal impulsive forcing (which consists of forcing at the initial time of and with the structure of the optimal initial condition that maximizes perturbation growth at $T = 20$), although formally different, in this strongly unstable example have both assumed the structure of the adjoint of the most unstable mode of the model, which is shown in the second panel from the top on the left-hand side of Fig. 12.

8. Conclusions

In this work we have studied the response of a linear dynamical system to distributed deterministic and stochastic forcing. Specific atmospheric applications of this study are to the presumed unstable forward forecast error system and the necessarily stable data assimilation system that are each continually perturbed by disturbances that degrade the forecast and assimilation, respectively. These disturbances may be physical in origin and may also arise from numerical model deficiencies. The disturbances may be deterministic or stochastic. We analyzed the effect of temporally distributed forcing on stable and unstable error systems by obtaining a spanning set of optimal distributed forcings over a

chosen interval making use of the adjoint system propagator applied to the finite-time covariance matrix eigenvectors. In the stable case modeling a data assimilation system, the optimal forcing is concentrated near the end of the interval of the optimization, while in the unstable forecast error system the optimal forcing is concentrated near the beginning of the interval. The optimals use both accumulation of forcing variance as well as variance transfer from the mean in their growth process, although transfer from the mean was dominant in the examples. The structure of the response to the optimal distributed forcing is the first eigenvector of the finite-time covariance $\mathbf{C}(T) = \int_0^T \mathbf{\Phi}(T, t)\mathbf{\Phi}^H(T, t)dt$. In the case of stochastic forcing, the structure of the forcing producing the greatest response is the finite-time stochastic optimal, which is the first eigenvector of the stochastic optimal matrix $\mathbf{B}(T) = \int_0^T \mathbf{\Phi}^H(T, t)\mathbf{\Phi}(T, t)dt$, and this forcing generally tends to have the structure of the optimal initial condition. Finally, in the case of point impulsive forcing, the greatest response is obtained by a forcing concentrated at the global optimizing time in the past and with the structure of the global optimal initial condition.

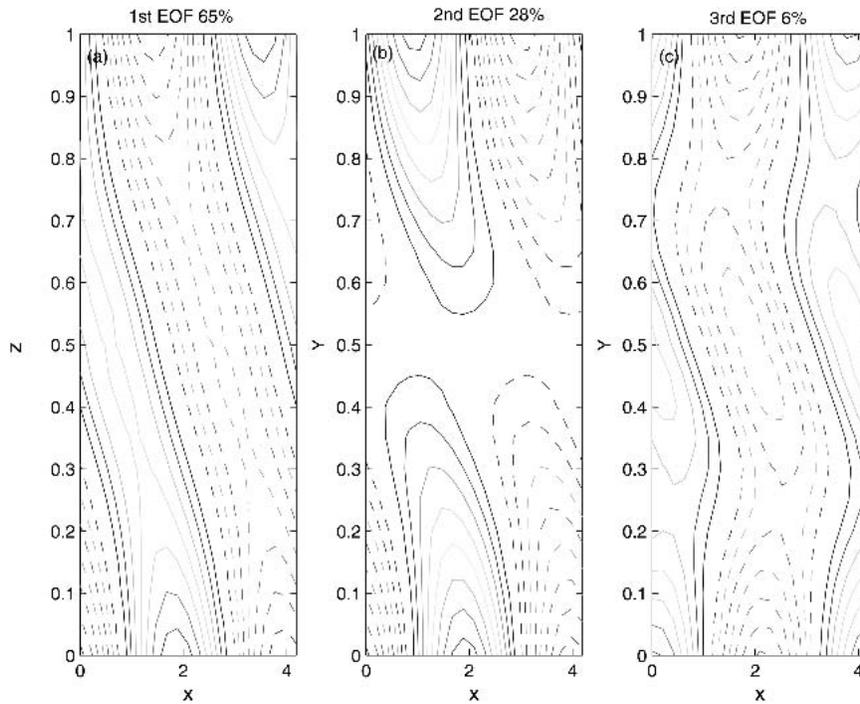Most forcings make little or no contribution to fore-

FIG. 9. Structure of the (left to right) first three eigenvectors of $\mathbf{C}(T)$ (EOFs) at $T = 40$ maintained by the stochastic optimal in the observer-stabilized Eady model. The parameters are as described in Fig. 4.
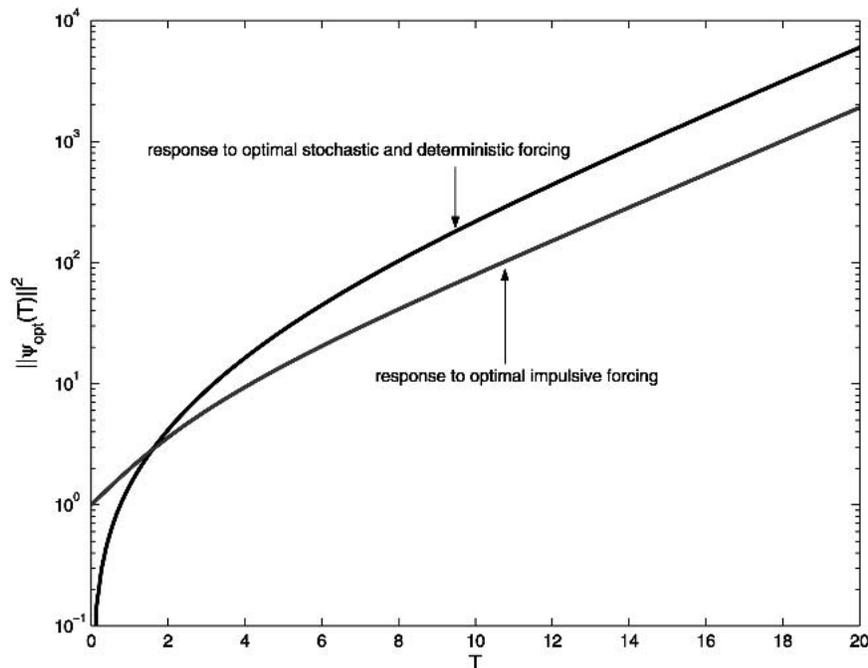

FIG. 10. For the unstable Eady forecast error model: the square norm of the state, $\|\psi_{\mathrm{opt}}(T)\|^2$, at optimizing time, T, as a function of optimizing time under deterministic forcing of unit norm, which, for this strongly unstable error model, coincides with the optimal expected square error norm under statistically stationary stochastic forcing. For comparison, the optimal square error norm under optimal impulsive forcing is also shown. The zonal wavenumber is $k = 1.5$ and the Rayleigh damping coefficient is $r = 0.15$. The forecast system is unstable with a growth rate $1/2.5$ day^{-1}.
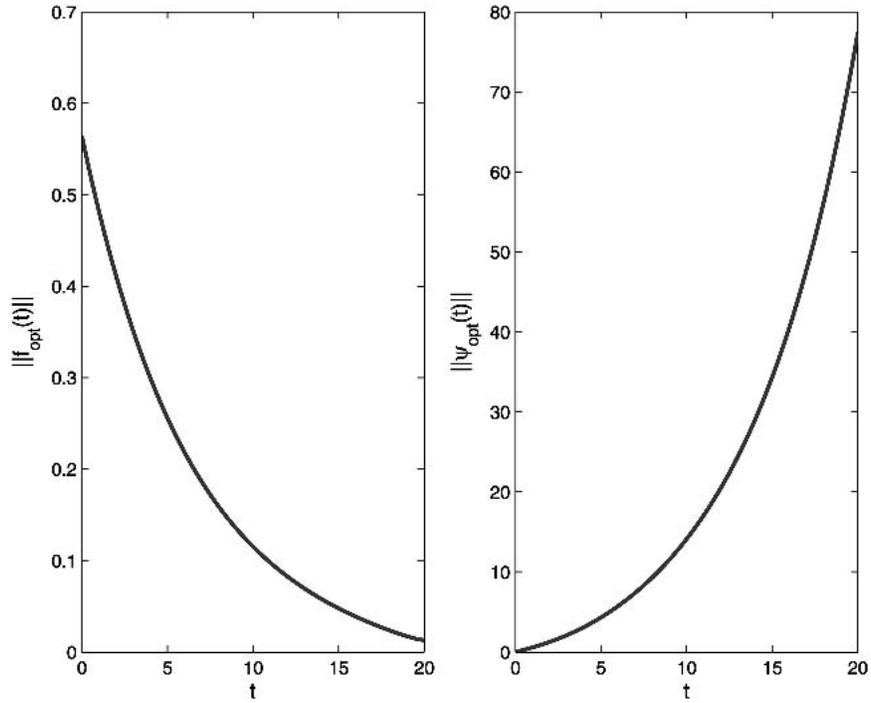
FIG. 11. For the unstable Eady forecast model example: (a) norm of the deterministic forcing that optimizes the energy of the state at $T = 20$. (b) Corresponding norm of the state $|\psi_{opt}(t)|$. The parameters are the same as described in Fig. 10.

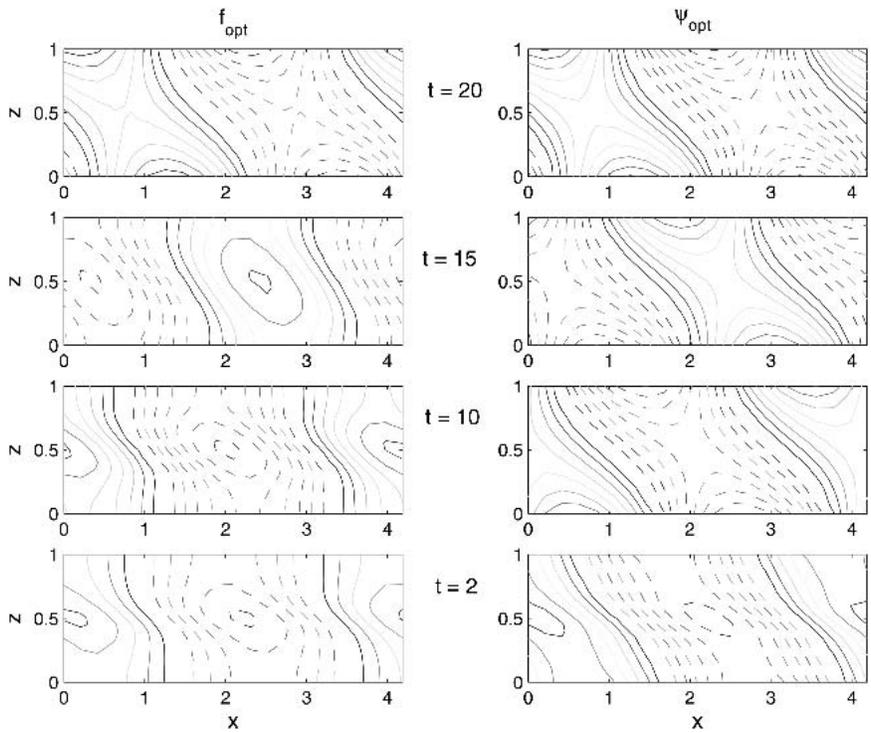FIG. 12. For the unstable Eady forecast model: (left) structure of the deterministic optimal forcing Re $[\mathbf{f}(z, t)e^{ikx}]$ that optimizes the state norm at $T = 20$. The optimal forcing is shown at times $t = 2, 10, 15, 20$. (right) the corresponding structure of the state Re $[\psi(z, t)e^{ikx}]$. The parameters are as described in Fig. 10.

cast and assimilation error, and a small subspace of forcings produce most of the error that is realized primarily in a small subspace of error states. These forcings are the optimal forcings that project on the dominant eigenvectors of the finite-time covariance matrix.

Forecast error may arise primarily from error in the initial analysis, or it may be forced by error sources distributed over the forecast interval. Initial error influence on forecast is characterized by the set of optimal perturbations, while error distributed in time is characterized by optimal forcing functions. It remains to determine the relative importance of these error sources in practical forecast situations.

Consider macroscopic cyclone formation as distinct from perturbation error growth. The cyclone may arise from forcing distributed over a given interval or from an initial disturbance at the beginning of the interval. Improved understanding of the origin of cyclones depends, in part, on a better characterization of the relative role of initial conditions and distributed forcings in the cyclogenesis process.

REFERENCES

Arnold, L., 1998: *Random Dynamical Systems.* Springer-Verlag, 586 pp.

Bamieh, B., and M. Dahleh, 2001: Energy amplification in channel flows with stochastic excitation. *Phys. Fluids,* **13,** 3258–3269.

Barkmeijer, J., T. Iversen, and T. N. Palmer, 2003: Forcing singular vectors and other sensitive model structures. *Quart. J. Roy. Meteor. Soc.,* **129,** 2401–2424.

Bewley, T. R., and S. Liu, 1998: Optimal and robust control and estimation of linear paths to transition. *J. Fluid Mech.,* **365,** 23–57.

Brockett, R. W., 1970: *Finite Dimensional Linear Systems.* John Wiley and Sons, 244 pp.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.,* **52,** 1434–1456.

Butler, K. M., and B. F. Farrell, 1992: Three dimensional optimal perturbations in viscous shear flow. *Phys. Fluids,* **4A,** 1637–1650.

D'Andrea, F., and R. Vautard, 2000: Reducing systematic errors by empirically correcting model errors. *Tellus,* **52A,** 21–41.

Dullerud, G. E., and F. Paganini, 2000: *A Course in Robust Control Theory: A Convex Approach.* Springer-Verlag, 440 pp.

Eady, E. A., 1949: Long waves and cyclone waves. *Tellus,* **1,** 33–52.

Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.,* **39,** 1663–1686.

——, 1988a: Optimal excitation of perturbations in viscous shear flow. *Phys. Fluids,* **31,** 2093–2102.

——, 1988b: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.,* **45,** 163–172.

——, 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.,* **46,** 1193–1206.

——, 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.,* **47,** 2409–2416.

——, and P. J. Ioannou, 1993a: Stochastic dynamics of baroclinic waves. *J. Atmos. Sci.,* **50,** 4044–4057.

——, and ——, 1993b: Stochastic forcing of the linearized Navier–Stokes equations. *Phys. Fluids,* **5A,** 2600–2609.

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

——, and ——, 1996b: Generalized stability. Part II: Nonautonomous operators. *J. Atmos. Sci.,* **53,** 2041–2053.

——, and ——, 1998: Turbulence suppression by active control. *Phys. Fluids,* **8,** 1257–1268.

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

——, and ——, 2001a: Accurate low-dimensional approximation of the linear dynamics of fluid flow. *J. Atmos. Sci.,* **58,** 2771–2789.

——, and ——, 2001b: State estimation using a reduced-order Kalman filter. *J. Atmos. Sci.,* **58,** 3666–3680.

——, and ——, 2002a: Perturbation growth and structure in uncertain flows. Part II. *J. Atmos. Sci.,* **59,** 2647–2664.

——, and ——, 2002b: Optimal perturbation of uncertain systems. *Stochastic Dyn.,* **2,** 395–402.

Glover, K., 1984: An optimal Hankel-norm approximation of linear multivariable systems and their $\mathbf{L}^\infty$-error bounds. *Int. J. Control,* **39,** 1115–1193.

Högberg, M., T. R. Bewley, and D. S. Henningson, 2003a: Linear feedback control and estimation of transition in plane channel flow. *J. Fluid Mech.,* **481,** 149–175.

——, ——, and ——, 2003b: Relaminarization of $Re_\tau = 100$ turbulence using linear state-feedback control. *Phys. Fluids,* **15,** 3572–3575.

Jovanović, M. R., and B. Bamieh, 2001: Modeling flow statistics using the linearized Navier–Stokes equations. *Proc. 40th IEEE Conf. on Decision and Control,* Orlando, FL, IEEE, 4944–4949.

Kalman, R. E., 1964: When is a control system optimal? *J. Basic Eng.,* **83D,** 95–108.

Kim, J., 2003: Control of turbulence. *Phys. Fluids,* **15,** 1093–1105.

Kleeman, R., and A. M. Moore, 1997: A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.,* **54,** 753–767.

Lacarra, J., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus,* **40A,** 81–95.

Luenberger, D. G., 1971: An introduction to observers. *IEEE Trans. Autom. Control.,* **16,** 596–602.

Molteni, F., and T. N. Palmer, 1993: Predictability and finite-time instability of the Northern Hemisphere winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.,* **55,** 633–653.

Reddy, S., and D. S. Henningson, 1993: Energy growth in viscous channel flows. *J. Fluid Mech.,* **252,** 209–238.

Zhou, K., and J. C. Doyle, 1998: *Essentials of Robust Control.* Prentice Hall, 411 pp.