ENSEMBLE DYNAMICS AND BRED VECTORS

By

Nusret Balci

Anna L. Mazzucato

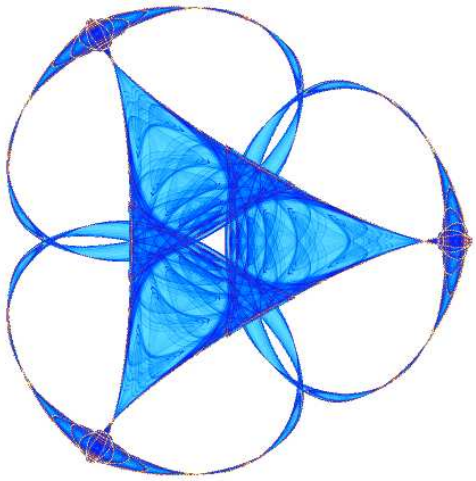
Juan M. Restrepo

and

George R. Sell

IMA Preprint Series # 2353

(November 2010)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

Ensemble Dynamics and Bred Vectors

Nusret Balci [1], Anna L. Mazzucato [2], Juan M. Restrepo [3*] & George R. Sell [4]

[1] *Institute for Mathematics and its Applications*

University of Minnesota, Minneapolis MN 55455 U.S.A.

[2] *Department of Mathematics*

The Pennsylvania State University, University Park, PA 16802 U.S.A.

[3] *Department of Mathematics and Department of Physics*

University of Arizona Tucson, AZ 85721 U.S.A.

[4] *School of Mathematics*

University of Minnesota, Minneapolis MN 55455 U.S.A.

November 14, 2010

[*] *Corresponding Author:* restrepo@math.arizona.edu

Abstract

We propose a variant of the Bred Vector (BV) algorithm, orginally introduced by Z. Toth and E. Kalnay (*Bulletin of the American Meteorological Society* 74:2317 2330, 1993) to assess the sensitivity of model outputs to changes in initial conditions for weather forecasting. The new algorithm, which we call the *Ensemble Bred Vector* or EBV, is based on collective dynamics in an essential way. As such, it features distinctive qualities compared to the classical breeding algorithm. By construction, the EBV produces one or more dominant vectors, and is less prone to spurious results than the BV algorithm. It retains the attractive features of the BV with regard to being able to handle legacy codes with minimal additional coding.

We investigate the performance of EBV, comparing it to the BV algorithm as well as the finite-time Lyapunov Vectors. With the help of a continuous-time adaptation of the algorithms we give a theoretical justification to the observed fact that the vectors produced by BV, EBV, and the finite-time Lyapunov vectors are similar to each other for small amplitudes. The continuum theory is also used to establish the relationship between the BV and general directional derivatives.

Numerical comparisons of BV and EBV for the 3-equation Lorenz model and for a forced, dissipative partial differential equation that arises in the thermohaline circulation demonstrate that the EBV yields a size-ordered description of the perturbation field, and is more robust than the BV in the higher nonlinear regime. The EBV yields insight into the fractal structure of the Lorenz attractor, and of the inertial manifold for the Cahn-Hilliard-type partial differential equation.

Keywords: Bred vectors, Lyapunov vectors, sensitivity, dynamic stability, Cahn-Hilliard, Lorenz.

PACS 94.05sx, 92.60.Ry, 92.60.Wc, 93.65.+e, 92.70.-j, 02.50.-r

1 Introduction

Central to weather prediction is the analysis of the sensitivity of a physical or computer-coded model to initial conditions. Model sensitivity to parameters is also important in model inter-comparisons, and in the process of obtaining a better understanding of the role played by these parameters in model outcomes.

Sensitivity and predictability are often interwined in the context of weather prediction and have been the subject of extensive research (see Buizza et al. (1993) and references contained therein.) These are not exclusively weather-related issues and thus geophysical fluid dynamics will often mine other physical, computational and mathematical disciplines, for ideas with which to assess dynamic sensitivity. Practical sensitivity methodologies must contend with the evolution and dynamics of highly coupled, complex, high-dimensional systems, riddled with subscale parameterizations and empirical relations, which are the norm in large-scale climate and meteorology models.

In forecasting, which is increasingly common in climate but has been a primary aspect of meteorology, causality is an obvious constraint on sensitivity methodologies and thus forward sensitivity is a more common tool than backward sensitivity. (See Errico and Lagland (1999a,b), and references therein, for a comparison between forward and backward sensitivity analysis. We refer the reader to the book by Trefethen and Embree (2005) for a discussion in the context of numerical linear algebra).

We will introduce a variant Bred Vector (BV) algorithm, which we call the *Ensemble Bred Vector* algorithm (EBV). Bred Vectors have been proposed as a mean to assess the forward sensitivity in weather and climate models. The

BV algorithm was first introduced by Toth and Kalnay (1993). While several modifications to the original algorithm have been presented since then (we review some of the relevant literature on BV in Section 2), we will focus on the original algorithm.

The BV algorithm is a finite-time, forward sensitivity methodology which, in addition to being useful in characterizing model sensitivity to initial conditions, has been proposed as a means to produce a reduced-rank representation of the background error in data assimilation and forecast error-covariance approximations (see Corazza et al. (2003), for example),

In the BV algorithm, one follows an initial condition of the time-discrete nonlinear system, along with a family of nearby solutions. (Since this algorithm is used to sample the error space, an ensemble of initial perturbations is bred simultaneously.) The perturbations at the initial time are fixed with a common small amplitude ϵ. After each cycle, the outcome of the perturbations is rescaled to the same amplitude ϵ. For the BV algorithm, the rescaling of each perturbation is independent of the others, and there is no mechanism for comparing the dynamics of nearby perturbations. The new variation that we propose here, the EBV algorithm, differs from the BV algorithm, in the rescaling rule. In particular, for the EBV those perturbations that are not the same size as the largest perturbation, play a reduced role after the rescaling. Thus the rescaling used in the EBV algorithm serves us better in separating the dominant dynamics. Therefore, the EBV algorithm is linked to the ensemble dynamics of the underlying non-linear model, hence its name. Even when all initial perturbations in the BV algorithm are chosen of the same size initially, the outcomes of the BV and EBV will be generally different. Both these algorithms are discussed in Section 2.

A more familiar and more complete strategy for assessing sensitivity of model outcomes to perturbations of initial conditions, along solution paths, is to compute the finite-time singular values of the tangent linear model (see Molteni and Palmer (1993) and Palmer et al. (1998)). In addition to the complexities of obtaining the tangent linear model and its adjoint for a large scale model, the SVD itself can be exceedingly demanding computationally. The forward/backward finite-time Lyapunov vector calculation can yield a wealth of sensitivity information; in some situations it can provide as much information as the SVD itself. For the finite-time Lyapunov calculation you still need to have a tangent linear model. Wolfe and Samelson (2007) propose a way to use a partial computation of the SVD to obtain Lyapunov vectors. The BV is considered an algorithmic alternative that approximates the outcomes of a finite-time Lyapunov vector calculation. One of the key advantages of the BV algorithm, retained by the new EBV algorithm, is its ease of use, which does *not* require coding the tangent linear model.

The Lyapunov Vectors for a nonlinear problem with a global attractor arise in the study of solutions of the associated tangent linear model. Except for the trivial case where the attractor contains exactly one point, the tangent linear equation is always a non-autonomous equation, and hence "nonautonomous" techniques, such as those developed in Lyapunov (1992), must be used. The Multiplicative Ergodic Theorem (MET for short), which exploits the family of invariant probability measures on the attractor, is a relatively new tool bag for the study of the longtime properties of the related tangent linear model. Some researchers have used the MET in related studies, for example, Wolfe and Samelson (2007). As shown by Johnson et al. (1987) by combining the MET with the theory of exponential dichotomies

one can describe the exponential separation between some solutions of the tangent linear model with different exponential growth rates.

It was observed by Toth & Kalnay *op. cit.* (see also Toth and Kalnay (1997)) in several experimental runs that BVs resemble the *leading* finite-time Lyapunov vectors. In order to make a more quantitative comparison between Lyapunov vectors and BVs, we study, in Section 3, the continuum-in-time limits of the BV and EBV algorithms and show how they are related to the tangent linear model dynamics.

The BV algorithm and its variants are discrete-time algorithms based on finite-dimensional approximations of weather models, obtained either by mode projection as in the Lorenz (1963) system, or by spatial discretization of partial differential equations (PDEs), as in the CY92 model, which will be described below. Section 3 also contains a further discussion of some of the desirable and interesting features of the new EBV. For instance, there is a natural ordering of the ensemble members of an EBV, and as we will show, it is possible to observe perturbations with smaller sizes than the dominant one but with very strong growth.

One of the main goals in Section 4 is to make comparisons between the outcomes of the BV algorithm and the EBV, and where appropriate to the finite time Lyapunov vector algorithm. We will see that in some metrics, the BV and the EBV yield comparable results, while in other metrics, the EBV outcomes are clearly superior.

We consider two models. The first is the familiar Lorenz63 model introduced by Lorenz (1963). It has a well-known global attractor. The second is a nonlinear forced and dissipative partial differential equation of the Cahn-Hilliard type. The equation forms part of a model proposed by Cessi and

Young (1992) of the oceanic thermohaline circulation. We will denote the equation associated with this model the CY92 in this study. As it turns out, it is a good example of the typical climate-related model dynamics. However, the CY92 is special, since it has an inertial manifold. Consequently, the long-time dynamics of this partial differential equation is completely contained in the attractor of a finite dimensional ordinary differential equation. As we will show the EBV yields insights into the structure of the attractor of the Lorenz63, and of the inertial manifold to the CY92.

The BVs are sensitive to the amplitude and frequency content of the initial perturbation. In contrast, the outcome of the EBV algorithm shows a clear hierarchy among its members, and the first few members already generate an unambiguous characterization of the perturbation field at both large and small amplitudes. In the nonlinear regime, however, the EBV will be shown in Section 4 to be less likely to produce spurious results than the BV. In Section 6 we will address implementation issues of the EBV.

2 Bred Vector Algorithms and Finite-Time Lyapunov Vectors

In this section we present the definitions and methodology for computing the two Bred Vector algorithms, the BV and the EBV. We also review the basic theory of Lyapunov Vectors, Lyapunov exponents, and their finite-counterparts. We will compare these different tools later in the paper.

For the BV algorithm, we use the one originally proposed by Toth and Kalnay (1993). BV is purely algorithmic. It is "equation-free" and thus with

additional minimal computer coding it can handle legacy code representing even extremely complex models. Most alternatives for obtaining estimates of forward sensitivity will involve non-trivial additional coding. For example, in order to obtain the finite-time Lyapunov vectors and exponents, one needs to derive and make use of the tangent linear model.

Several articles in the literature have addressed applications of the BV algorithm in weather modeling. See, in particular Toth and Kalnay (1993), Toth and Kalnay (1997), Kalnay (2003), and reference therein. For a comparison of the BV algorithm and other methods, such as Monte-Carlo perturbed observations, we refer for example to Cheung (2001); Gneiting and Raftery (2005); Hansen and Smith (2000); Trevisan and Legnani (1995); Wei and Toth (2003). For an application of the BV algorithm to ensemble Kalman filters see Wang and Bishop (2003). A few articles discuss the BV algorithms and their variants in a more theoretical context. Deremble et al. (2009) make use of singular vectors, as well as a variation of the concept of a BV, to study the regime predictability in some reduced weather models with two degrees of freedom. Primo et al. (2008) discuss universality in a variant BV algorithm, where rescaling is by the geometric mean (see also Hallerberg et al. (2010)).

We consider the following initial value problem

$$\begin{aligned} \frac{dy}{dt} &= G(y), \quad t > 0, \\ y(0) &= y_0, \end{aligned} \tag{1}$$

where t represent time and $G = G(y)$ is a map that has at least a bounded gradient. Since our main applications involve autonomous differential equations, we assume that G does not explicitly depend on time. The basic theory we present here has a routine extension to non-autonomous problems.

The vector $y = y(t)$ can live in a finite- or infinite-dimensional normed

linear space. In the former case, (1) is an (autonomous) system of ordinary differential equations, while in the latter case, (1) is an (autonomous) system of partial differential equations, modeling a time dependent, spatially extended system. For systems of partial differential equations, we assume that either periodic boundary conditions or non-flux boundary conditions are prescribed. The use of other boundary conditions may lead to a related theory, but we do not address the issue here.

If the system contains evolution partial differential equations, the system, along with the boundary conditions, are discretized in space or projected onto a finite-dimensional space compatible with the boundary conditions. Consequently, we usually assume that (1) is a system of ordinary differential equations of dimension K, which may be large. Since most large-scale weather and climate circulation models presently use explicit-in-time integrators, we will focus on numerical models of this type.

2.1 Bred Vector Algorithms: BV and EBV

We first review the BV algorithm, in order to compare it with the new EBV algorithm, which we introduce afterwards.

We treat the function $y = y(t)$ as a given (continuous time) solution of (1). We then turn to an approximate solution $Y_n = Y(t_n)$, which is defined on the time grid: t_n, for $n = 0, 1, 2, \cdots$, where $t_{n+1} = t_n + \delta t_n$, for $n = 0, 1, 2, \ldots$. We set $t_0 = 0$. We assume that $\delta t_n = \delta t$ is positive and small, and that it does not depend on n, for $n \geq 0$.

For the autonomous case, the initial value problem approximated using an explicit numerical integration scheme, leads to consideration of the difference

equation

$$\begin{aligned} Y_{n+1} &= M(Y_n, \delta t), \quad n = 0, 1, 2, .., \\ Y_0 &= Y(0). \end{aligned} \tag{2}$$

where $Y(t_n)$ is an approximation of the vector $y(t)$ at $t = t_n$, while Y_0 an approximation of the initial condition y_0. The points Y_n are in the K-dimensional Euclidean space $\mathbb{R}^K$ and $M = M(Y, \delta t)$ is the discrete-time solution operator on $\mathbb{R}^K$ generated by the ordinary differential equation (1). We denote the components of a vector Y in $\mathbb{R}^K$ by Y^j, $j = 1, \ldots, K$.

The BV algorithm generates a sequence $\{(Y_n, \delta\mathcal{Y}_n)\}$, for $n = 1, 2 \ldots$, where $(Y_n, \delta\mathcal{Y}_n - \delta\mathcal{Y}_{n-1})$ is an approximation of

$$(y(t), v(t)), \qquad \text{at } t = t_n,$$

and $v(t)$ is a vector positively proportional to the projection of the directional derivative of the map G in (1) in the direction of $\delta\mathcal{Y}_n$ on a plane with normal $\delta\mathcal{Y}_n$. Thus, if we set $w(t_n) = DG\,|_{(y(t_n))}\,[\delta\mathcal{Y}_n]$, then $\delta\mathcal{Y}_n - \delta\mathcal{Y}_{n-1}$ is approximately and positively proportional the vector $w(t_n) - (w(t_n) \cdot (\delta\mathcal{Y}_n))\delta\mathcal{Y}_n/(\|\delta\mathcal{Y}_n\|^2)$. The initial datum for this algorithm is $(Y_0, \delta\mathcal{Y}_0)$, where Y_0 is the initial *base point* (i.e., an approximation of y_0), and $\delta\mathcal{Y}_0 \neq 0$ is the initial *base vector* (*i.e.*, an initial choice for the directional derivative at y_0).

For $n = 0, 1, 2, \ldots, N-1$, given the nth base point Y_n, and base vector $\delta\mathcal{Y}_n$,

1. Y_{n+1} denotes the $(n+1)^{\text{st}}$ base point, and it is determined by (2);

2. $\delta\mathcal{Y}_{n+1}$, the $(n+1)^{\text{st}}$ base perturbation vector, is given by

$$\delta Y_{n+1} = M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t), \quad \text{and} \quad \delta\mathcal{Y}_{n+1} = R_{n+1}\,\delta Y_{n+1}, \tag{3}$$

where R_{n+1} is a rescaling rule.

The set $\{\delta\mathcal{Y}_n\}_{n=0}^{N-1}$ gives the sequence BV(t_n). The choice of N is completely arbitrary and the dynamics of the model in question is often used to suggest an appropriate value. Nevertheless, it is taken to be the smallest value for which change in the sequence BV(t_n) is acceptably small.

One rescaling rule, proposed by Toth and Kalnay, *op. cit.*, consists of rescaling the t_n base vector to the previous one: $\|\delta\mathcal{Y}_{n+1}\| = R_{n+1}\,\|\delta Y_{n+1}\| = \|\delta\mathcal{Y}_n\|$, for $n \in [0, N-1]$. Equivalently,

$$R_{n+1}(Y_0) := \frac{\|\delta\mathcal{Y}_n\|}{\|\delta Y_{n+1}\|} \;=\; \frac{\|\delta\mathcal{Y}_0\|}{\|\delta Y_{n+1}\|}, \tag{4}$$

so that $R_{n+1} = R_{n+1}(Y_0)$ depends on the initial base point Y_0, as well as the initial base vector $\delta\mathcal{Y}_0$. An alternate rescaling rule consists in rescaling periodically, at $t_{1+m\,k}$, where k is an integer $k \geq 2$, and $m = 0, 1, 2, \cdots$. In this case, one uses the rule (4) when $n = mk$, and

$$R_{n+1}(Y_0) \;=\; 1, \qquad \text{for } mk < n < (m+1)k. \tag{5}$$

We do not use this alternate rule in this paper. (For a discussion of rescaling time and regime predictability in some reduced models, see Deremble et al. (2009) and references therein).

The *Ensemble Bred Vector* or EBV for short, contains intrinsically the notion of an ensemble of data and outcomes. In present-day weather forecasting several concurrent and independent runs of the BV algorithm are usually performed.

Instead of considering a single base vector $\delta\mathcal{Y}_0$, as one does for the BV, we now consider a family of initial base vectors $\delta\mathcal{Y}_0(\iota)$, where ι is an element of an index set $\mathbb{I}$. We require that the norms of this family of base vectors

satisfy $\|\delta\mathcal{Y}_0(\iota)\| = \epsilon$, for all $\iota \in \mathbb{I}$. The amplitude ϵ is assumed small, but how small ϵ must be taken is typically model dependent. Instead of using equation (3), where the scaling rule $R_{n+1}(Y_0(\iota))$ depends on index ι, we use a uniform scaling $\boldsymbol{R}^{\min}_{\boldsymbol{n+1}}$, which is the same for all $\iota \in \mathbb{I}$. In particular, we replace equation (3) with

$$\begin{aligned} \delta Y_{n+1}(\iota) &= M(Y_n + \delta\mathcal{Y}_n(\iota), \delta t) - M(Y_n, \delta t) \\ \delta\mathcal{Y}_{n+1}(\iota) &= \boldsymbol{R}^{\min}_{\boldsymbol{n+1}}\, \delta Y_{n+1}(\iota), \end{aligned} \tag{6}$$

for all $\iota \in \mathbb{I}$, where

$$\boldsymbol{R}^{\min}_{\boldsymbol{n+1}} = \epsilon \left[\max_{\iota \in \mathbb{I}} (\|\delta Y_{n+1}(\iota)\|) \right]^{-1}. \tag{7}$$

We define the EBV to be the ensemble $\delta\mathcal{Y}_n(\iota)$, for $\iota \in \mathbb{I}$, as described above, that satisfies (6) and (7). (Alternatively, a periodic rescaling rule utilizing (5) above can be employed.)

The crucial difference between the BV and EBV algorithms is that, even when the BV is run concurrently over an ensemble of initial data, the outcome of the algorithm for each given datum does not depend on the other members of the ensemble. In contrast, the evolution of the ensemble members is interdependent in the EBV. Nevertheless, by construction, the EBV should exhibit similar behavior to the BV at small amplitudes. Indeed, one of the design principles for the EBV algorithm is to reinforce this aspect by providing it with a built-in acceleration mechanism.

The time step used to compute the base trajectory and the time intervals between normalizations need not be the same. This is the case for both the BV algorithm as well as EBV. Both the BV algorithm and EBV outcomes depend on the choice of vector norm used to define the rescaling rule in either (4) or (7). In principle, since we restrict to finite K-dimensional systems

herein, all norms are equivalent. In practice however, this finite-dimensional system arises from a discretization of a continuum infinite-dimensional system. Thus the dimension K can be very large and typically the constants expressing the norm equivalence may grow unboundedly with K. Hence, the choice of norm used has an impact on the implementability and performance of these algorithms. In Riviére et al. (2008), it was shown that the outcome of the BV algorithm may depend strongly on the choice of norm. This is actually a consequence of the non-negligible nonlinear effects in the system. In Section 6, we will elaborate further on the issue of norm dependence.

As already discussed, several alternatives to the BV algorithm have been proposed and employed in the literature. BVs have been viewed as non-linear analogs of finite-time Lyapunov vectors. Similarly nonlinear analogs of singular vectors have been proposed, for instance conditional nonlinear optimal perturbations proposed by Mu and Jiang (2008) and non-linear singular vectors proposed by Riviére et al. (2008), although they entail computationally expensive optimization.

2.2 Lyapunov Vectors

As discussed in the Introduction, we will be applying the BV and EBV algorithms to two models consisting of systems of (nonlinear) autonomous ODEs of the form (1). The second model arises from the discretization of a PDE. In each model, there is a global compact attractor $\mathfrak{A}$, which is an subset of $\mathbb{R}^K$ invariant for the time evolution of the system. (We refer the reader to Chapter 2 in Sell and You (2002) for more information on attractors and global attractors.) We will let θ denote a typical point in the attractor

$\mathfrak{A}$, and we will let $\theta \cdot t = y(t)$ denote the unique solution of (1) that satisfies $y(0) = \theta$. Since $\mathfrak{A}$ is invariant, one has $\theta \cdot t \in \mathfrak{A}$, for all $t \in \mathbb{R}$.

In order to study the sensitivity with respect to initial conditions on the attractor, the Tangent Linear Model is used, defined as follows. Let

$$\partial_t x = A(\theta \cdot t)\, x, \tag{8a}$$

$$x(0) = x_0 \in \mathbb{R}^K, \tag{8b}$$

be given, where $A(y) = DG(y)$ is the Jacobian matrix of G. Hence $A = A(\theta \cdot t)$ is the linearization of (1) along the solution $\theta \cdot t$. We observe that, even if G does not explicitly depend on time, (8a) is generally nonautonomous, since $\theta \cdot t$ changes with time.

We let $U(\theta, t)$, denote the solution operator of (8a), which takes the initial data to the solution at time t, so that $U(\theta, t)x_0$ is the solution of the initial value problem for (8). Such an operator is well defined by the uniqueness of solutions to the problem (8). Uniqueness of solutions also readily implies the cocycle identity:

$$U(\theta, \tau + t) = U(\theta \cdot \tau, t)\, U(\theta, \tau), \qquad \text{for all } \theta \in \mathfrak{A} \text{ and all } \tau,\, t \in \mathbb{R}. \tag{9}$$

Next we consider a family of mappings $\Pi = \Pi(t)$, which are defined for $t \in \mathbb{R}$ by the relation

$$\Pi(t)(\theta, x_0) := (\theta \cdot t, U(\theta, t)x_0), \qquad \text{for } (\theta, x_0) \in \mathfrak{A} \times \mathbb{R}^K. \tag{10}$$

We note that $\Pi(t)$ maps $\mathfrak{A} \times \mathbb{R}^K$ into itself, for each $t \in \mathbb{R}$; it is jointly continuous in (t, θ, x_0); it satisfies $\Pi(0)(\theta, x_0) = (\theta, x_0)$, (i.e., $\Pi(0) = I$, the identity operator; as well as the evolution property:

$$\Pi(\tau + t) = \Pi(t)\, \Pi(\tau), \qquad \text{for all } \tau,\, t \in \mathbb{R}. \tag{11}$$

This group property is a consequence of the cocycle identity (9). In fact, the group property is equivalent to the cocycle identity. Note that the θ-component of Π does *not* depend on the x_0-component. Thus Π is called a *skew product* flow. Since Π is linear in x_0, it is sometimes called a *linear* skew product flow. In summary, the Tangent Linear Equation over the attractor $\mathfrak{A}$ generates a linear skew product flow. (For more information on the theory of skew product flows in the context of non-autonomous dynamics, see the multiple works of Sacker and Sell, for example Sacker and Sell (1977), Sacker and Sell (1978) and Sacker and Sell (1980).) The dynamics of Π are crucial for understanding the sensitivity and predictability of the underlying model.

In his opus magnum, which was published in Ukraine in 1892, Lyapunov presented his theory of stability for finite-dimensional ordinary differential equations. This work includes his study of the non-autonomous linear problem (8), see Lyapunov (1992). One of Lyapunov's goals was to develop an analogue of the well-known eigenvalue-eigenvector theory, for the solutions of the autonomous problem, to the study of solutions of general non-autonomous equation (8).

The Lyapunov exponent associated to $\theta \in \mathfrak{A}$ and $x_0 \neq 0$ in $\mathbb{R}^K$ is defined by

$$\lambda(\theta, x_0) = \lambda(x_0) := \limsup_{t\to\infty} \frac{1}{t} \log(\|U(\theta, t)x_0\|), \tag{12}$$

and the vector x_0 is referred to as a Lyapunov Vector associated with $\lambda = \lambda(x_0)$.

In this definition of the Lyapunov exponent $\lambda(\theta, x_0)$, we used the limit $\limsup_{t\to\infty}$. One might have used instead three other limits: $\limsup_{t\to-\infty}$, $\liminf_{t\to\infty}$, and $\limsup_{t\to-\infty}$. When all these four limits are equal, $\lambda(\theta, x_0)$

is said to be a *strong Lyapunov exponent*, and one has

$$\lambda(\theta, x_0) \;=\; \lim_{t\to\infty} \frac{1}{t} \log(\|U(\theta,t)x_0\|) \;=\; \lim_{t\to-\infty} \frac{1}{t} \log(\|U(\theta,t)x_0\|). \tag{13}$$

In Section 5, we give an answer to the question: When is a given Lyapunov exponent a strong Lyapunov exponent?

One should note that if $\lambda(x_0)$ satisfies (12), then $\lambda(\alpha\, x_0) = \lambda(x_0)$, for any real number $\alpha \neq 0$, i.e., $\lambda(\alpha\, x_0) = \lambda(x_0)$. Thus a Lyapunov vector is not an isolated vector, rather it spawns a line of Lyapunov vectors (through the origin) in $\mathbb{R}^K$. That is to say, a Lyapunov vector is a *point* in $\mathcal{P}^{K-1}$, the $(K-1)$-dimensional projective space. For a given vector $v \neq 0$ in $\mathbb{R}^K$, we will use $[v] \in \mathcal{P}^{K-1}$ to denote the unique line in $\mathbb{R}^K$ that contains v. (Note that $[v] = [-v]$.) Conversely, when one maps a line $[u]$ in $\mathcal{P}^{K-1}$ to a vector $u \in \mathbb{R}^K$, we require that the pre-image u lie on the line $[u]$ and that $\|u\| = 1$. One should note that $\mathcal{P}^{K-1}$ is a metric space. We will let $d([u],[v]) = \min \|u \pm v\|$ denote the induced metric on $\mathcal{P}^{K-1}$. We note that

$$\lambda(\theta, x_1 + x_2) \;\leq\; \max\{\lambda(\theta, x_1), \lambda(\theta, x_2)\}. \tag{14}$$

As we now show, for each $\theta \in \mathfrak{A}$, the finite dimensional problem (8), admits only a finite number of Lyapunov Exponents ordered as

$$\lambda_k < \cdots < \lambda_2 < \lambda_1, \tag{15}$$

where $k = k(\theta)$, $\lambda_i = \lambda(\theta, x_i)$, and $x_i \neq 0$ is a Lyapunov vector associated with λ_i, for $1 \leq i \leq k$. Since the Lyapunov exponents are distinct, it follows that, for any choice of Lyapunov vectors, x_i, the solutions $\{U(\theta,t)\, x_i \,:\, 1 \leq i \leq k\}$ are linearly independent. Consequently, it follows that $k \leq K$. The collection of exponents λ_i, which is enumerated in (15) and which depends

on θ, is called the *Lyapunov spectrum*. For example, if $A(t) = A_0$ is an autonomous matrix, then the solution operator forms a group, which we will denote by $U_A(t) := e^{t\,A_0}$. If A_0 is in addition normal, then $U_A(t) = C\,e^{tD}\,C^{-1}$, where C is invertible and D is a diagonal matrix. In this case, the Lyapunov Spectrum of $A(t)$ consists of all real numbers λ that satisfy $\lambda = \text{Re}\,\nu$, where ν is an eigenvalue of A_0, i.e., ν is a diagonal entry in the matrix D.

By using the relation (15), for each $\theta \in \mathfrak{A}$, one can find a basis of vectors

$$\mathbb{B} = \mathbb{B}(\theta) = \mathbb{B}_k \cup \mathbb{B}_{k-1} \cup \cdots \cup \mathbb{B}_2 \cup \mathbb{B}_1, \tag{16}$$

where $x \in \text{Span}(\mathbb{B}_j) \setminus \{0\} \implies \lambda(\theta, x) = \lambda_j$, for $k(\theta) \geq j \geq 1$. Furthermore, the linear subspaces $V_j = \text{Span}(\mathbb{B}_j)$, for $k \geq j \geq 1$ satisfy $\mathbb{R}^K = V_k + V_{k-1} + \cdots + V_2 + V_1$. The *multiplicity* of the j^{th} exponent, $\text{mult}(\lambda_j)$, is defined as the dimension, $\text{mult}(\lambda_j) = \dim(V_j)$.

The study of the properties of the Lyapunov Exponents and the associated Lyapunov vectors has been an ongoing task, *ab initio*. There are two milestones in this study which are of special interest here. The first is the contribution of Perron in 1930, see Perron (1930a,b), who showed that one can represent the solution operator $U(\theta, t)$ as a triangular matrix. By taking appropriate limits (as $t \to \infty$), of the diagonal entries of this triangular matrix representation, one obtains the Lyapunov Spectrum.

The second milestone is the discovery of the Multiplicative Ergodic Theorem (MET) in 1968, when the proofs of both Oseledec and Millionščikov were published, see Oseledec (1968) and Millionščikov (1968). A more detailed description of the MET and its role in this study of the sensitivity with

respect to initial conditions, appears in Section 5. Suffice it to say here, that the MET is based on a study of the Invariant Probability Measures (IPMs) on the attractor $\mathfrak{A}$. Among other things, it is correct to say that, with probability $= 1$ the Lyapunov exponents are strong Lyapunov exponents.

Since the solution operator $U(\theta, t)$ of the linear problem (8) maps lines in $\mathbb{R}^K$ onto lines, one can use this operator to define a related flow $\Sigma(\theta, t)$ on $\mathcal{P}^{K-1}$ by means of the relation

$$\Sigma(\theta, t)\,[u] \;=\; [U(\theta, t)\,u], \qquad \text{for } [u] \in \mathcal{P}^{K-1}.$$

Using this, one obtains an equivalent flow $[\Pi]$ on $\mathcal{P}^{K-1} \times \mathfrak{A}$, where

$$[\Pi](t)(\theta, [x_0]) \;=\; (\theta \cdot t, \Sigma(\theta, t)[x_0]), \qquad \text{for } (\theta, [x_0]) \in \mathfrak{A} \times \mathcal{P}^{K-1}, \qquad (17)$$

compare with (10). Since $\mathcal{P}^{K-1} \times \mathfrak{A}$ is a compact space, there exists at least one invariant probability measure on $\mathcal{P}^{K-1} \times \mathfrak{A}$. More on these invariant measures appears in Section 5.

One obtains additional information about the dynamics of the flow $[\Pi(t)]$ focusing on the Lyapunov vectors. Assume that the largest Lyapunov Exponent, λ_1 has multiplicity $m_1 \;=\; 1$, and let $\mathcal{B}_1 \;=\; \{w\}$, where $\|w\| = 1$ is a Lyapunov vector associated with λ_1. The following is an important property of the flow $[\Pi]$ in this setting:

Let $w \in R^K$ be given as above, and let $v \neq 0$ be another vector in $\mathbb{R}^K$. Then the solutions satisfy:

$$\lim_{t \to \infty} d\,(\Sigma(\theta, t)\,[v]\,,\, \Sigma(\theta, t)\,[w]) \;=\; 0, \qquad (18)$$

where d is a induced metric on $\mathfrak{A} \times \mathcal{P}^{K-1}$.

2.3 The Finite-Time Lyapunov Vectors

As discussed previously, every non-zero vector is a Lyapunov vector associated to one of the Lyapunov exponents. Due to the relation (18), we see that the vectors associated with the largest Lyapunov exponent λ_1 define the direction(s) of maximal growth. Next we turn our attention to the question of finding good finite-time approximations of these Lyapunov vectors.

We begin by constructing a piecewise autonomous approximation of the Tangent Linear Equation for (1). To this end, we replace (8) by

$$\partial_t Y(t) \;=\; A_n\, Y(t), \;\text{ for } t_n < t \le t_{n+1}, \tag{19a}$$

$$Y(0) = Y_0, \tag{19b}$$

$$t_n = n\,\delta t, \qquad n = 0, 1, 2, \cdots, N-1. \tag{19c}$$

where $A_n = A(y(t_n)) := \frac{\partial G}{\partial y}|_{y=(y(t_n))}$ and $y(t_n)$ is the value of an exact solution of (1) at $t = t_n$. The solution of this system at the grid points t_n is, explicitly and recursively, given by

$$Y_{n+1} \;=\; e^{\delta t\, A_n}\, Y_n, \qquad \text{for } n = 0, 1, 2, \cdots, N-1. \tag{20}$$

Consequently, Y_N, the solution at time $T = N\,\delta t$, is given by

$$Y(T) \;=\; Y_N \;=\; Z(T)\, Y_0, \tag{21}$$

where

$$Z(T) \;:=\; e^{\delta t\, A_{N-1}} e^{\delta t\, A_{N-2}} \cdots e^{\delta t\, A_1} e^{\delta t\, A_0}.$$

We now define the approximation of the Lyapunov Vector associated with the largest Lyapunov Exponent λ_1, at the time $t = T$ - *the finite-time Lyapunov vector*, as the direction of steepest ascent for the matrix $Z(T)$. For example,

if one had used an explicit Euler scheme, then $W(T)$ would be an Eulerian approximation of $Z(T)$, where

$$W(T) = (I + \delta t\, A_{N-1}) \cdots (I + \delta t\, A_1) \cdots (I + \delta t A_0).$$

The finite-time Lyapunov Vector is the singular vector corresponding to the largest singular value of $Z(T)$.

3 Continuum Limits of the Bred Vector Algorithm and EBV

We will elucidate further the relationship between the BV algorithm and EBV and the dynamics of the underlying system. We also will formalize the connection between these algorithms and the linear tangent equation, and thus the finite-time Lyapunov vectors. To accomplish these aims we revert to a continuum formulation and allow for the vector $y(t)$ in (1) to be a member of a possibly infinite-dimensional normed vector space, henceforth denoted by H. Since energy norms are used in many geophysical fluid mechanics problems, we will assume H is a complete inner-product space (that is, a Hilbert space) and denote the inner product in H by $(\cdot, \cdot)$. In the following calculation, we assume that G has the necessary smoothness to make all mathematical steps rigorous. Let y and $y + \delta y$ be two solutions of (1) that satisfy the initial conditions $y(0) = y_0$ and $(y+\delta y)(0) = y_0 + \delta y_0$. Then $\delta y(t)$ solves

$$\frac{d(\delta y)}{dt} = \left[G(y + \delta y) - G(y)\right],\ \delta y(0) = \delta y_0. \tag{22}$$

Even if G is autonomous, the resulting equation (22) for δy is nonautonomous. Integrating (22) we obtain

$$\delta y(t_0+\delta t) = \left([y(t_0)+\delta y(t_0)] + \int_{t_0}^{t_0+\delta t} G(y+\delta y)\, dt \right) - \left(y(t_0) + \int_{t_0}^{t_0+\delta t} G(y)\, dt \right). \tag{23}$$

We next define a family of vectors on a sphere centered at 0 with radius $\|\delta y(0)\|$ in H by

$$v(t) = \frac{\delta y(t)}{\|\delta y(t)\|}\|\delta y(0)\|.$$

Above, the norm is that induced by the inner product on H (in practice an ℓ^2-based norm): $\|u\| = \sqrt{(u,u)}$.

Assuming that δy is regular enough, we calculate the evolution of $v(t)$:

$$\begin{aligned} \frac{1}{\|\delta y(0)\|}\frac{dv}{dt} &= \frac{1}{\|\delta y(t)\|^2}\left[\left(\frac{d}{dt}\delta y\right)\|\delta y(t)\| - \delta y(t)\frac{d}{dt}\|\delta y\| \right] \\ &= \frac{1}{\|\delta y(t)\|^2}\left[G_y(\delta y,t)\|\delta y(t)\| - \frac{1}{2}\frac{\delta y(t)}{\|\delta y(t)\|}\cdot\frac{d}{dt}\|\delta y\|^2 \right], \end{aligned} \tag{24}$$

which makes explicit the fundamental dependence of the time evolution of v on the norm. Here, $G_y(\delta y,t) = G(y+\delta y) - G(y)$. Taking the scalar product of (22) with $\delta y(t)$, we also have

$$\frac{1}{2}\frac{d}{dt}\|\delta y\|^2 = (G_y(\delta y,t), \delta y(t)).$$

Substituting this relation in (24) gives after some simplifications,

$$\frac{1}{\|\delta y(0)\|}\frac{dv}{dt} = \frac{G_y(\delta y,t)}{\|\delta y(t)\|} - \left(\frac{G_y(\delta y,t)}{\|\delta y(t)\|}, \frac{v(t)}{\|\delta y(0)\|}\right)\frac{v(t)}{\|\delta y(0)\|}. \tag{25}$$

We integrate (25) independently on $[0,\delta t]$, $[\delta t, 2\delta t]$, ..., $[(N-1)\delta t, N\delta t = T]$ and denote $v(t)$ on each of the intervals $[(k-1)\,\delta t, k\,\delta t]$ by $v_k(t)$, obtaining N integral equations with N free parameters (the integration constants) $v_k^0, k =$

$1, \ldots, N$:

$$v_k(t) = v_k^0 + \|v_k^0\| \int_{(k-1)\delta t}^{t} \left(\frac{G_y(\delta y, \tau)}{\|\delta y(\tau)\|} - \left(\frac{G_y(\delta y, \tau)}{\|\delta y(\tau)\|}, \frac{v_k(\tau)}{\|v_k^0\|} \right) \frac{v_k(\tau)}{\|v_k^0\|} \right) d\tau,$$
$$(k-1)\delta t \leq t < k\,\delta t, \quad k = 1, \ldots, N. \tag{26}$$

We observe that the integrand in (26) above is simply the component of the vector $\frac{G_y(\delta y, t)}{\|\delta y(t)\|}$ perpendicular to $v_k(t)$ in H at time t.

We now take

$$v_1^0 = \delta y(0), \qquad v_k^0 = \lim_{t \to (k-1)\delta t^-} v_{k-1}(t), \quad 1 \leq k \leq N, \tag{27}$$

assuming the limits exists, and extend the domain of definition of v_k by setting $v_k = 0$ if $t < (k-1)\delta t$ or $t \geq k\delta t$.

We then let $V_N = v_1 + \cdots + v_N$, and note that V_N is continuous in time. We further assume that

$$V = \lim_{N \to \infty} V_N$$

exists and is continuous in time, and that the family of the piecewise differentiable curves δy_N on $[0, T]$, obtained by solving (22) together with $\delta y((k-1)\delta t) = v_k^0, k = 1, \ldots, N$ converges to V as $N \to \infty$. We can guarantee this assumption if the convergence as $N \to \infty$ is true in the supremum norm, since δy_N interpolate V_N at at least N points. By uniqueness of solutions to (22), letting $N \to \infty$ we conclude that

$$\delta y(t) = V(t), \quad t \geq 0.$$

Thus, we obtain the following representation formula for V:

$$V = V(0) + \int_0^T \left(G_y(V(t), t) - \left(G_y(V(t), t), \frac{V(t)}{\|V(t)\|} \right) \frac{V(t)}{\|V(t)\|} \right) dt. \tag{28}$$

Equivalently,

$$U(T) \;\; = U(0) + \textstyle\int_0^T \left[G_y^\varepsilon\left(U(t), t\right) - \left(G_y^\varepsilon\left(U(t), t\right), U(t) \right) U(t) \right] dt, \tag{29}$$

where for all $t \geq 0$,

$$\varepsilon = \|V(0)\| = \|V(t)\|, \qquad U(t) = V(t)/\varepsilon; \quad G_y^\varepsilon\left(U(t), t\right) = \frac{1}{\varepsilon} G_y\left(\varepsilon U(t), t\right)$$

(by construction V_N has constant norm).

We now make the initial perturbation δy_0 smaller and smaller, so that ϵ is approaching 0. Since $G_y^\varepsilon\left(U(t), t\right) = \dfrac{G(y + \varepsilon U(t)) - G(y)}{\varepsilon}$, it follows that

$$\lim_{\varepsilon \to 0} G_y^\varepsilon\left(U(t), t\right) = D_{U(t)} G(y),$$

where the right-hand-side is the directional derivative of G in the $U(t)$-direction.

We next discuss rescaling in time in the context of the Linear Tangent Model for the problem (1), which we recall is given by

$$\frac{d(\delta f)}{dt} = A(t)\delta f, \quad A(t) = DG(y)|_{y=y(t)}, \tag{30a}$$

$$\delta f(t_0) = \delta f_0, \quad \delta f \in H, \tag{30b}$$

where DG is the Jacobian matrix of the map $G(y)$, and $y(t)$ is the solution of (1). As in Section 2, $U(t, s)$ denotes the solution operator of (30), taking the solution at time s to the solution and time t, and we write $U(t) := U(t, 0)$. We recall also that U is a semiflow, *i.e.*,

$$U(t + s) = U(t + s, t)\, U(t), \qquad \forall t, s \in \mathbb{R}. \tag{31}$$

We consider various strategies to obtain a vector with a given amplitude $\varepsilon > 0$ from the linear tangent equation: we can only rescale the end result at time $t = T$, we can rescale from time to time, or we can rescale *continuously*. This last strategy is identical to the rescaling strategy in the continuous breeding we have just discussed.

By linearity it is clear that any discrete-time rescaling is equivalent to rescaling at the final time. For notational convenience, we choose a uniform partition of the time interval $[0,T]$: $t_0 < t_1 = \delta t < \cdots < t_n = n\delta t < \cdots < T = N\delta t$. We let ϵ_k denote the size of the perturbation $\delta f(t_k)$, and set $V_k = \delta f(t_k)/\epsilon_k$. Then, the semiflow property (31) implies, by induction on $1 \le k \le N$, that

$$\begin{aligned}
&\hat{V}_1 = \delta f(t_1) = U(\delta t; t_0)\delta f_0,\\
&V_1 = \frac{\varepsilon_0}{\varepsilon_1}\delta f(t_1);\\
&\quad\vdots\\
&\hat{V}_{k+1} = U(\delta t; t_0 + k\delta t)\, V_k = \frac{\varepsilon_0}{\varepsilon_k} U(\delta t; t_0 + k\delta t)\, \delta f(t_k) = \frac{\varepsilon_0}{\varepsilon_1} U((k+1)\delta t; t_0)\, \delta f_0;\\
&V_{k+1} = \frac{\varepsilon_0}{\varepsilon_k} \times \frac{\varepsilon_k}{\varepsilon_{k+1}} \delta f(t_{k+1}) = \frac{\varepsilon_0}{\varepsilon_{k+1}} \delta f(t_{k+1}),
\end{aligned}$$

so that

$$V_N = \frac{\varepsilon_0}{\varepsilon_N}\delta f(T), \tag{32}$$

which is precisely the result of rescaling only at the final time T.

We next discuss rescaling continuously in time, by which we mean solving the following differential equation:

$$\frac{d(V)}{dt} = A(t)V - (A(t)V, V)V, \tag{33a}$$

$$V(t_0) = \frac{\delta f_0}{\|f_0\|}, \quad V \in H. \tag{33b}$$

To justify this ansatz, we assume for simplicity that $H = \mathbb{R}^K$ (the infinite-dimensional case follows then by a density argument), and for a fixed time t set $\mathbf{r} = V(t)$ (and assume that it is larger than a fixed positive number, say ϵ). We then decompose the infinitesimal change dV to $V(t)$ after an

infinitesimal time increment dt as

$$dV = d\mathbf{r} + d\boldsymbol{\omega},$$

where $d\mathbf{r}$ is parallel to $\mathbf{r}$ and $d\boldsymbol{\omega}$ is perpendicular to $\mathbf{r}$, so that they are independent increments.

We continue to assume throughout that $A(t)$ is a sufficiently regular operator, and in particular that $D := Max_{[0,T]}\|A(t)\|_{\text{op}} < \infty$, where $\|A(t)\|_{\text{op}}$ is the norm of $A(t)$ as operator from H to H. Then $\|dV\| = O(D\,dt)$. Since

$$\frac{\|\mathbf{r}\|}{\|\mathbf{r} + d\mathbf{r} + d\boldsymbol{\omega}\|}(\mathbf{r} + d\mathbf{r} + d\boldsymbol{\omega}) = \frac{\|\mathbf{r} + d\mathbf{r}\|}{\|\mathbf{r} + d\mathbf{r} + d\boldsymbol{\omega}\|}[\mathbf{r} + d\boldsymbol{\omega} + o(d\boldsymbol{\omega}; d\mathbf{r})],$$

rescaling $\mathbf{r} + dV$ to the size of $\mathbf{r}$ is equivalent to changing the infinitesimal increment to $d\boldsymbol{\omega}$.

Given the assumed regularity of $A(t)$, and hence of the solutions to (30), any discrete approximation $\dfrac{\|\mathbf{r} + \delta\mathbf{r}\|}{\|\mathbf{r} + \delta\mathbf{r} + \delta\boldsymbol{\omega}\|}$ will be uniformly bounded from above and below, and will converge to 1 uniformly on $[t_0, T]$ as the mesh size δt approaches zero. By (32) and a straightforward approximation argument, we conclude that the solution V of (33) is parallel to V_N at $t = T$. Therefore, rescaling continuously in time is equivalent to rescaling at the final time only.

We point out that, while for any fixed time $t_0 < t$, (33) is the limit of the breeding vector evolution equation (29) as the amplitude of the perturbation δy vanishes, it does not necessarily follow that the respective solutions must have the same long-time behavior. For instance, when (1) has a fixed point at the origin, $DG(0) \equiv 0$ and hence the tangent linear equation fixes every perturbation, yet the breeding algorithm may exhibit a wide range of different behaviors.

We next investigate the continuum limit for the EBV algorithm. Writing

a system of differential equation for the EBV algorithm is not as straightforward as for the BV algorithm. The perturbations $\delta y(\iota)$ in the ensemble that grow fastest at the largest amplitude (before rescaling) at time t will obey the equation (29) at least for short time. All the other perturbations will satisfy the following differential equation:

$$\frac{d(\delta y(\iota))}{dt} = G_y(\delta y(\iota), t) - \frac{\delta y(\iota)}{\varepsilon_0}\frac{d\varepsilon(t)}{dt}, \tag{34}$$

where $\iota \in \mathbb{I}$, and $\frac{d\varepsilon(t)}{dt}$ is the maximum growth rate among all ensemble members with the maximum amplitude at time $t = T$. Assuming enough smoothness on $\epsilon(t)$ (piecewise locally Lipschitz, for instance) as needed, (34) is justified by taking the right time derivative of the vector $\delta y(\iota)(t)\frac{\varepsilon_0}{\varepsilon(t)}$, where $\varepsilon(\tau), \tau \geq t$ is the maximum norm of the ensemble members at time τ in absence of any rescaling after time t. As ε_0 approaches zero for time t fixed, all ensemble members that have fallen short of the maximum amplitude or maximum growth rate point in directions along which (34) approaches the linear tangent equation, since $d\varepsilon(t)/dt \to 0$ by continuity of G_y at zero with $G_y(0, t) = 0$. Therefore, the convergence to the linear tangent equation is expected to be quicker for the EBV compared to the BV in general.

Given the high degree of complexity of physical systems represented by (1), in particular weather models, it is desirable to reduce the dependence of V in (28) to a few parameters. For breeding algorithms, it seems of particular importance to eliminate the effect of the choice of initial perturbation, and reduce it to a dependence on the size of the perturbation. (We must assume that a choice of norm has been made *e.g.* by physical considerations). Even when the initial ensemble is chosen randomly, the rescaling rule in the BV may prolong the life of "artificial" instabilities. Therefore, we take as initial

ensemble *all possible perturbations of size* ε, *i.e.*, a sphere $\mathbf{S}_\varepsilon(t=0)$ of radius ε centered at 0 in a (possibly high dimensional) vector space of functions, and evolve all the vectors on the sphere at once. Assuming that the dynamics is sufficiently regular, the initial sphere is deformed continuously into a "topological sphere." The deformation will be maximal along the directions that grown the most among *all* the possible perturbations of the same size. A region on this deformed sphere that grows relatively slowly will shrink in size. By continuity, there is at least one point on this deformed sphere the size of which achieves the maximum. As with the discrete EBV, vector are simultaneously rescaled by the factor

$$\left(\sup_{h\in\mathbf{S}_\varepsilon(t)}\|h\|\right)^{-1}\varepsilon \tag{35}$$

at time t. After rescaling the maximum amplitude on the topological sphere is always ϵ. The largest growing and most important instabilities are part of the topological sphere, with amplitude near ε, *i.e.*, the peak of the deformed sphere.

In nonlinear systems, it is possible that a perturbation will not grow very rapidly in the zone of perturbations with size near ε, but instead, a multiple of it can be dominant among the smaller perturbations and might grow quite fast there. In fact, if we choose a (sufficiently fine) mesh $\mathbf{\Gamma}_\varepsilon(t_{n-1})$ on the deformed sphere $\mathbf{S}_\epsilon(t_{n-1})$ and $V(t_{n-1})\in\mathbf{\Gamma}_\varepsilon(t_{n-1})$, $\bar{V}(t_n)$ is the evolved vector under the breeding algorithm but before rescaling, and $V(t_n)$ the corresponding ensemble member after rescaling, then we have

$$\frac{\|V(t_n)\|}{\|V(t_{n-1})\|}=\mathbf{R}_n^{\min}\frac{\|\bar{V}(t_n)\|}{\|V(t_n)\|},$$

where $\mathbf{y}^{\min}$ is defined similarly to (7). In a nonlinear system, this ratio can be larger than, or very close to one, for some range of perturbation sizes

smaller than ε. This points to another interesting and desirable feature of the EBV algorithm: we not only obtain significant perturbations for the initial size perturbation size ε, but also observe the perturbations with smaller sizes with strong growth, provided that they can overcome the suppressing impact of the ratio R_n. A case in point will be presented in Section 4 for the CY92 model. Even at late times and small amplitudes, we can see perturbations surviving the harsh rescaling strategy and they resemble what one would get from the usual BV algorithm at those sizes (and the finite-time Lyapunov vectors) very closely, whereas the dominating perturbation of size ε resembles a particular BV of size ε. One expects that many of the BVs cannot survive the EBV algorithm, unless the nonlinear part in the governing system of equations is negligible.

4 Applications of the Bred Vector Algorithms

We will be comparing the BV and the EBV on two problems. The Lorenz equations, or Lorenz63 (see Lorenz (1963)), and a dissipative and forced nonlinear partial differential equation that arises in modeling the thermohaline circulation (see Cessi and Young (1992)). The latter will be denoted as the CY92. It is a Cahn-Hilliard equation and it will shown to have an inertial manifold. We will also have occasion to compare the BV and EBV results to the finite-time Lyapunov vector outcomes.

Throughout we will use an explicit fourth order Runge-Kutta time marching scheme, for the calculation of the base solutions as well as for the calculations of the BV, EBV, and finite-time Lyapunov vectors.

It turns out that our comparisons are very much influenced by mult(λ_1),

the multiplicity of the largest Lyapunov exponent, see (15). The two models we study here differ in the sense that mult$(\lambda_1) = 1$ (and λ_2 is "not too close" to λ_1) for the CY92, while mult$(\lambda_1) \geq 2$ (or λ_2 is very "close" to λ_1) for the Lorenz63.

4.1 The Lorenz63 Model

The finite-dimensional, nonlinear Lorenz63 model has often been used as benchmark for testing sensitivity and, in particular, as a test problem for BV (see for example Toth & Kalnay *op. cit.*) The Lorenz model is described by the solutions of the nonlinear system:

$$\partial_t X \;=\; A\,X + N(X), \tag{36}$$

where $X = (X_1, X_2, X_3) \in \mathbb{R}^3$, with

$$A = \begin{pmatrix} -\sigma & \sigma & 0 \\ r & -1 & 0 \\ 0 & 0 & -b \end{pmatrix},\; N(X) = \begin{pmatrix} 0 \\ -X_1 X_3 \\ X_1 X_2 \end{pmatrix},\; DN(X) = \frac{\partial}{\partial X} N(X). \tag{37}$$

$DN(X)$ is a 3×3 matrix-valued function. The associated Tangent Linear Model is the skew product system

$$\begin{aligned} \partial_t X \;&=\; A\,X + N(X) \\ \partial_t U \;&=\; (A + DN(X))\;U, \end{aligned} \tag{38}$$

where $U = (U_1, U_2, U_3) \in \mathbb{R}^3$. The U-equation in (38), which is a linear equation, is of special interest to us.

We will use the notation for the Tangent Linear Model introduced in Section 2.2. With the initial condition $\theta = X_0 \in \mathbb{R}^3$, we let $\theta \cdot t = S(\theta, t) = X(t)$ denote the solution of the nonlinear equations (36) that satisfies $S(\theta, 0) = \theta$.

In this study we set $r = 28, b = 8/3, \sigma = 10$. It is well-known that for these values of the parameters the Lorenz63 model has a chaotic global attractor. For the nonautonomous U-equation; $\partial_t U = (A + DN(\theta \cdot t))\, U$, we let

$$U(t) \;=\; U(\theta, t)\, U_0, \qquad \text{denote the solution,} \tag{39}$$

where $U_0 \in \mathbb{R}^3$, and $U(0) = U_0$.

Figure 1 shows the time evolution of the BVs and the finite-time Lyapunov vectors, obtained with a time step of $\delta t = 0.005$. The computations of the BVs and finite time Lyapunov vectors are performed with a time interval equal to the dynamics time step. At $t = 0$, $X = 0.1493$, $Y = 6.2575$, $Z = 1.8407$. The perturbation vector $\delta\mathcal{Y}_0$ was $[1, 1, 1]$. We use the same initial conditions and perturbations for the finite time Lyapunov vectors calculations. In our numerical simulations for the Lorenz63 system we find that the BVs and finite-time Lyapunov vectors can spend long intervals of time being nearly coincidental. Figure 2, shows a superposition of the BVs on the numerically-obtained base solution, in the time interval $[20, 40]$, corresponding to Figure 1. The vectors are computed at each dynamic time step, however, we show here the vectors at 0.2 time intervals. A rotation of the figure about the Z axis would show that the BVs occupy a reduced-dimensional manifold, which contains the strange attractor. Moreover, a plot such as the one presented in Figure 2 showing the finite-time Lyapunov vectors instead of the BVs would look similar, as would be surmised from the small differences between the time histories of the finite-time Lyapunov vectors and BVs in Figure 1.

While we do not calculate the Lyapunov spectrum, as it might occur in reference to some ergodic, invariant measure on the attractor, it seems

clear that, for the Lorenz attractor, one faces one of two possibilities for the spectra:

$$\text{either} \quad \lambda_k < \lambda_1 \quad \text{or} \quad \lambda_k < \lambda_2 < \lambda_1.$$

In the first case, one has $\text{mult}(\lambda_1) = 2$. In the second case, it appears that λ_2 must be very close to λ_1. By treating the Lyapunov vectors associated with λ_1 and λ_2 together, one has thereby creating another multiplicity 2 environment.

We will use the term "attractor" to refer to a compact, invariant set $\mathfrak{A}$ that attracts a neighborhood of itself. As a result, $\mathfrak{A}$ is Lyapunov stable, as a set. Consequently, for $0 < \varepsilon < 1$, there is a family of ε-neighborhoods, N_ε, of $\mathfrak{A}$, where each neighborhood is positively invariant and $\mathfrak{A} = \cap_{\varepsilon>0} N_\varepsilon$. When we write that θ is near $\mathfrak{A}$, we mean that $\theta \in N_\varepsilon$, for some small $\varepsilon > 0$. See Chapter 2 in Sell and You (2002), for a history of this concept and more information.

We believe that the dynamical properties of the BV and EBV algorithms that arise in case of the Lorenz attractor will be typical of what happens in other environments, where mult ≥ 2.

By using a somewhat different –but equivalent formulation– Toth & Kalnay have suggested that the time evolution $BV(t)$ is a good approximation of the tangent linear solution $U(t)$, over bounded time-intervals. The qualitative comparisons presented above certainly bear this point. However, as a consequence of the continuum limit theory in Section 3, we see that this perceptive observation has a solid mathematical basis, provided that the time-step δt is small and the perturbation amplitude is small, as well.

In order to find a numerical validation of the continuum limit theory

described in Section 3, we will calculate the distance

$$D(t, \delta t) := d\left([BV(t)]\,,\,[U(t)]\right), \tag{40}$$

where d is an induced metric on $\mathfrak{A} \times \mathcal{P}^{K-1}$ (see (18)), and $BV(t)$ is a BV at time t. For this calculation, we assume that both $BV(t)$ and $U(t)$ satisfy the same initial condition $(y_0, \delta\mathcal{Y}_0)$ at $t = 0$, and that $y_0 = \theta$ is near the attractor $\mathfrak{A}$. We also use identical ensembles of perturbation vectors for both the BV and the EBV algorithms. We then fix $T > 0$, and we examine the distances $D(T, \delta t)$ (in the projective space), for different choices of δt. Our goal is to show that $D(T, \delta t)$ becomes smaller, as δt gets smaller. The max and min values, for both the BV and the EBV algorithms, are reported in Table 1. For the calculations used to generate the data in Table 1, we had set the initial conditions for both $BV(t)$ and $U(t)$, at $t = 0$, so that $X = 0.5688$, $Y = 0.4694$, $Z = 0.0119$, and $\delta\mathcal{Y}_0 = [1, 1, 0]$. For the ensemble, we took the 3-fold Cartesian product of the set

$$\{-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1\}$$

in $\mathbb{R}^1$, and projected this product onto the sphere of radius 0.1, centered at the origin, in $\mathbb{R}^3$. By eliminating the repetitions introduced with this projection, one obtains an ensemble of 584 distinct points in $\mathbb{R}^3$. (Note that the initial conditions are "near" the Lorenz attractor.) The trajectory for the nonlinear problem was computed with a step size $dt = 0.0001$, so that any observed variations are only due to the differences in the algorithms and the relative affects on the perturbations. In Table 1, we fixed $T = 2$ and made two choices for δt, namely $\delta t = 0.004$ and 0.001. The results compare the outcomes of the BV and the EBV algorithms.

Table 1: *Maximum and minimum values of* $D(T, \delta t)$ *with* $T = 2$*, for two* δt*. See (40).*

δt	max BV	min BV	max EBV	min EBV
0.004	1.40×10^{-1}	5.26×10^{-4}	1.01×10^{-1}	1.10×10^{-3}
0.001	7.77×10^{-2}	1.03×10^{-4}	3.48×10^{-2}	4.16×10^{-5}

The maximal values of $D(T, \delta t)$, for both the BV and EBV algorithms, are essentially the same for both choices of δt, and the minimal values are essentially the same for $\delta t = 0.004$. When one moves from $\delta t = 0.004$ to $\delta t = 0.001$, both minimal values decrease, as expected. However the drop in the minimal value for the EBV algorithm is substantialy larger than the drop for the BV algorithm. This demonstrates a significant advantage in using the EBV over the BV.

One can, of course, use differing initial conditions for the two solutions $BV(t)$ and $U(t)$ and/or larger perturbations. However, one cannot expect to replicate the results seen in Table 1 in that case. First, there is a transient phase, which ends when the two solutions are "near" the attractor. Even if this transient phase is short, one still has an obstacle that the flow on the attractor is not Lyapunov stable. Essentially all pairs of nearby orbits diverge over long time intervals.

Figure 3 is devoted to the results of a calculation of $EBV(t)$ alone. The figure depicts the time evolution of the absolute value of a 98 distinct ensemble members in the EBV calculation. The initial perturbations are made to sample a perturbation sphere of amplitude 1 about the initial conditions. Same parameters and base solution, as in Figures 1 and 2. The figure high-

lights the rapid decay of many of the vectors and the eventual size-ordering that is inherent in the EBV algorithm. For $0 < t \leq 13$, the results in Figure 3 are rather chaotic. However, for $t \geq 13$ a very interesting pattern evolves. Focus on the 3 graphs at the top, for $t \geq 13$. The topmost graph takes on the value 1, for all $t \geq 0$, which is expected. (This corresponds to the Lyapunov vector with the largest exponent - most likely $\lambda_1 = 0$.) For $t \geq 13$, the next 2 graphs at the top represent the Lyapunov vectors with exponents $\lambda_3 < \lambda_2 < 0$. This separation of scales is a unique feature of the EBV.

This feature of the EBV can not be replicated by BV, even a BV with an ensemble of perturbations. The reason for this is that the rescaling rule for the BV algorithm forces all the perturbations to have the same norm, for each $t \geq 0$. Thus the Figure 3 for the EBV would be replaced by a figure for the BV, where all of the perturbations are plotted on the top line only.

An interesting outcome of the use of the EBV algorithm on Lorenz63 is that it exposes the fractal behavior (see Mandelbrot (1977)), of the Lorenz attractor. We computed a 584-member EBV, using the same parameter values, initial conditions and base trajectory as was used in generating Table 1, for the $\delta t = 0.001$ case. We constructed plots by merging all the vectors between time 24 and 30, at intervals of $dt = 0.1$. In that time, at any given instant, one only observes a couple of members reaching highest amplitude. Figure 4 show three views of the same plot. Zooming into Figure 4a by a factor of 8, 32, and 60, respectively, we obtain Figure 5. The nonautonomous nature of the Lorenz63 Time Linear Model is the cause of the differing and interesting patterns seen in Figures 4 and 5.

4.2 A Cahn-Hilliard Equation Example

In their work on the thermohaline dynamics Cessi and Young (1992) proposed a coupled model for the circulation, salinity, temperature, and density of the oceans, with atmospheric forcing. The crux of the model is the partial differential equation for salinity: The slow-time dynamics of the ocean salinity $S(x,t)$, zonally-averaged, and as a function of latitude $x \in [-\pi, \pi]$ and time $t \geq 0$, is described by

$$\begin{aligned} \frac{\partial S}{\partial t} &= \alpha \frac{\partial^2}{\partial x^2}[f(x) + \mu S(S - \sin(x))^2 + S - \gamma \frac{\partial^2 S}{\partial x^2}], \quad t > 0, \\ S(x,0) &= S_0(x). \end{aligned} \tag{41}$$

The equation is subject to zero-flux and zero-stress boundary conditions at the poles, however, we will be considering periodic boundary conditions (for steady and periodic forcing $f(x)$ as well as equilibrium solution $\sin(x)$ the conclusions that follow apply to the zero-flux case). The parameter α makes explicit the relationship between the fast and slow time scales. It will be used here to control the rate of the evolution in the numerical examples, set to 3.5×10^{-3}. We will fix $\gamma = 0.001$, and $\mu = \sqrt{10}$. The forcing $f(x)$ is a prescribed function that reflects balances of evaporation and precipitation of freshwater; it can be symmetric, about the Equator ($x = 0$), but is more typically, asymmetric. The second derivative of the forcing function, will be chosen to be

$$\partial_{xx} f(x) = \begin{cases} \frac{2}{27}(36 - 39\mu^2 \cos^2 x - 81 \cos^2 x + 8\mu^2 + 25\mu^2 \cos^4 x) \sin x, \\ \qquad \text{for} \quad x \in [-\pi, 0], \\ -\frac{2}{9}(3\mu^2 \cos^2 x - 3 - \mu^2) \sin x, \\ \qquad \text{for} \quad x \in (0, \pi]. \end{cases} \tag{42}$$

It is shown in Figure 6a. The initial condition chosen for this computation $S_0(x) = \cos(x)$. The base solution obtained numerically is shown in Figure

6b.

A great deal is known and can be said about the mathematical structure of CY92 and its solutions. By letting $S = \partial_x u$ and $f = \partial_x F$ the CY92 can be related to the Cahn-Hilliard Equation (CHE) with forcing: The general CHE, for u above, is

$$\partial_t u + \nu \triangle^2 u \, = \, \triangle(g(u)) + F \qquad \text{for } x \in \Omega \text{ and } t \geq 0, \tag{43}$$

where $u = u(t, x)$ is a scalar field, $\nu > 0$ is a constant, $\triangle$ is the Laplacian operator, $\triangle^2$ is the bi-harmonic operator, and g is a polynomial of degree 3: $g(u) = \sum_{j=1}^{3} a_j \, u^j$, with $a_3 > 0$. The term $F = F(x)$ is the forcing function. In general, the domain Ω may be an open bounded domain in the Euclidean space $\mathbb{R}^m$, with $1 \leq m \leq 3$. However, we restrict our attention, to the case where $m = 1$ and Ω is the interval $\Omega \, = \, (-\pi, \pi)$, with boundary $\Gamma = \{\pm\pi\}$. As we will see, the multiplicity of the largest Lyapunov exponent for this problem is 1.

For the analysis of the solutions of the CHE, one will use the standard Sobolev spaces $H^j = H^j(\Omega)$, where $j \geq 0$ is an integer and $H^0 = L^2(\Omega)$. As usual, the inner product and norm on H^0 is denoted by $\langle \cdot, \cdot \rangle \, = \, \langle \cdot, \cdot \rangle_0$ and $\| \cdot \| \, = \, \| \cdot \|_0$. For $j \geq 0$ one uses

$$\|u\|_{j+1}^2, = \, \|u\| j^2 + \sum_{|\alpha| = j+1} \int_\Omega |D^\alpha u|^2 \, dx,$$

where

$$D^\alpha = \frac{\partial^\alpha}{\partial x^\alpha}, \quad \alpha \geq 0,$$

and

$$\langle u, v \rangle_{j+1} \, = \, \langle u, v \rangle_j + \sum_{|\alpha| = j+1} \int_\Omega (D^\alpha u) \, (D^\alpha v) \, dx.$$

The Bi-harmonic Operator $B\,u \,=\, \triangle^2\,u$ on Ω and the linear Bi-harmonic equation:

$$\partial_t u + \nu \triangle^2\, u \,=\, 0 \tag{44}$$

play a basic role in the study of the CHE. We assume for now that the Operator B satisfies either the non-flux boundary conditions:

$$\frac{\partial u}{\partial n} \,=\, 0 \quad \text{and} \quad \frac{\partial(\triangle u)}{\partial n} \,=\, 0, \qquad \text{on } \Gamma, \tag{45}$$

or the periodic boundary conditions. (We will use BC to refer to either of these boundary conditions.) Thus the fourth order linear operator B is a nonnegative, selfadjoint operator with compact resolvent. Hence B is a sectorial operator and $-B$ is the (infinitesimal) generator of the analytic semigroup e^{-Bt}, see Section 3.7 in Sell and You (2002). The related Fourier series for B can be found in Section 3.2 of Sell and You (2002).

The null space $\mathcal{N}(B)$, is the $m = 1$-dimensional space $\mathbb{R}$, which is imbedded in H^0. For the study of the CHE, it is convenient to delete this space from the problem. We do that by using $\mathbb{P}_0$, the orthogonal projection of H^0 onto $\mathcal{N}(B)^\perp$. We use the spaces $H_0^j \,:=\, \mathbb{P}_0\, H^j$ and we define the operator $B_0 \,:=\, \mathbb{P}_0\, B$. (The operator $-B_0$ is also the infinitesimal generator of an analytic semigroup $e^{-B_0 t}$.) We define the spaces

$$V_0^j \,:=\, \{u \in H_0^j \,:\, BC \text{ holds}\}, \qquad \text{for } j \geq 1.$$

The domain of B_0 is

$$\mathcal{D}(B_0) \,=\, V_0^4 \,=\, \{u \in H_0^4 \,:\, \text{ BC holds}\}.$$

In the sequel, we will assume that the forcing function F in (43) satisfies

$F \in H_0^j$, for some $j \geq 0$. That is to say,

$$\int_\Omega F\, dx \,=\, 0.$$

Thus, by integrating (43), one observes that any solution $u = u(t,x)$ satisfies

$$\overline{u}_0 \,:=\, \frac{1}{|\Omega|}\int_\Omega u_0(x)\, dx \,=\, \frac{1}{|\Omega|}\int_\Omega u(t,x) \,=\, \overline{u}(t), \qquad \text{for } t \geq 0. \tag{46}$$

One seeks solutions of the CHE (43) in the Sobolev space V_0^2. A *mild* solution $u = u(t)$ of the initial value problem is given by

$$u(t) \,=\, e^{-\nu B_0 t}u_0 + \int_0^t e^{-\nu B_0 (t-s)}\, [\triangle(g(u(s))) + F]\, ds. \tag{47}$$

We will denote the maximally defined solutions of (47) by $u(t) \,=\, S(t)u_0$. With $u_0 \in V_0^2$, this mild solution is uniquely determined, with $S(t)u_0 \in V_0^2$, for $0 \leq t < T(u_0)$, where $0 < T(u_0) \leq \infty$.

Because of (46), we see that, for $j \geq 2$, the spaces

$$H_0^j \,:=\, \{\phi \in H^j \,:\, \overline{\phi} = 0\} \tag{48}$$

are positively invariant spaces for the solutions of (43). Thus we will focus only on solutions $S(t)\, u_0$ that satisfy $\overline{u}_0 = 0$. We also denote the collection of stationary solutions of the CHE by

$$Q \,:=\, \{u_0 \in H^2 \,:\, S(t)u_0 = u_0 \text{ for all } t \in \mathbb{R}\},$$

and $Q_0 = Q \cap H_0^2 = Q \cap V_0^2$.

The CHE with $F \equiv 0$: The basic problem with forcing $F \equiv 0$ is of special interest. The equation (43) becomes:

$$\partial_t u + \nu\triangle^2 u \,=\, \triangle(g(u)) \qquad \text{for } x \in \Omega \text{ and } t \geq 0. \tag{49}$$

To study the solutions of (49), one uses the Landau-Ginsburg functional:

$$J(u) := \int_\Omega \left[\frac{\nu}{2}|\nabla u|^2 + G(u)\right] dx, \qquad \text{where } G(z) = \int_0^z g(s)\, ds. \tag{50}$$

In addition, the Landau-Ginsburg functional satisfies

$$\partial_t J(S(t)u_0) \;=\; -\|\nabla K(S(t)u_0)\|_0^2, \qquad \text{for } 0 < t < T(u_0), \tag{51}$$

where $K(u) = -\nu\triangle u + g(u)$. Furthermore, there exist positive constants f_0 and C_0 such that $-f_0 \le G(s) \le C_0 s^4 + f_0$ and

$$\frac{\nu}{2}\|\nabla S(t)u_0\|_0^2 - f_0|\Omega| \le J(u_0). \tag{52}$$

Since $J(u_0)$ is bounded below, see (52), and decreasing along orbits, see (51), it follows that $J(S(t)u_0)$ is a Lyapunov function, see LaSalle and Lefschetz (1961). This implies that the mild solution $S(t)u_0$ is defined for all $t > 0$, i.e., $T(u_0) = \infty$, for every $u_0 \in V^2$. It is fact that, whenever $F \equiv 0$ and $\overline{u}_0 = 0$, then $\omega(u_0)$, the omega limit set of the solution $S(t)\, u_0$ is a nonempty, compact, connected invariant set in Q_0.

Let us now return to the CHE with forcing (43). By integrating the equation (43), where $u_0 \in H_0^2$ and $F \in H_0^2$, as well, then $S(t)\, u_0 \in H_0^2$, for all $t \ge 0$.

The global attractor $\mathfrak{A}_0$. Assume that B satisfies the boundary conditions BC, i.e., either the non-flux condition (45) holds, or the periodic boundary conditions hold. Let $S(t)$ be the semiflow on V_0^2 described above. Then the following hold.

1. $S(t)$ has a nonempty, compact global attractor $\mathfrak{A}_0$ in V_0^2, and $\mathfrak{A}_0$ attracts all bounded sets in V_0^2. The attractor $\mathfrak{A}_0$ depends continuously on the forcing function $F \in H_0^2$.

2. When $F \in H_0^3$, then the attractor $\mathfrak{A}_0$ is a compact, invariant set in V^{4r}, for each r with $0 \leq r \leq 1$.

3. The set Q_0 is nonempty, compact, and invariant with $Q_0 \subset \mathfrak{A}_0$.

4. Lastly, there is an Inertial Manifold for the solutions of the infinite dimensional system (43), see Foias et al. (1988) and Sell and You (2002). One finds this manifold by using the orthogonal projection $\mathbb{P}_N$ onto the lowest N nodes, that is to say, into Span$\{e_k : k \leq N\}$, where $\{e_k\}$ are the eigenfunctions for the Bi-harmonic operator B. One then makes a change of variables $u = v + w$, where $v = \mathbb{P}_N u$, $w = \mathbb{Q}_N u$, and $\mathbb{Q}_N = I - \mathbb{P}_N$. The (v, w) system:

$$\begin{aligned} \partial_t v + Bv &= \mathbb{P}_N \left(\triangle(g(v+w)) + F \right) \\ \partial_t w + Bw &= \mathbb{Q}_N \left(\triangle(g(v+w)) + F \right), \end{aligned} \tag{53}$$

is equivalent to the CHE (43). One then shows that, for N large, the variable w is enslaved to the variable v in some neighborhood of $\mathfrak{A}_0$. That is to say, $w = \Phi(v)$, for a suitable function Φ. It turns out the the longtime dynamics of (43) is equivalent to the longtime dynamics of the finite-dimensional ordinary differential equation:

$$\partial_t v + Bv = \mathbb{P}_N \left(\triangle(g(v + \Phi(v))) + F \right). \tag{54}$$

See Sell and You (2002) and the references contained therein, for more details.

The finite-time Lyapunov vector algorithm for the CY92 model is derived, by first rewriting the equation as

$$\partial_t S = \alpha \partial_{xx} \left[\mu^2 S \left(S - \eta \right)^2 - r f(x) + S - \gamma^2 \partial_{xx} S \right] =: F(S).$$

We opt here to first linearize and then discretize. To obtain the tangent linear equation, we need to find the linear map $\mathcal{L}$ such that

$$F(S+\delta S) - F(S) = \mathcal{L}(\delta S) + o(\delta S) \text{ as } \|\delta S\| \to 0,$$

where $\|\cdot\|$ is a norm. Typically, this will be the norm in the space where solutions live, such as the Sobolev space H^1 for S, or dictated by physical considerations. In this infinite-dimensional model, different norms are not equivalent.

First, we note that

$$\{S+\delta S\}\left(\{S+\delta S\}-\eta\right)^2 - S\left(S-\eta\right)^2 = \left(S-\eta\right)\left(3S-\eta\right) + o(\|\delta S\|).$$

We let L be a finite-difference approximation to $\mathcal{L}$. We will employ second order finite-difference discretizations in space. The discrete tangent linear equation becomes the product of two matrices AB, where B is the matrix whose entries are found by discretizing the operator

$$\mu^2\left(S-\eta\right)\left(3S-\eta\right) + Id - \gamma^2\partial_{xx},$$

where Id is the identity matrix, and A is the matrix corresponding to the discretization of $\alpha\partial_{xx}$. The finite-time Lyapunov algorithm is obtained by solving

$$\partial_t\delta\Theta = L\,\delta\Theta = AB\,\delta\Theta = A(B(\delta\Theta)),$$

subject to some initial vector perturbation $\delta\Theta(0) = s$, where $\delta\Theta$ is the discretization of the perturbation δS.

In the following simulations, the time step is $\delta t = 0.01$. In the example calculations that follow we applied second-order centered finite differences in space, and used 121 grid points, $-\pi = x_0, x_1, ..., x_{120} = \pi - \delta x$.

By construction the finite-time Lyapunov vectors are not amplitude sensitive, however the BVs are, thus different amplitude perturbations will yield different BVs, in the case of a general nonlinear problem. This outcome has important practical implications, if one would like to use BV to either infer the structure in the field, or the degree of sensitivity of the outcomes to perturbations in initial conditions. A challenging problem could thus arise in the context of large-scale simulations: what is considered a large structural change or a highly sensitive outcome is physics-dependent, perhaps even difficult to surmise quantitatively; the physics in question may not be fully understood and thus the a reasonable perturbation amplitude is simply guessed.

We ran the finite-time Lyapunov vector algorithm and the BV, using the same initial condition, forcing and perturbation. First we examine the effect of the size of the amplitude of the perturbation on the outcomes: we do so by keeping the shape of the perturbation fixed, changing only the overall amplitude. Figure 7a and c show the final perturbed minus the non-perturbed fields, when the perturbation is maximally 0.25 and then ten times smaller. Figure 7b and d are plots of the final BVs and finite-time Lyapunov vectors, corresponding to these size perturbations. When the perturbations are small the BVs and the finite-time Lyapunov vectors are qualitatively consistent and, nearly so, quantitatively. The outcome shown here is very typical: the qualitative and quantitative disagreement grows with increased in the amplitude of perturbations.

The shape or spectral content of the perturbation mattered as well. The spectrum of the perturbation is clearly important when projecting onto spectral bases for a reduced representation. We used the CY92 model, with

the same forcing as before and same initial condition. We examine next monochromatic sine wave perturbations with wavenumber $j = 1, 2, ..., 6$ of the form

$$\delta Y_0 = \epsilon_j \sin[jx + \frac{1}{3}(j-1)\exp(1)], \tag{55}$$

where all of the perturbation amplitudes remain the same, $\epsilon_j = 0.25$. In Figure 8a we show the BVs associated with each of the perturbations ($j = 1, 2..., 6$), at $t = 70$. Figure 8b shows the EBVs at $t = 70$. Even at $t = 5$, shown in Figure 8c, we already see a height-ordered structure in the EBV vector ensemble. The space-time plot of all EBVs is plotted in Figure 8d. Evident is the change in structure, after about time $t = 4$. The EBV calculation was run with identical parameters to those used in Figure 8a, with the ensemble consisting of the same initial perturbations in (55). In the BV case we are getting outcomes that do not have the reductive appearance of the EBV; there is no ambiguity to the prevailing ensemble member in the EBV case. This ensemble member has a clear correspondence in structure to the path and its perturbation field. The BVs should eventually agree with the EBVs, nevertheless, it is safe to assume that this will happen only after a very long time, longer than the time interval that might be suggested by the base solution.

We also compared BV, for each of the perturbations in (55), to the outcomes of the finite-time Lyapunov vector algorithm. Figure 9 compares the BVs and finite-time Lyapunov vectors for each j wavenumber perturbation. The amplitude ϵ_j of each of the perturbations was 0.25. The structure of the finite-time Lyapunov vectors, to within a sign, is qualitatively the same, regardless of the frequency of the perturbation. The BVs are not. If the

perturbation was made considerably smaller the differences between the BVs and the finite-time Lyapunov vectors would become small, as expected: As shown in Section 3, the BVs and the finite-time Lyapunov vectors must be akin to each other, provided the perturbations are small enough. For larger amplitudes, the BVs might show more structure as nonlinear effects may play a role in the structure of the BVs. Apparently, $\epsilon = 0.25$ is already in the range of large perturbations of the CY92 about the solution chosen. Whether this was a large perturbation or a small perturbation experiment was only clear to us after comparison to the finite-time Lyapunov vector outcomes and not from the dynamics of CY92.

For the CY92 model runs we can conclude thusly: for sufficiently small perturbations and renormalizing time intervals the finite-time Lyapunov vector, the BV algorithm, and the EBV are similar in outcomes. As we depart from these conditions, we see differences in the outcomes of the three methods. The BV outcomes are most sensitive to the amplitude and the frequency of the perturbations.

4.2.1 BV and EBV in the Higher Nonlinear Regime

We revisit the CY92 simulations of BV and EBV, using the same forcing and initial conditions used Section 4.2 but set the time scale parameter $\alpha = 0.01$, increase the spatial resolution to 320 points, and set the integration time step at $\delta t = 0.000015$. The perturbation field remains structurally the same as in (55), in what follows, however, we will increase the size of the perturbation of each of the components. For $\epsilon_j = 0.6, 0.8$, and 1.2, with $j = 1, 2, ..., 6$, in (55) we obtain Figure 10, which shows all BVs at $t = 19.025$. We expect a simple structure in the perturbation field and thus the BV should reflect this. For

comparison, the EBV results appear in Figure 11. Comparison of Figures 10 and 11 show how the BV outcomes look qualitatively different from their EBV counterparts as we increase the initial amplitude of the ensemble. The comparison is further aided by reference to Figure 12, in which the each ensemble member of the EBV has been rescaled to L^2-norm of 1. The EBV yields 3 distinct types of perturbations, and the picture remains robust to the increase in the initial perturbation size. The outcomes are ordered in size, as usual. Among all the vectors present in BV outcomes, one finds its way to the EBV final result, all the others are suppressed, in the EBV. The BV outcomes practically lose track of, and coherency with, the perturbation field if we consider the dominant perturbations with size less than or equal to the initial perturbation size parameter as the desired information. We can go a little further and inquire what structure is obtained in the smaller EBV ensemble members. To sumarize the outcomes, when the nonlinear effects cannot be neglected, the algorithms differ essentially and they diverge in their results. Our findings favor the EBV algorithm, because it is much more robust when nonlinear effects are nonnegligible. How these outcomes generalize to arbitrary dynamics and equations is not clear, but suggests a good issue for future systematic study.

5 The Multiplicative Ergodic Theorem

As noted above, the second milestone in the theory of Lyapunov exponents/vectors is the MET (Multiplicative Ergodic Theorem) in 1968 when the proofs of both Oseledec (1968) and Millionščikov (1968). Also the work of Johnson et al. (1987), which contains a study of the links between the Lya-

punov Spectrum, the Sacker-Sell Spectrum, and various ergodic theoretic properties, such as ergodic measures and/or invariant probability measures of the nonlinear flow on the attractor of the nonlinear problem. The Sacker-Sell Spectrum, which uses the theory of exponential dichotomies, leads to a "continuous" splitting of the vector bundle generated by the Tangent Linear Equations, i.e., the nonautonomous vector spaces generated by (8), while the MET leads to a "measurable" refinement of this continuous splitting, as is noted in Remark 4.2.9 on pages 177-179 in Arnold (1998). The latter reference is of special note because it contains various extensions of the MET to problems not originally envisioned in the pioneering works of an earlier generation.

Based on the presentation in Johnson et al. (1987), we continue here the Lyapunov theory begun in Section 2.2. We let M denote a compact, invariant set in the attractor $\mathfrak{A}$. For the MET, we use **PRB**, the collection of invariant probability measures on M, and **ERG**, the collection of ergodic measures on M, with **ERG**$\subset$**PRB**. (Recall that a probability measure μ is non-negative and satisfies $\mu(M) = 1$.) We set $p = (k; \widehat{m})$, where k is an integer with $1 \leq k \leq K$, and $\widehat{m}$ is a vector with $\widehat{m} = \{m_1, \cdots, m_k\}$, such that $m_j \geq 1$ is an integer, for $1 \leq j \leq k$, with $\sum_{j=1}^{k} m_j = K$.

Next we return to the Tangent Linear Equation described in Section 2.2. We let $U(\theta, t)$ denote the solution operator for the linear equation (8), where $\theta \in \mathfrak{A}$. Let μ be a probability measure in **PRB**. The MET states that the following hold:

1. there is a measurable set $M_\mu \subset M$ with $\mu(M_\mu) = 1$ and one has $M_\mu = \cup_p M_\mu(p)$, for $p = (k, \widehat{m})$, where each $M_\mu(p)$ is invariant, such

that;

2. there are measurable mappings $\lambda_1, \cdots, \lambda_k : M_\mu(p) \to \mathbb{R}$ such that

$$\lambda_k(\theta) < \cdots < \lambda_2(\theta) < \lambda_1(\theta), \qquad \text{for } \theta \in M_\mu(p),$$

3. and there are measurable mappings: W_j from $M_\mu(p)$ into the collection of linear subspaces of $\mathbb{R}^K$, with

$$\dim W_j(\theta) \,=\, m_j, \qquad \text{for } \theta \in M_\mu(p).$$

Furthermore, one has

$$U(\theta, t)\, W_j(\theta) \,=\, W_j(\theta \cdot t), \qquad \text{for } (\theta, t) \in M_\mu(p) \times \mathbb{R}, \qquad (56)$$

and $1 \leq j \leq k$. In addition, for $\theta \in M_\mu(p)$, one has:

4. the family $\{W_1(\theta), \cdots, W_k(\theta)\}$ is linearly independent;

5. $\mathbb{R}^K \,=\, W_1(\theta) + \cdots + W_k(\theta)$, for $\theta \in M_\mu(p)$;

6. (Strong Lyapunov Exponents): If $x \in W_j(\theta)$, $x \neq 0$, then $\lambda(\theta, x) = \lambda_j(\theta)$, for $1 \leq j \leq k$, i.e., $\lambda_j(\theta)$ is a strong Lyapunov exponent, and it does not depend on x.

7. Moreover, if $\mu \in \mathbf{ERG}$, then precisely one $M_\mu(p)$ has positive measure, with $M_\mu(p) = M_\mu$, and the term k, as well as, the mappings $\lambda_j : M_\mu \to \mathbb{R}$ are constant, for $1 \leq j \leq k$.

The relation (56) implies that each measurable vector bundle

$$\{(\theta, x) \,:\, x \in W_j(\theta),\, \theta \in M_\mu(p)\}$$

is invariant under the skew product flow $\Pi(t)$, for each $j \in \{1, \cdots, k\}$. The Lyapunov Spectrum LYΣ =LY$\Sigma(\mu, \theta)$ is the collection $\{\lambda_k, \cdots, \lambda_1\}$ in Item 2. This depends on the measure $\mu \in \mathbf{PRB}$ and $\theta \in M_\mu(p)$. When $\mu \in \mathbf{ERG}$, then LY$\Sigma(\mu)$ does not depend on $\theta \in M_\mu$, due to Item 7. above.

A related concept is the Sacker-Sell Spectrum (also known as the Dynamical Spectrum), which is denoted by SSΣ, and is defined to be the collection of $\lambda \in \mathbb{R}$ for which the shifted flow $U_\lambda(\theta, t) := e^{-\lambda t} U(\theta, t)$ fails to have an exponential dichotomy over M, (see Sacker and Sell (1978, 1980); Sell and You (2002)).

By using the spectral theory developed in Sacker and Sell (1978, 1980), one observes that the Spectrum SSΣ is the union of n-nonoverlapping compact intervals, SS$\Sigma = \cup_{i=1}^{n}[a_i, b_i]$, where $1 \le i \le n \le K$. For each spectral interval $[a_i, b_i]$, there is an invariant vector bundle

$$\mathcal{V}_i \,=\, \{(\theta, x) \,:\, x \in \mathcal{V}_i(\theta)\}, \qquad \text{for } 1 \le i \le n,$$

where $\mathcal{V}_i(\theta)$ is a linear subspace of $\mathbb{R}^K$ with $\dim(\mathcal{V}_i(\theta)) = n_i$, $n_i \ge 1$ is an integer, and $\sum_{i=1}^{n} n_i = K$. The family of spaces $\mathcal{V}_i(\theta)$, for $1 \le i \le n$, is linearly independent, and one has

$$\mathbb{R}^K \,=\, \mathcal{V}_1(\theta) + \cdots + \mathcal{V}_n(\theta), \qquad \text{for } \theta \in M.$$

Furthermore, one has

$$U(\theta, t)\, \mathcal{V}_i(\theta) \,=\, \mathcal{V}_i(\theta \cdot t), \qquad \text{for } (\theta, t) \in M \times \mathbb{R}. \tag{57}$$

Hence these subbundles are invariant under the skew product flow $\Pi(t)$. We assume that the spectral intervals are in order, from left to right, i.e., $a_1 \le b_1 < \cdots < a_n \le b_n$. More information about the consequences of

the absence of an exponential dichotomy can be found in the Alternative Theorem (see Sacker and Sell (1976a,b, 1994)).

There is an interesting relation between these two spectral concepts, when the invariant set M is "dynamically connected," which means that M cannot be written as the union of two disjoint, nonempty, closed, invariant sets. In this case, one finds that

$$\text{boundary SS}\Sigma \subseteq \bigcup_{\mu \in \mathbf{ERG}} \text{LY}\Sigma(\mu) \subseteq \text{SS}\Sigma. \tag{58}$$

The relation (58) contains some useful information concerning the Lyapunov exponents, see Section 2.2. Since $b_n \in$ boundary SSΣ, it follows from (58) that $b_n \in$ LY$\Sigma(\mu)$, for some $\mu \in$ **ERG**. Hence, Item 7 implies that $b_n = \lambda_1(\theta)$, for all $\theta \in M_\mu$. That is to say, b_n is the "largest" Lyapunov exponent associated with the ergodic measure μ. The same considerations apply for the boundary points b_i, for $1 \leq i \leq (n-1)$. Of course, the ergodic measures may change, as i changes.

In addition to the implication that $b_i \in$ LY(μ), for some $\mu \in$ **ERG**, one can show the following. The relation (58) implies that for every j with $1 \leq j \leq k$ (see Item 2 above), there is an i with $1 \leq i \leq n$, such that $W_j(\theta) \subseteq \mathcal{V}_i(\theta)$. Furthermore, one has $n \leq k$. Thus the measurable collection $\{W_1, \cdots, W_k\}$ is a refinement of the continuous collection $\{\mathcal{V}_1, \cdots, \mathcal{V}_n\}$. The Sacker-Sell spectral theory is a valuable tool for the study of the temporal exponential separation of solutions of the tangent linear model.

One can find the Lyapunov exponents $\lambda_j(\theta)$ and the multiplicities $m_j(\theta)$ by using the wedge products and taking suitable limits, as $|t| \to \infty$. See Theorem 2.4 in Johnson et al. (1987). One does not need to use the bases of

the subspaces $W_j(\theta)$ for this calculation.

6 Implementation Issues

The main thrust of our work has been to propose an ensemble based vector breeding algorithm, the Ensemble Bred Vector (EBV) algorithm. The first breeding algorithm, the Bred Vector (BV) algorithm was introduced by Toth and Kalnay (1993, 1997). We compared the EBV to the BV algorithm and the finite-time Lyapunov Vector algorithms. In the EBV, an ensemble of initial perturbations is bred concurrently and then rescaled by the size of the largest member of the bred ensemble. The uniform normalization of all the ensemble members after each cycle is the distinctive trait of the EBV algorithm, which leads to some profound differences when compared to the BV algorithm.

As expected, when initial perturbations are sufficiently small, the EBV, the BV, and the finite-time Lyapunov vector algorithms lead to similar results. We gave a theoretical justification of this phenomenon by looking at the corresponding time-continuum analogues of the BV and EBV algorithms. We rescale frequently: The algorithms and results are formulated, for simplicity, assuming that the rescaling is done after every discrete time step. We do not consider the variants with $O(1)$ rescaling, mainly because we do not run the algorithms for long time intervals, and thus, it is not clear what purpose a rescaling would serve.

While the BV algorithm has to be completed by a measure of growth, the ordering of the outcomes based on their growth characteristics is a built-in feature of the EBV: we do not consider growth rate measures *specific* to ei-

ther, or growth rate measures that intricately depend on the EBV or BV. In fact, a growth rate measure should be independent of how the final perturbation vectors are obtained, and it should be possible to obtain this growth rate information regardless of whether the vector is possibly an outcome of a particular algorithm. This is the only type of growth rate measure that could be used for comparison purposes. There is an obvious way to achieve this: one simply perturbs the initial data in (1), and studies the evolution of the perturbed versus unperturbed systems without rescaling, starting at a time t_0 until an appropriate time T. Here t_0 and T are the initial and final times and these are to a large extent, arbitrary. Though problem-specific, the resulting time series can be then interpreted in the most appropriate way. For linear problems, exponential growth/decay rate is a natural choice. For many interesting nonlinear problems, this may not be the case: The long term dynamics may not exhibit exponential growth, nor exponential decay, while the instantaneous or very short term dynamics may be riddled with transient effects even in linear problems. Hence, one size does not fit all.

The outcome of the BV algorithm depends on the choice of norm used for rescaling (see Riviére et al. (2008).), if the nonlinear effects are not negligible. To illustrate, we consider the 2-dimensional system

$$\begin{aligned} \frac{dX_1}{dt} &= X_2, \\ \frac{dX_2}{dt} &= -\sin(8\,X_1), \quad t > 0, \end{aligned} \tag{59}$$

subject to the (same) initial conditions $(X_1(0), X_2(0)) = (0.8, -1)$. Figure 13 shows the outcome of the BV for three different choices of norms, starting from perturbations of size no greater than 0.15. The norms we employ are all standard l_p norms. The resulting BVs reflect the characters of these

norms. That different outcomes are obtained is not surprising: the unit sphere changes shape depending on what ℓ_p norm is used. Figure13 is meant to convey that these differences can be qualitatively drastic.

In practice, norms are dictated by practical as well as theoretical considerations. Conservation laws provide guidance for the most appropriate norms for the given dynamics (for instance, in Rayleigh-Bérnard convection, the temperature enjoys a maximum principle, while the velocity does not), but other considerations will play a role (*e.g.*, a sup-norm may be an obvious choice in determining the location of severe weather events). It will not be uncommon, thus, that a mixture of norms may be necessary in multi-physics problems.

One may ask if it is generally possible to dispense with the restrictions of the numerical integration scheme at least to compute the perturbed field. If we compute the base trajectory accurately, maybe, due to frequent rescaling, we can get away by using larger time steps or less accurate procedures. This is probably true for certain cases, but in general it is not so. The problem manifests itself in various forms and with various symptoms. As we saw for the CY92, when the time resolution is not high enough and just near the borderline, even when the base trajectory is computed more or less accurately, the finite-time Lyapunov vector, BV and/or EBV computations fail and output completely spurious vectors. In Lorenz63 computations, with time step $\delta t = 0.02$, we report that a mixture of spurious and accurate vectors were obtained when we ran either the BV and EBV algorithms on the ensembles we presented in this paper. However, with small enough time steps, this problem disappeared in both the BV and EBV calculations.

In the case of chaotic systems like Lorenz63, there is another fundamental

limitation one should be aware of: due to very strong Lyapunov instability of the dynamics on the attractor, it is impossible to accurately integrate the base trajectory for long periods of time. Divergence from the base trajectory is reflected in the resulting bred vectors as clearly illustrated by the plots in Section 4. An implementation of EBV (or BV) thus requires careful testing and validation, as suggested above. Other numerical implementation issues are enumerated now.

In Figure 15 we show the calculation for the BV for the linear system

$$\frac{dX}{dt} = AX, \quad t > 0, \qquad X(0) = X_0. \tag{60}$$

Here, A is an upper-triangular Jordan-block matrix of dimension 5, where A has a single eigenvalue λ, which is repeated on the main diagonal, while the diagonal directly above the main diagonal has non-zero entries, which for simplicity we assume to be the same value ρ. We assume further that $\lambda = -1$ and $\rho = 1$. The general solution operator contains "generalized" eigen-solutions with growth rates $t^k\, e^{-t}$, where $0 \leq k \leq 4$. (In this case, the eigenvalue has (algebraic) multiplicity 5.) The solution operator for equation (60) is not normal. Even though the equation (60), per se, is not apt to arise in any application of the BV algorithm, we present a brief analysis of the discretized version of equation (60) to illustrate the difficulties of making good approximations when similar non-normal behavior does occur in an application. For this linear problem we expect to see, provided we take t sufficiently long to forget the transient, a very trivial outcome to BV, or EBV. We found, however, that the calculation of the BVs, using (3), blows up if we took t sufficiently long. We employed an integration time step of 0.001, a perturbation initially of magnitude 0.01. We performed the calculation in

double precision, but without regard to how numerical sensitivity is handled. The integration time step was 0.001, and though it is in the stable regime, for ERK4 for this problem, judging from a convergence test, it is close to being too large. In Figure 15a we show the BVs, *i.e.*, $\mathcal{Y}$, as a function of time. Up to $t = 30$ the calculated solution is qualitatively acceptable. Nevertheless the 2-norm of $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$ monotonically increases, as is shown in Figure 15b, even though all the factors $t^k\, e^{-t}$, for $0 \leq k \leq 4$ in the solution operator, do decay for $t > 4$.

Figure 15c shows the 2-norm of $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$, using a numerically well-conditioned implementation of the BV algorithm. The strategy used to obtain a well-conditioned outcome was to normalize the elements in the required subtraction before computing their difference. The well-conditioned calculation was capable of qualitatively good results, without further requiring a smaller integration time step. The example serves as a warning as well as a suggestion for future systematic study. In this specific problem EBV was found to be less numerically sensitive to outcomes than BV. The strategy adapted here to increase the numerical stability was by no means generally applicable to all problems, nor was it optimal.

With regard to computational cost and speed, both BV and EBV share the same implementation advantage: a minimal amount of coding in order to apply these to legacy codes. In present-day applications ensembles of asynchronous and independent BV are run. In the EBV the ensemble members need to be run concurrently. For small problems this extra cost is trivial to minimal both in code design and in computational complexity. On larger problems the EBV communication issue becomes a manageable challenge.

When comparing BV and EBV, the extra cost of the latter, if any, may

be offset by the outcomes: the EBV accomplishes what one would want from the BV: a finite-time, algorithmically-based procedure for assessing the structure of the perturbation field and estimates of sensitivity of perturbations to initial conditions, which is only moderately more computationally expensive than current implementations of the BV algorithm. In addition, when the model has an underlying reduced structure, the perturbation field is suggestive of this structure well captured by the dominant members of the EBV. In contrast to the BV, the EBV is much more robust under the nonlinear effects as illustrated in Section 4. This last aspect is very important and it should be studied in the context of more chaotic equations in fluid mechanics, meteorology, and geophysics.

7 Concluding Remarks

The original BV algorithm attracted interest among researchers because it offered an approach to approximating solutions of the tangent linear model, without first solving the computationally intensive U-equation, see 38. In Section 3, we develop a solid mathematical basis for both the BV and EBV algorithms and show that each algorithm results in good approximations of the solutions of the tangent linear model, when the step-size δt is small. In Table 1, we also give a numerical confirmation of this theory for both algorithms, but there is more.

One of the metrics shows a substantial advantage for the EBV over the BV, in the sense that the drop in the minimal EBV error is substanntially better than the corresponsing drop in the minimal BV error.

The Figures 3, 4, and 5, which describe properties of only the EBV, are of special note because they cannot be replicated by the presently constituted BV. In Figure 3 one finds a clear separation of the 3 Lyapunov vectors associate with the Lorenz model, and in Figures 4 and 5, one can view the delicate fractal behavior within the Lorenz global attractor.

Lastly, in Figures 10, 11, and 12, we examine a series of related test problems that depart from the tangent linear model. The point here is to get a comparison of the performance of the two algorithms, as one moves further into nonlinear regime. Once again, one sees a significant advantage of the EBV over the BV.

In conclusion, for the applications described in this article, we believe that the new EBV algorithm has been shown to be clearly superior to the traditional BV algorithm.

Acknowledgements

JMR was partially supported by NSF grant DMS-0335360. ALM was partially supported by NSF grant DMS-0708902, DMS-1009713 and DMS-1009714. NB, ALM, JMR, and GRS wish to thank the Institute for Mathematics and its Applications (IMA) for their support and hospitality. Research at the IMA is supported by the National Science Foundation and the University of Minnesota.

References

L. Arnold. *Random Dynamical Systems*. Springer, New York, 1998.

R. Buizza, J. Tribbia, F. Molteni, and T. Palmer. Computation of optimal unstable structures for numerical weather prediction models. *Tellus*, 45A: 388–407, 1993.

P. Cessi and W. R. Young. Multiple equilibria in two-dimensional thermohaline circulation. *Journal of Fluid Mechanics*, 241:291–309, 1992.

K. K. W. Cheung. Ensemble forecasting of tropical cyclone motion: comparison between regional bred modes and random perturbations. *Meteorology and Atmospheric Physics*, 78:23–35, 2001.

M. Corazza, E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt, and J. A. Yorke. Use of the breeding technique to estimate the structure of the analysis errors of the day. *Nonlinear Processes in Geophysics*, 10:233–243, 2003.

B. Deremble, F. DAndrea, and M. Ghil. Fixed points, stable manifolds, weather regimes, and their predictability. *Chaos*, 19:043109, 2009.

R. M. Errico and R. Lagland. Notes on the appropriateness of bred modes for generating initial perturbations used in ensemble prediction. *Tellus*, 51A:431–441, 1999a.

R. M. Errico and R. Lagland. Reply to: Comments on notes on the appropriateness of bred modes for generating initial perturbations. *Tellus*, 51A: 450–451, 1999b.

C. Foias, G. R. Sell, and R. Temam. Inertial manifolds for nonlinear evolutionary equations. *Journal of Differential Equations*, 73(2):309 – 353, 1988.

T. Gneiting and A. E. Raftery. Atmospheric science - weather forecasting with ensemble methods. *Science*, 310:248–249, 2005.

S. Hallerberg, D. Pazo, J. M. Lopez, and M. A. Rodriguez. Logarithmic bred vectors in spatiotemporal chaos: Structure and growth. *Physical Review E*, 81(6, Part 2), JUN 9 2010.

J. A. Hansen and L. A. Smith. The role of operational constraints in selecting supplementary observations. *Journal of the Atmospheric Sciences*, 57: 2859–2971, 2000.

R. A. Johnson, K. J. Palmer, and G. R. Sell. Ergodic properties of linear dynamical systems. *SIAM Journal on Mathematical Analysis*, 18(1):1–33, 1987. ISSN 0036-1410. doi: 10.1137/0518001. URL `http://dx.doi.org/10.1137/0518001`.

E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 2003.

J. P. LaSalle and S. Lefschetz. *Stability by Lyapunov's Direct Method with Applications*. Academic Press, New York, 1961.

E. N. Lorenz. Deterministic nonperiodic flow. *Journal of atmospheric Science*, 20:130–141, 1963.

A. M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3):521–790, 1992. Translated by A. T. Fuller

from Édouard Davaux's French translation (1907) of the 1892 Russian original, (historical introduction) by Fuller, a Smirnov, and the bibliography Barrett,

B. B. Mandelbrot. *Fractals: form, chance, and dimension*. W. H. Freeman and Co., San Francisco, Calif., revised edition, 1977. Translated from the French.

V. M. Millionščikov. Metric theory of linear systems of differential equations. *Mat. Sb. (N.S.)*, 77 (119):163–173, 1968.

F. Molteni and T. N. Palmer. Predictability and finite-time instability of the northern winter circulation. *Quarterly Journal of the Royal Meteorological Society*, 119:269–298, 1993.

M. Mu and Z. N. Jiang. A new approach to the generation of initial perturbations for ensemble prediction: Conditional nonlinear optimal perturbation. *Chinese Science Bulletin*, 53:2062–2068, 2008.

V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968. ISSN 0134-8663.

T. N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza. Singular vectors, metrics and adaptive observations. *Journal of Atmospheric Sciences*, 55: 633–653, 1998.

O. Perron. Über eine Matrixtransformation. *Mathematische Zeitschrift*, 32 (1):465–473, 1930a. ISSN 0025-5874. doi: 10.1007/BF01194646. URL `http://dx.doi.org/10.1007/BF01194646`.

O. Perron. Die Stabilitätsfrage bei Differentialgleichungen. *Mathematische Zeitschrift*, 32(1):703–728, 1930b. ISSN 0025-5874. doi: 10.1007/BF01194662. URL `http://dx.doi.org/10.1007/BF01194662`.

C. Primo, M. A. Rodriguez, and J. M. Gutierrez. Logarithmic bred vectors. a new ensemble method with adjustable spread and calibration time. *Journal of Geophysical Research-Atmospheres*, page D05116, 2008.

O. Riviére, G. Lapeyre, and O. Talagrand. Nonlinear generalization of singular vectors: Behavior in a baroclinic unstable flow. *Journal of the Atmospheric Sciences*, 65:1896–1911, 2008.

R. J. Sacker and G. R. Sell. Lifting properties in skew-product flows with applications to differential equations. *Memoirs of the American Mathematical Society*, 11(190):iv+67, 1977.

R. J. Sacker and G. R. Sell. A spectral theory for linear differential systems. *Journal of Differential Equations*, 27(3):320 – 358, 1978.

R. J. Sacker and G. R. Sell. The spectrum of an invariant submanifold. *Journal of Differential Equations*, 38(2):135 – 160, 1980.

Robert J. Sacker and George R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. II. *Journal of Differential Equations*, 22(2):478–496, 1976a. ISSN 0022-0396.

Robert J. Sacker and George R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. III. *Journal of Differential Equations*, 22(2):497–522, 1976b. ISSN 0022-0396.

Robert J. Sacker and George R. Sell. Dichotomies for linear evolutionary equations in Banach spaces. *J. Differential Equations*, 113(1): 17–67, 1994. ISSN 0022-0396. doi: 10.1006/jdeq.1994.1113. URL `http://dx.doi.org/10.1006/jdeq.1994.1113`.

G. R. Sell and Y. You. *Dynamics of Evolutionary Equations*. Springer, New York, 2002.

Z. Toth and E. Kalnay. Ensemble forecasting at NCEP: the generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317–2330, 1993.

Z. Toth and E. Kalnay. Ensemble forecasting at ncep: the breeding method. *Monthly Weather Review*, 125:3297–3318, 1997.

L. N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, Princeton, 2005.

A. Trevisan and R. Legnani. Transient error growth and local predictability - a study of the lorenz system. *Tellus Series A - Dynamics Meteorology and Oceanography*, 47:103–117, 1995.

X. G. Wang and C. H. Bishop. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences*, 60:1140–1158, 2003.

M. Z. Wei and Z. Toth. A new measure of ensemble performance: Perturbation versus error correlation analysis (PECA). *Monthly weather Review*, 131:1549–1565, 2003.

C. L. Wolfe and R. M. Samelson. An efficient method for recovering Lyapunov vectors from singular vectors. *Tellus*, 59A:355–366, 2007.

List of Figures

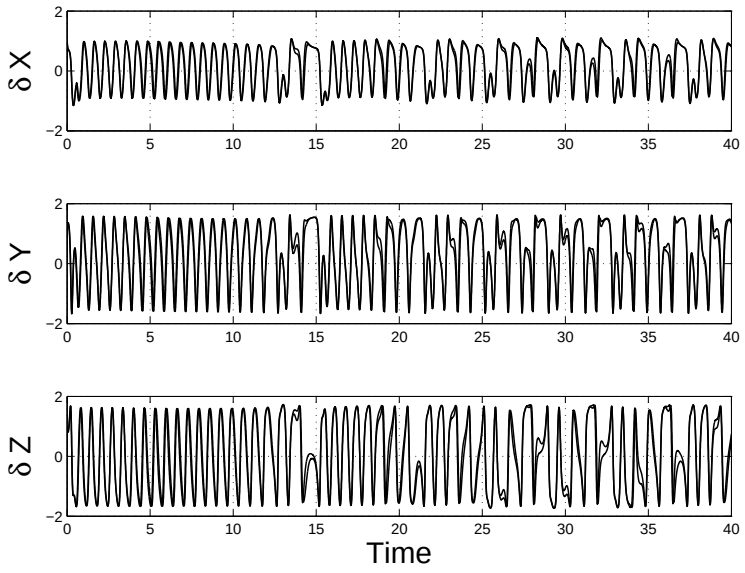


Figure 1: *The three components of BV and of the finite-time Lyapunov vectors, as a function of time, for the Lorenz63 system. (For parameters and initial conditions, see text). The vectors are nearly coincidental over the whole time span, even for large initial perturbations. The perturbation size is* 1*, in all of the components. See also Figure 2.*

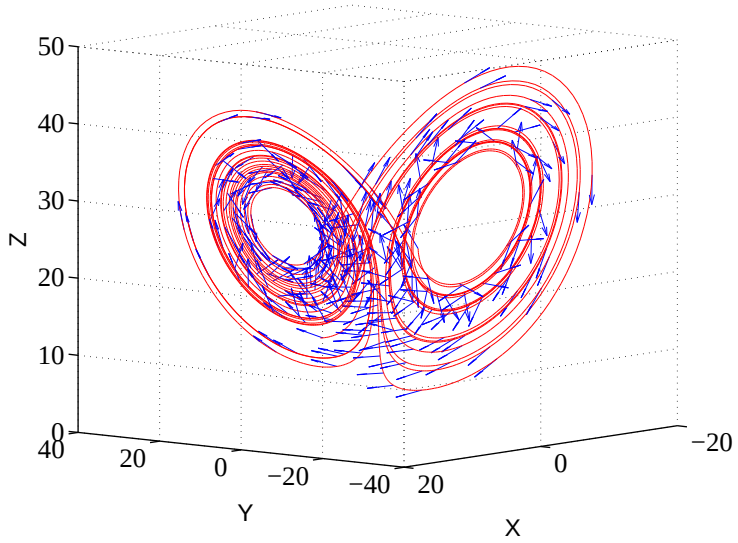


Figure 2: *For the time interval* $[20, 40]$*, the orbit. Superimposed on the computed solution, at* 0.2 *time step intervals, are the* $BV(t)$*. A superposition of the Lyapunov vectors on this plot would show very minimal differences when compared to the* $BV(t)$*. Same parameters, initial condition, perturbation as in Figure 1.*

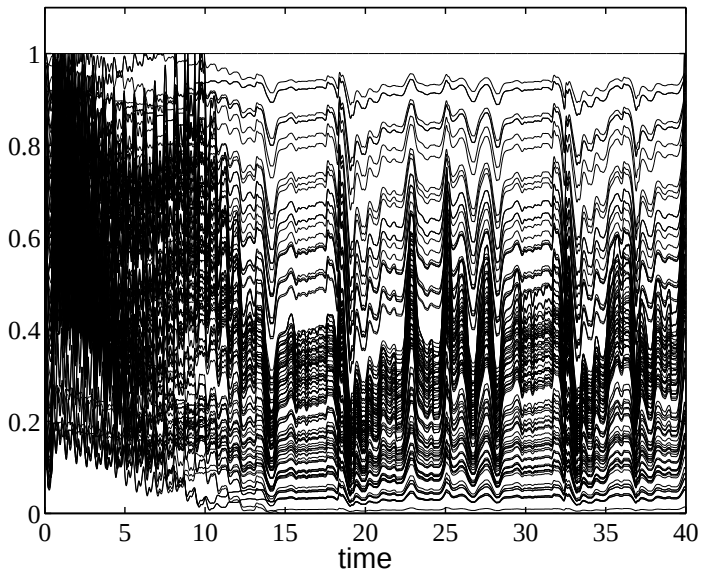


Figure 3: *Plot depicting the evolution of the 2-norm of an ensemble of 98 EBV*(t)*. The perturbation sphere has radius 1. Initially, in the interval* $[0, 13]$ *we note a very fast decay of some of the ensemble members, leading to a sorting in size, beyond that time. This outcome is one of the most distinguishing features of the EBVs when compared to an ensemble of individual BV outcomes. Same parameters and initial conditions, as in Figure 1 and Figure 2.*

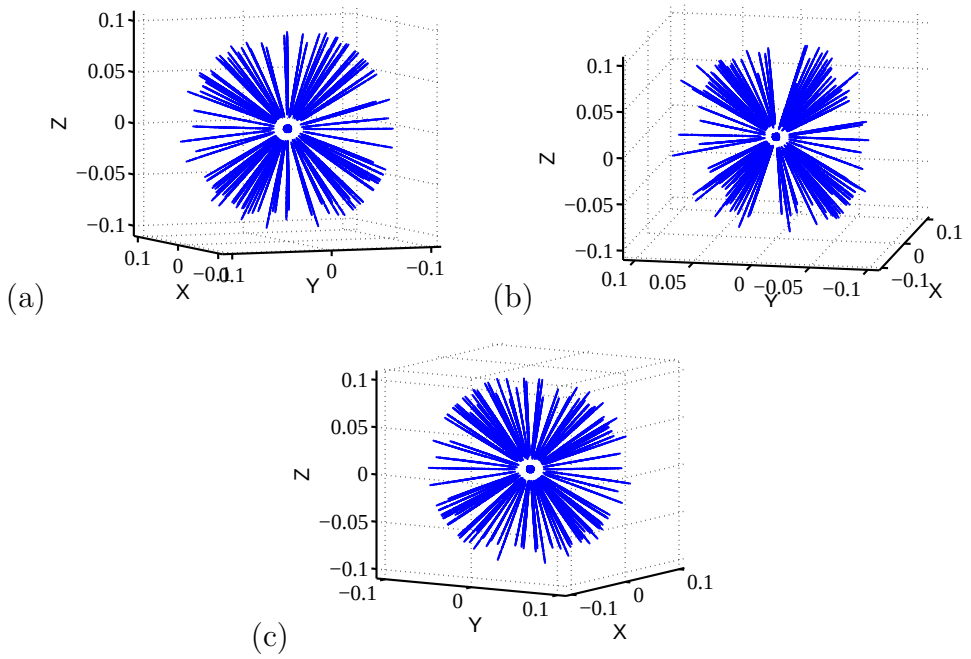
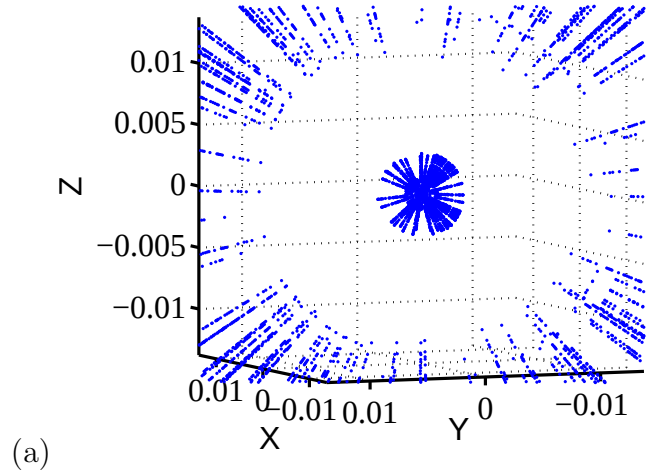
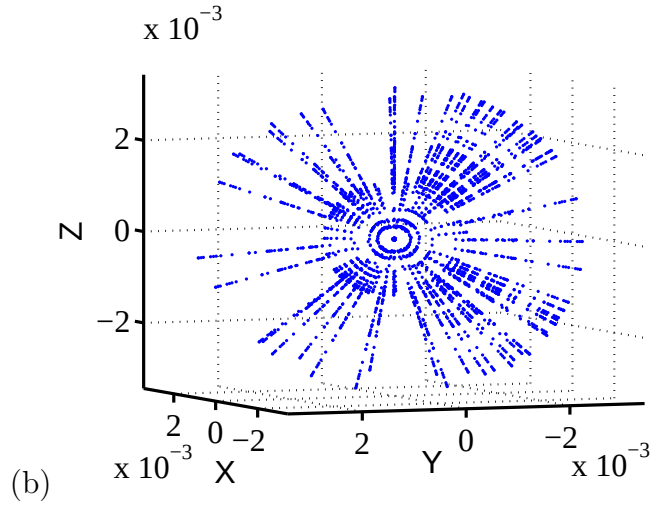


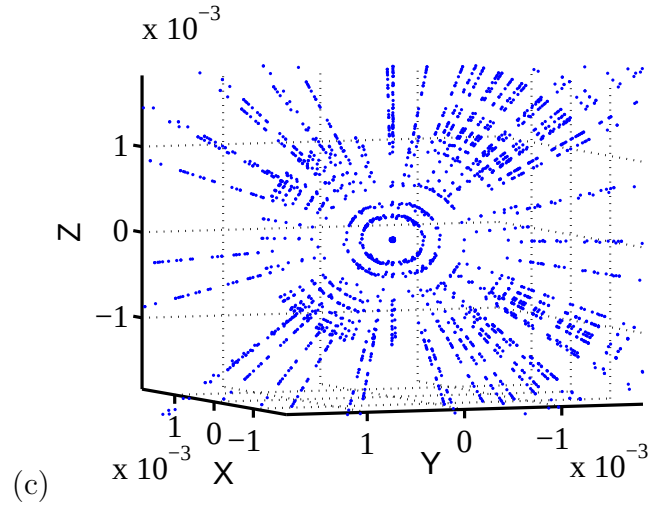
Figure 4: *Viewed from three different angles, (a)-(c). EBVs, for times* 24 *through* 30*, taken at* 0.1 *time intervals, for Lorenz63. Parameters and conditions are those used to generate Table 1.*



(a)



(b)



(c)

Figure 5: *Close-up views of Figure 4a: (a) zoomed in 8 times; (b) zoomed in 32 times; (c) zoomed in 60 times.*

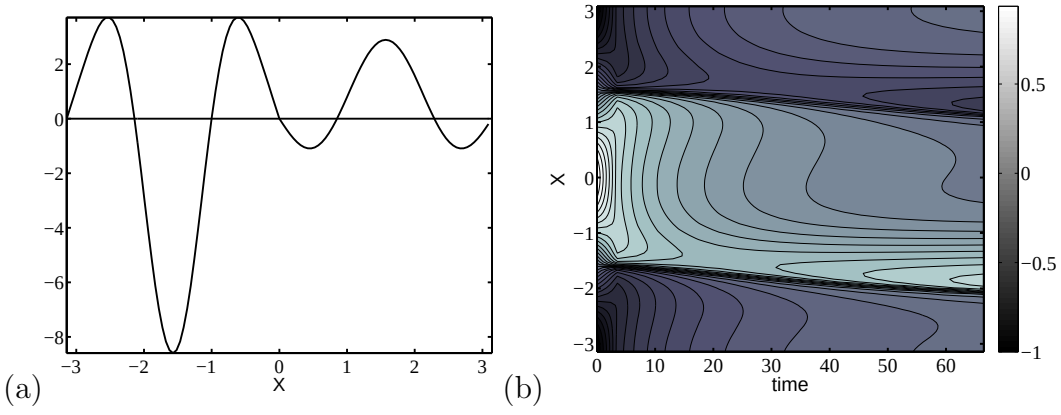


Figure 6: *(a) Forcing function, (42), used in the CY92 model; the numerical solution to CY92 is shown in (b). See Figure 7 for the BV algorithm and finite-time Lyapunov vector algorithm outcomes.*

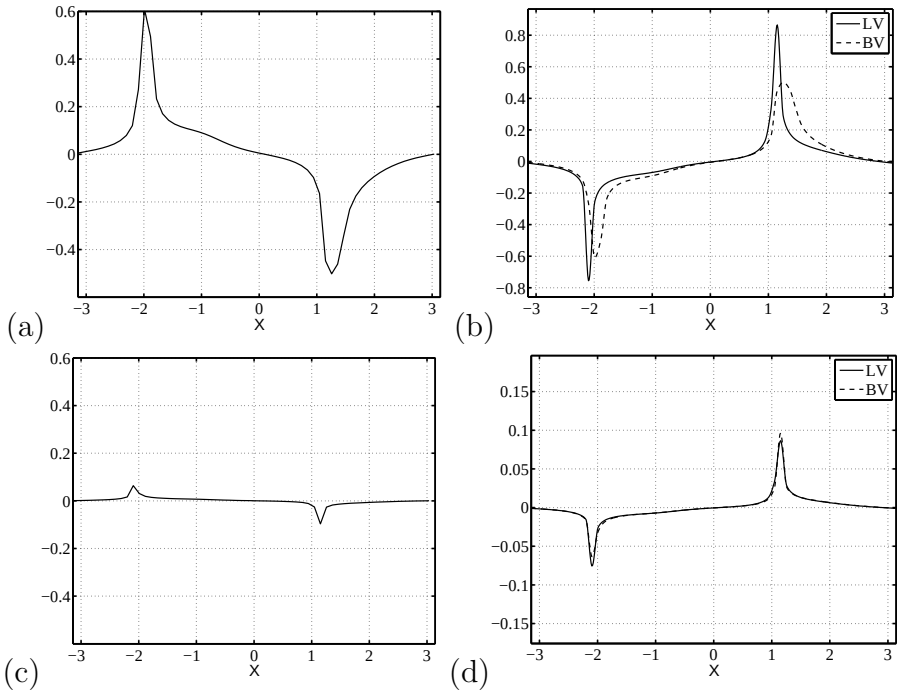


Figure 7: *Comparison of the finite time Lyapunov vectors and BVs. Effect of amplitude of perturbation on the CY92 Model, with initial conditions* $Y_0 = \cos(x)$. *The initial perturbation was* $\epsilon \sin(x)$. *At* $t = 70$, *(a) and (c)* $(Y + \delta\mathcal{Y}) - Y$. *(b) and (d) finite-time Lyapunov vectors and BVs. (a) and (b) computed using* $\epsilon = 0.25$*; (c) and (d) computed using* $\epsilon = 0.025$.

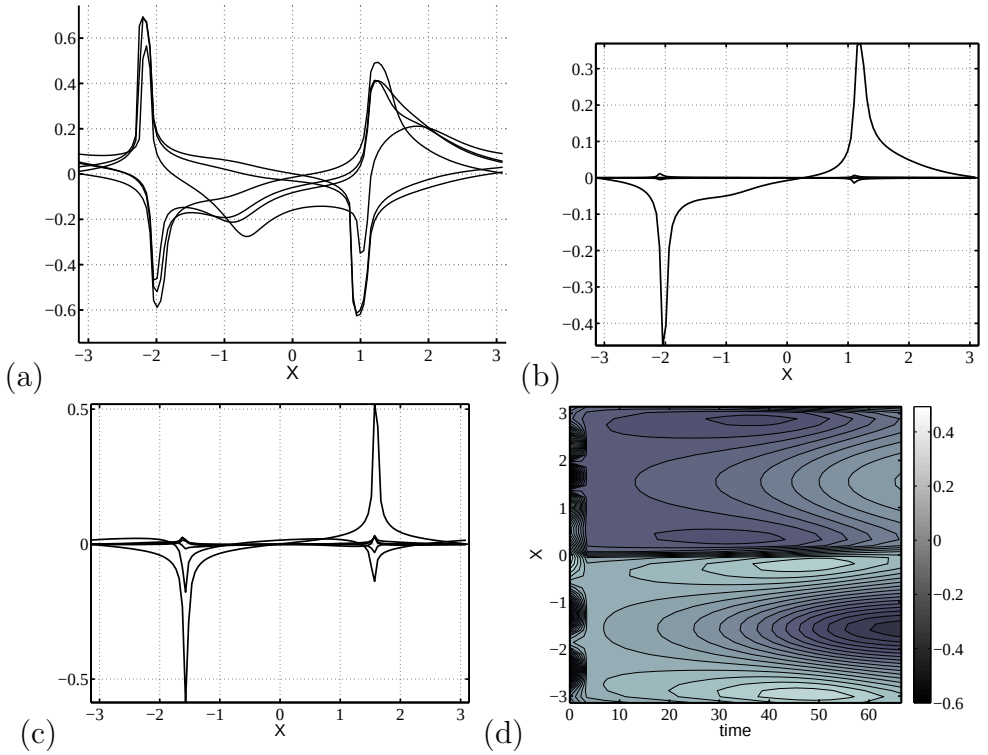


Figure 8: *(a) At $t = 70$, the superposition of 6 BV simulations, the j^{th} one corresponding to a sine wave perturbation as per (55). (b) Cross section of the EBVs, at $t = 70$, run with an ensemble of perturbations (55). It clearly showing a dominant vector. (c) Cross section of the same EBVs, at an earlier time: at $t = 5$; (d) Superposition of all EBVs, as a function of time and space.*

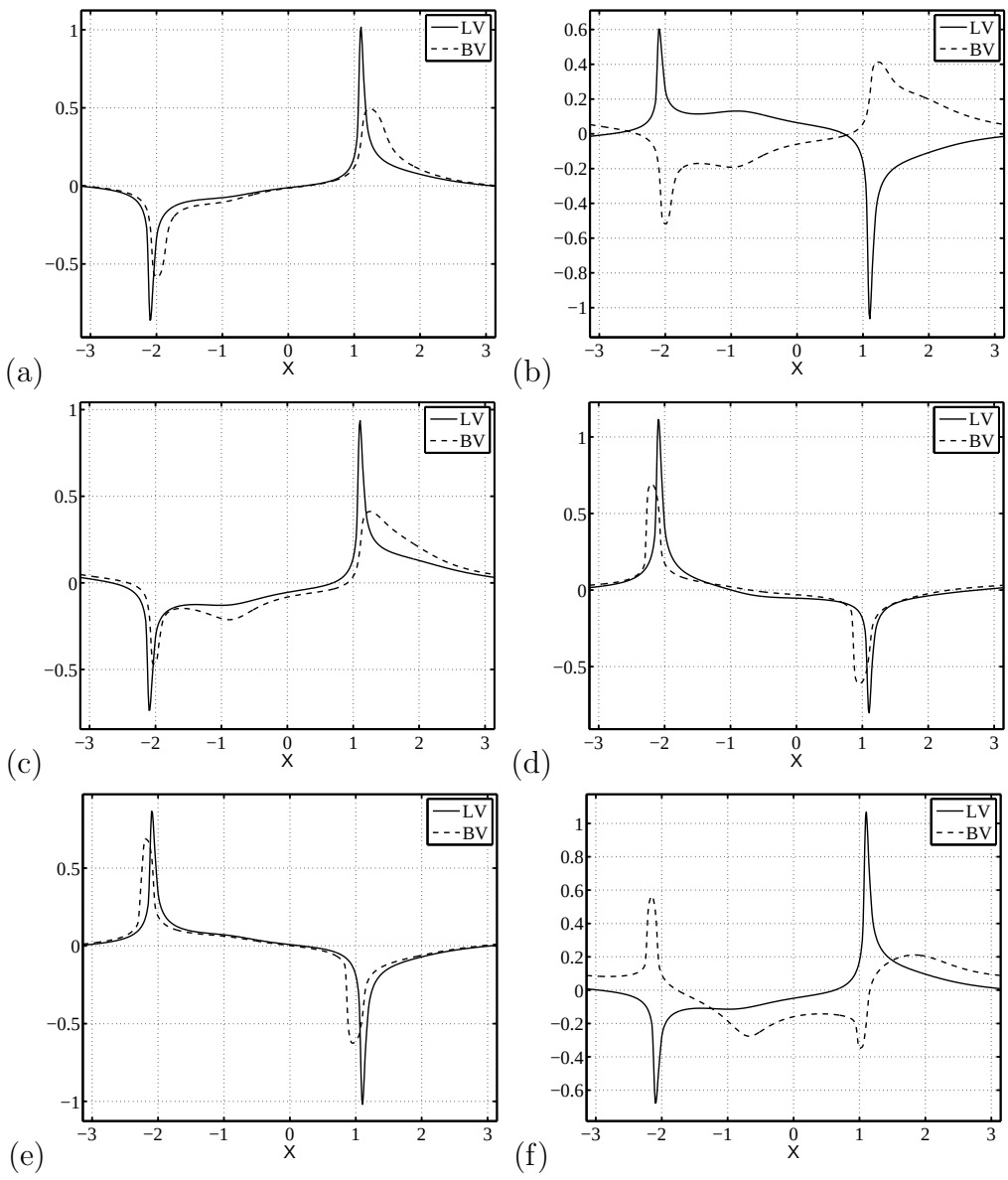


Figure 9: *Wavenumber dependence of BVs and finite-time Lyapunov vectors. CY92 Model, cosine initial condition and non-symmetric forcing (see Figure 6). Comparison of the finite-time Lyapunov vectors and the BVs for sine wave perturbations of size* $\epsilon_j = 0.25$. *The perturbations are the individual sine waves in (55). The correspondence is (a)* $j = 1$, *(b)* $j = 2$, *..., (f)* $j = 6$.

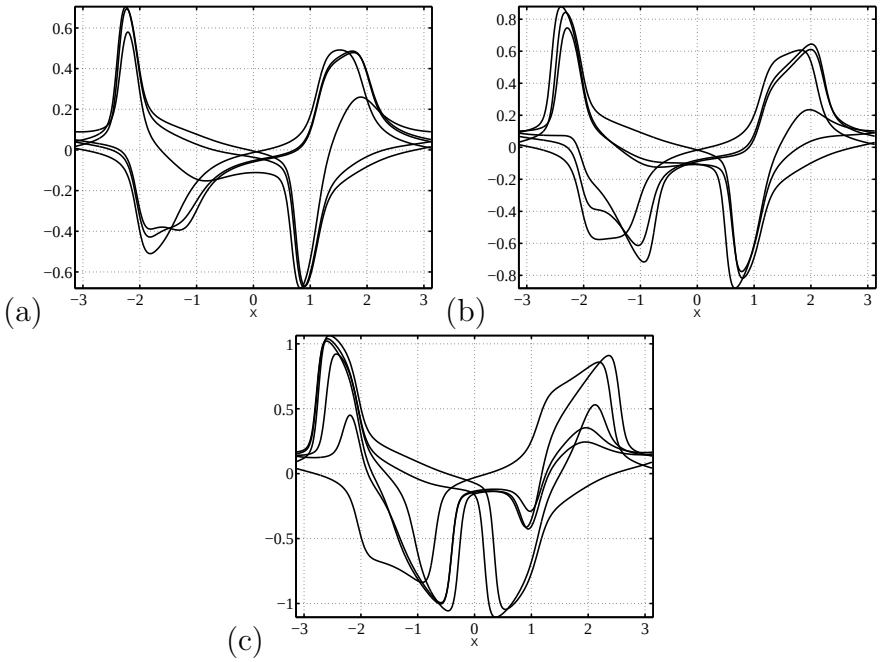


Figure 10: *BVs at* $t = 19.025$*, corresponding to* $\epsilon_j = \epsilon$ *set to (a) 0.6, (b) 0.8, (c) 1.2. The outcomes are very sensitive to nonlinearity and it is not clear what structural information about the perturbation field is contained in the results.*

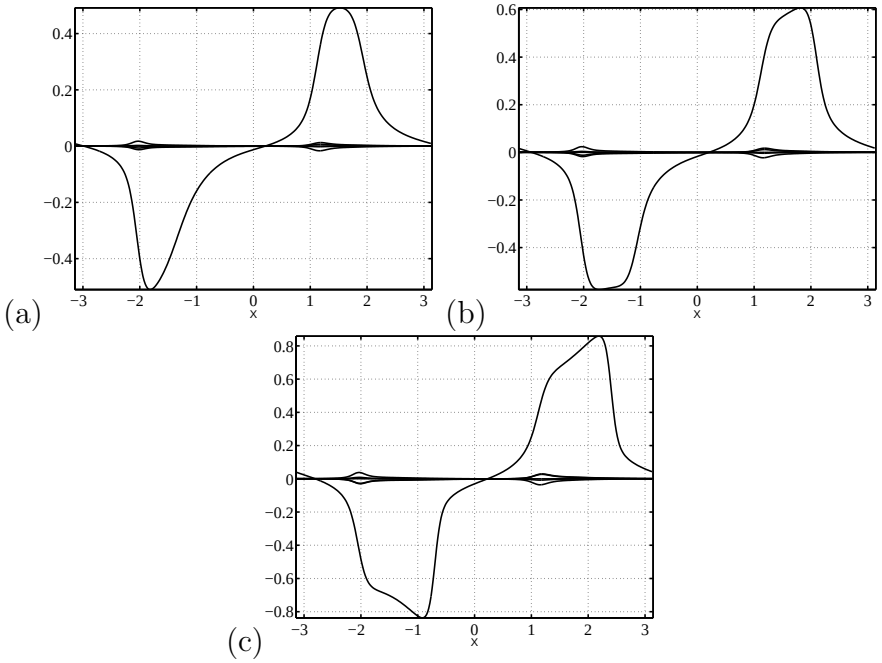


Figure 11: *EBV with ϵ set to (a) 0.6, (b) 0.8, (c) 1.2. Compare to Figure 10. The vectors are shown in their original scales.*

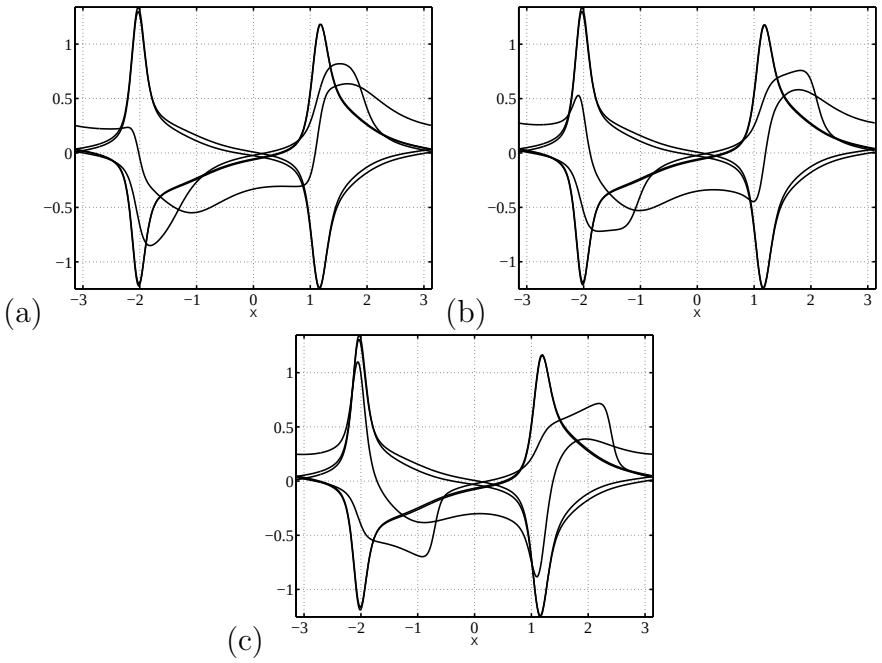


Figure 12: *EBV outcomes shown in Figure 11, with each vector is rescaled to L^2-norm 1 (only in plots), for visual comparison. Perturbation amplitudes (a) 0.6, (b) 0.8, and (c) 1.2. Note that as amplitude increases, the resemblance between the BV and the EBV outcomes, is lost almost completely, except for one of the vectors in this particular problem.*

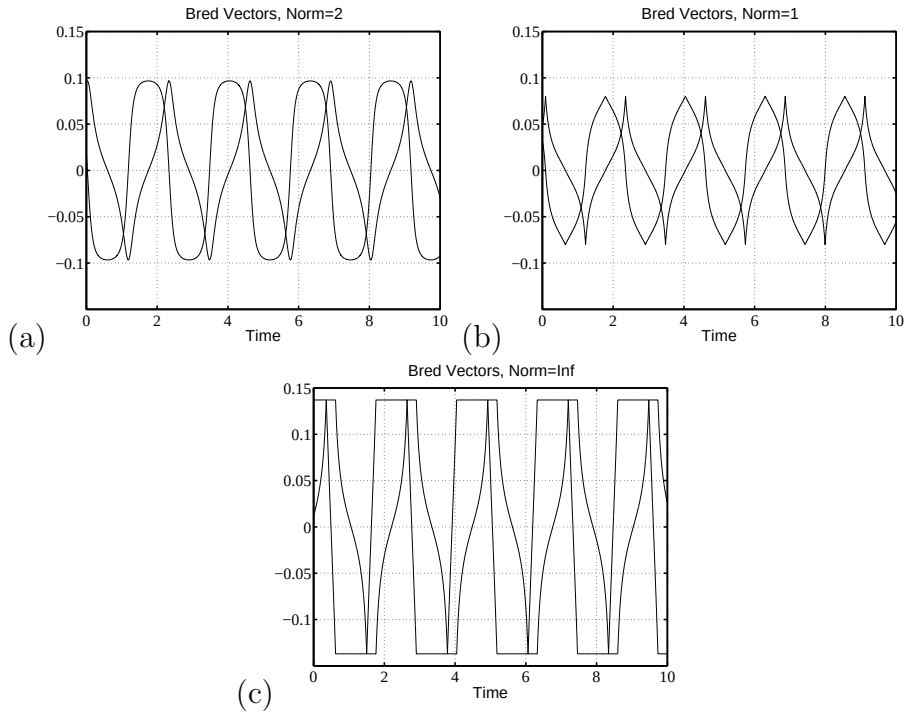


Figure 13: *(a) l_2-norm, (b) l_1-norm, and (c) l_∞-norm BV as a function of time, for the system in (59). The initial conditions are the same in all cases, the perturbations are no greater than* 0.15.

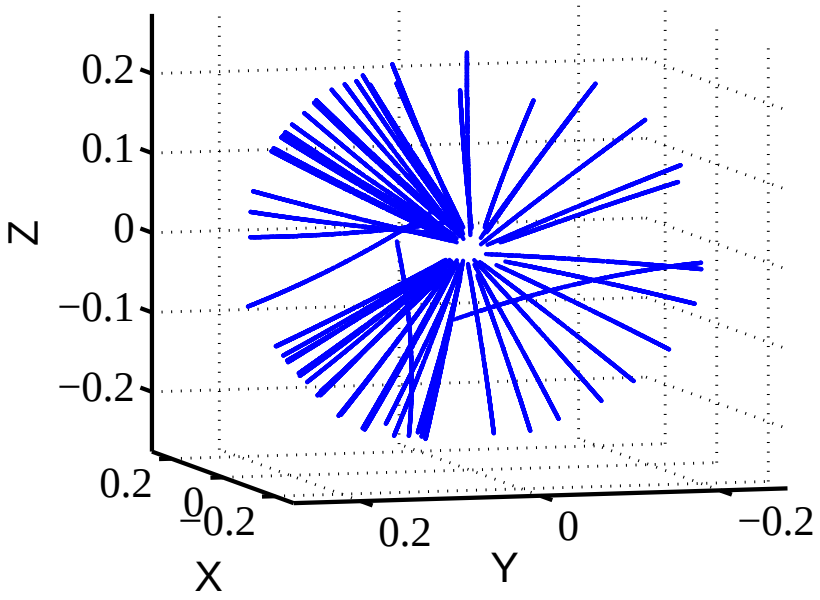


Figure 14: *Note the spurious EBV members with* $\delta t = 0.02$*, which were not present in Figure 5, as well as the qualitative overall differences.*

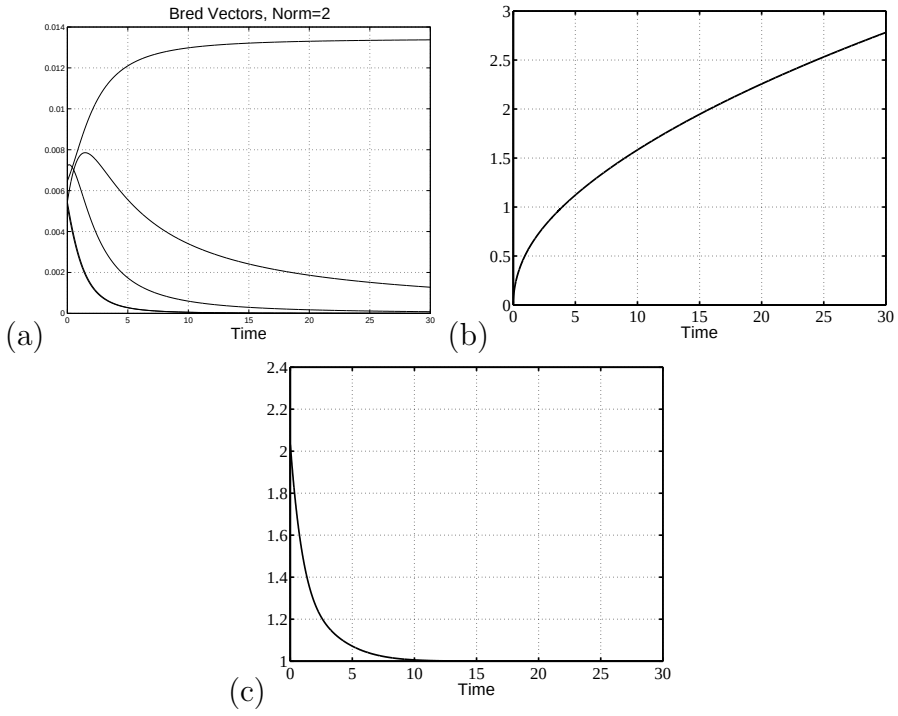


Figure 15: *(a) An ensemble of BV computations, as a function if time. The calculation diverges and eventually fails (not shown); (b) 2-norm of the numerically-approximated* $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$ *as a function of time, for the linear system* $\frac{dX}{dt} = AX$*. Here,* A *is an upper-triangular Jordan-block matrix of dimension 5, where* A *has a single eigenvalue* -1*, which is repeated on the main diagonal, while the superdiagonal has entries 1. (c) 2-norm of the numerically-approximated* $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$*, well-conditioned case.*