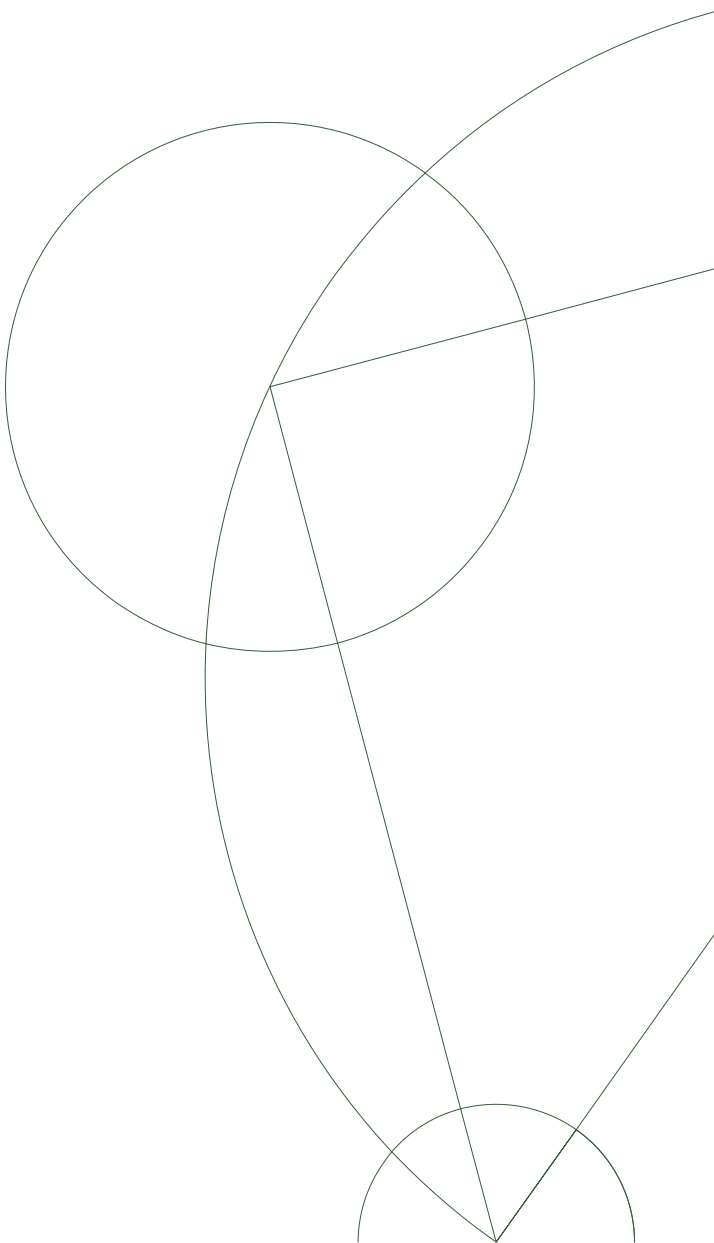
UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Perturbation Methods in Ensemble Weather Forecasts

Master's thesis in Physics



Martin Lyskjær Frølund
rhk101@alumni.ku.dk
Niels Bohr Institute, University of Copenhagen

Advisor: Peter D. Ditlevsen

Submitted: May 20, 2022

Abstract

The chaotic nature of the atmosphere makes weather prediction a fundamentally difficult task. Incomplete and uncertain knowledge of the current state of the atmosphere makes a prediction quickly diverge from the actual evolution of the atmosphere. For that reason, weather prediction centres perform ensembles of forecasts, all with different initial conditions, to estimate the uncertainty of their prediction. Various methods are used to perturb the initial conditions, the choice of which influences the uncertainty estimate. In this work we investigate seven different perturbation methods including the breeding vector (BV) method, the orthogonal breeding vector (BV-EOF) method, the Lyapunov vector (LLV) method, the singular vector (SV) method, and the random field (RF) method. We compare the methods theoretically and through numerical simulations with two low-dimensional non-linear models; the Lorentz-63 model and the Sabra shell model of turbulence. In evaluating the methods, we focus on the spatial/spectral distribution of the perturbations and the ability of the methods to produce large error growth relative to a reference. With the Lorentz-63 model, we find the largest exponential growth rates of the error with the SV method. In general, we observe good agreement with similar studies for all methods except the BV and RF method. The BV method produces larger growth rates than expected, which we attribute to differences in the method parameters. The RF method constructs perturbations based on historical states of the model but showed too small error growth rates. We attribute this to a difference in the procedure for sampling the historical states. For the Sabra shell model, we find the largest exponential growth rates of the error with the BV-EOF and the LLV method. Compared to those methods, we find low growth rates with the SV method. We argue that the reason lies in the sensitivity to numerical errors of the Lanczos algorithm, which we use in the SV calculation. Further work is needed to account properly for this sensitivity and to examine the performance of other algorithms.

Contents

List of Figures

List of Tables

Definitions

$\Re(x)$	Real part of x
$\Im(x)$	Imaginary part of x
$\mathbf{x}$	A vector
$\mathbf{A}$	A matrix
$\mathbf{x}^*$	Conjugate transpose
$\overline{\mathbf{x}}$	Complex conjugate
$\langle \mathbf{x}; \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y}$	The inner product between vector $\mathbf{x}$ and $\mathbf{y}$. The product is defined under the L_2 norm. When used to evaluate the projection of one vector onto the other, only the real part is used i.e. $\Re(\langle \mathbf{x}; \mathbf{y} \rangle)$
$\|\|\mathbf{x}\|\|$	The L_2 norm of $\mathbf{x}$.
$\delta_{i,j}$	Kronecker delta
$\langle x(y) \rangle_y$	The average of x with respect to y; e.g. time-average $\langle x \rangle_t$
$\iota = \sqrt{-1}$	The imaginary unit

1 Introduction

To introduce this thesis, we start out with a short outline of the history of weather forecasting. This will make the origin and relevance of ensemble forecasting clear. The question of how to start a weather forecast is then introduced along with the purpose of the thesis. In the end of the chapter the structure of the thesis is presented together with a short description of the code that has been developed.

1.1 Predicting the weather – from human- to supercomputers

The problem of predicting the future state of the atmosphere was first recognized and defined by Bjerknes, 1904 [2009] to be that of integrating a set of equations describing the evolution of the atmosphere from a set of initial and boundary conditions (Kalnay, 2002). The equations were the Newtons equations of motion, the equation of state, the continuity equation and the thermodynamic energy equation. Although the problem was now defined, a practical way of solving it did not materialize until 1910, where Richardson proposed solving the equations numerically (Richardson, 1922). By dividing the atmosphere into a grid Richardson made the first numerical 6-hour forecast of the atmosphere above central Germany the 20th of May 1910 [1]. The forecast utterly failed though[2], primarily because the initial conditions included fast moving gravity waves, which completely dominated the underlying meteorological signal (Kalnay, 2002). Despite the failure, the forecast

[1] Richardson, 1922 imagined gathering 64,000 people in a large hall to produce weather forecasts for the whole world - a kind of human weather forecast factory. Each human computer should solve the equations at their respective grid point and in the middle a director should orchestrate the factory. This fantasy bears remarkable similarities to the concept of a real computer although devised decades before the first research into computers.

[2] The forecast predicted a change in surface pressure of 146hPa in six hours, whereas the actual surface pressure did almost not change.

and Richardsons work laid the foundation for modern days numerical weather prediction (NWP).

Charney, 1949 and Eliassen, 1949 solved the problem with fast moving gravity waves by filtering the equations of motions using the quasi-geostrophic approximation. This led to the first successful one-day weather forecast performed on an electronic computer (Charney et al., 1950); successful because the forecast produced meteorological realistic results and correlated to some degree with observations. The filtering, however, permitted only the large scale atmospheric waves to evolve, which was a significant constraint. By lowering the duration of the time step in the numerical integration[3] J. G. Charney and J. C. Freeman (unpublished study; (Kalnay, 2002)) later performed a successful integration of the unfiltered equations, where the gravity waves died out through geostrophic adjustment[4]. This was an important discovery, because it showed that it was possible to predict the atmosphere with the more elaborate equations that permitted some small-scale meteorological phenomena.

Although J. G. Charney recognized that a limit to the predictability of the atmosphere existed, he attributed this to model deficiencies such as parametrization of processes happening on scales smaller than the grid size (e.g. turbulence; see sec. 2.3). By studying a simple three-dimensional deterministic model, Lorenz, 1963, however, discovered that small perturbations of the initial conditions could lead to diverging solutions to the model, which after some time were as different as random states of the model. In (Lorenz, 1965) Lorenz quantified the limit of predictability of the atmosphere to be about two weeks with a 28-variable atmospheric model. This showed that such a limit exists even for a perfect model and arise due to the model exhibiting *chaos* (ch. 2). Through this work, Lorenz also remarked that not only does such a limit exists, the predictability depend on the evolution of the atmosphere; some days the forecasts can predict accurately a week ahead, some days only a couple of days (Kalnay, 2002).

The value of the deterministic forecast was now questioned, and Epstein, 1969 recognized that a probabilistic approach could be fruitful. Instead of integrating the dynamical equations of motions from a set of observations, the observations should be used to define a probability density function (PDF) for the atmosphere. Integration of this PDF through the deterministic equations

[3]The Courant-Friedricks-Lewy (CFL) condition puts a requirement on the duration of the time step depending on the numerical scheme used (only for explicit schemes). If this requirement is not met, the solution will blow up (Kalnay, 2002).

[4]The process where a geostrophically unbalanced atmosphere is modified until in balance by radiation of gravity waves, which lead to adjustments of the wind and pressure fields (*Geostrophic adjustment* 2012; Holton et al., 2013).

of motions, would then give the evolution of the atmospheric PDF. The moments of this PDF would give estimates of the average evolution of the atmospheric state and the reliability of the forecast (from the variance). In that way the forecasters would not be forced to count on a single deterministic forecast, but get an indication of how probable different scenarios are.

The practical way of accomplishing this is by *ensemble forecasting*, which was first introduced operationally in (Molteni et al., 1996; Toth et al., 1993). Instead of integrating the atmospheric PDF into the future, which for a non-linear model can be cumbersome if even possible, the future PDF is sampled by running a finite number of deterministic forecasts all with different initial conditions, model configurations and/or boundary conditions (Molteni et al., 1996). How to perturb those conditions and configurations in order to sample the PDF properly is an art, and it influences the quality of the ensemble forecast.

No matter how it is done, ensemble forecasting has been and is a discipline for supercomputers due to the spatial resolution and the complexity of the atmospheric models. In that way, Richardson's fantasy of a human weather forecast factory has been fulfilled.

As we have seen, the discipline of predicting the weather has changed dramatically throughout the 20th century; from the first ideas by Bjerknes and Richardson to supercomputers running ensembles of forecasts. One of the main questions to address in relation to ensemble forecasting is:

1.2 How to start an ensemble weather forecast?

The quality of an ensemble forecast can be measured by comparing the average forecast (the *ensemble mean*) with the actual state of the atmosphere, e.g. as the root-mean-square-error (RMSE). The ensemble mean tend to improve the skill of the forecast compared to the individual forecasts (the *ensemble members*), since the uncertainty in periods where the individual forecasts disagree are averaged out[5] (Kalnay, 2002).

[5]Another way to understand the improved skill of the ensemble mean is by considering the properties of N dimensional phase spaces for large N. As described in (Christiansen, 2021), randomly drawn samples from an N-dimensional cube will tend to be located near the edges. The average of those samples (i.e. the ensemble mean), however, will be located in the centre of the cube. The distance between the average and a random sample will then be a factor $\sqrt{2}$ less than the distance between two random samples. For ensembles, this means e.g. lower RMSE of the ensemble mean relative to observations than any individual member.

The standard deviation of the ensemble members relative to the ensemble mean (the *ensemble spread*) gives a different indication of the ensemble quality. Since one seeks to sample the atmospheric PDF properly, the ensemble spread may indicate to what degree this is achieved as a function of forecast length. Since, however, a small/large spread may also reflect that the state of the atmosphere is predictive to a high/low degree, one often compare the RMSE and the ensemble spread vs forecast length. In general the ratio spread/RMSE should be ~ 1; if not, the ensemble is said to be *under-* (spread/RMSE < 1) or *overdispersive* (spread/RMSE > 1) (Fortin et al., 2014; Frogner et al., 2019).

An example of an ensemble forecast for the wind speed in Stockholm at the 925 hPa level in the atmosphere is shown in fig. 1.1. In the top panel the ensemble members (blue), ensemble mean (red), unperturbed forecast (green, dashed) and artificial observations[6] (cyan) are plotted. The RMSE and the ensemble spread is highlighted by the author (orange vertical lines). Note how the ensemble members and the unperturbed forecast agree well in the beginning 1-2 days after which the members spread more and more. The observations are drawn to show that an ensemble generally predicts well for a couple of days after which the prediction becomes less and less accurate. The PDF for the wind speed at day 4 calculated from the ensemble is shown in the middle panel. The cumulative PDF is shown in the bottom panel. From such a distribution one can get e.g. the probability for wind speeds larger than 15 m/s.

In addition to ensure a proper ensemble spread, the perturbation method used to perturb the initial conditions have to ensure that the ensemble samples the most unstable dynamics. This is because the unstable dynamics evolves fast and has a large impact on the evolution of the atmosphere (Molteni et al., 1996).

While Toth et al., 1993 initiated the ensemble by the so-called *breeding method*, which makes use of the full non-linear model, Molteni et al., 1996 used the *singular vectors* calculated from a linear approximation of the non-linear model (introduced in a meteorological context by Lorenz, 1965). As will be seen in ch. 4 there exists theoretical reasons for the latter method to be superior.

In a modern context, the advent of elaborate *data assimilation* methods[7] have introduced the *ensemble of data assimilations*, which is both used to estimate an uncertainty on the data assimilation and for generating initial

[6] Added to the figure by the author for the sake of example, i.e. not real observations

[7] The method of determining the (unperturbed) initial conditions of a forecast model from a limited number of typically non-uniformly distributed observations (Kalnay, 2002).

perturbations (Roberto Buizza et al., 2010; Isaksen et al., 2010). In terms of regional models, the COMEPS model, a regional variant of the HARMONIE Ensemble Prediction System (HarmonEPS; (Frogner et al., 2019)) used by the Danish Meteorological Institute (DMI), includes among others the possibility of calculating the perturbations from the difference between the members of a nested global ensemble model (the PertAna method). Another method available is the *random field* method originally introduced for ensemble forecasting in (Magnusson et al., 2009) (sec. 4.5).

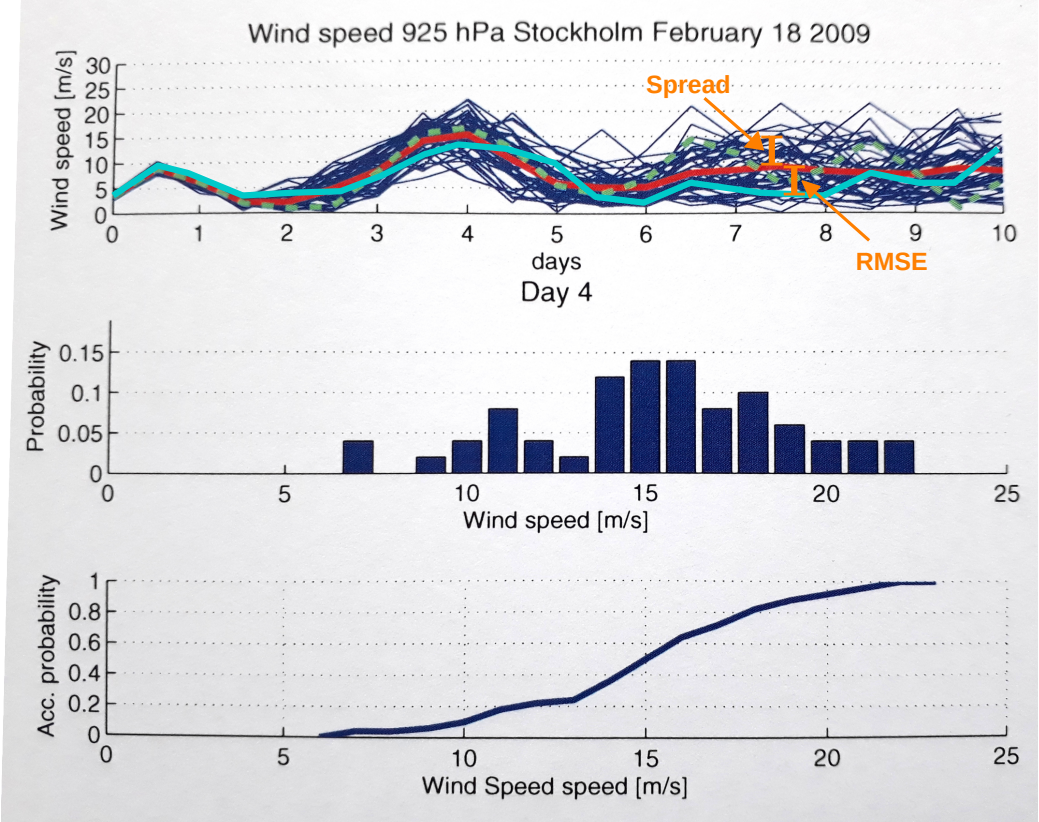


Figure 1.1: An example of an ensemble forecast for the wind speed in Stockholm at the 925hPa level in the atmosphere (top panel). Ensemble members (blue), ensemble mean (red) and unperturbed forecast (green, dashed). The PDF for the wind speed calculated from the ensemble at day 4 (middle panel). The cumulative PDF (bottom panel). From (L. Magnusson, 2009). The cyan line in the top panel represents artificial observations, while the spread and RMSE are highlighted with orange vertical lines. Those objects are added to the figure by the author for the sake of example (i.e. not real observations).

In this thesis, we investigate the question of starting an ensemble weather forecast by focusing on the process of perturbing the initial conditions. Thus, we will not deal with perturbations of the model configurations and boundary conditions. We study the original breeding and singular vector (BV, SV) methods as well as the orthogonal BV method (BV-EOF) and the newer random field (RF) method. We compare those methods with perturbations generated at random (RD), from an eigenvector analysis (normal mode (NM) method) and from the leading local Lyapunov vectors (LLV). The purpose is to present a detailed review of the methods and their behaviour when applied to two low-dimensional non-linear chaotic models; the famous Lorentz-63 model and the Sabra shell model of turbulence.

1.3 Structure of the thesis

This thesis is divided into eight chapters including the introduction. As a reference for all chapters specific mathematical definitions are listed in Definitions.

In ch. 2 fundamental concepts of chaos is introduced including a thorough outline of the theory of Lyapunov exponents and vectors. This serves as the basic theory needed in order to understand the behaviour of the two chaotic and non-linear models studied. The last section of the chapter presents the theory of turbulence upon which the Sabra shell model is introduced; ch. 3.

In ch. 3 the Lorentz-63 model is introduced as well, and for both models an analysis of the most relevant characteristics of the models is given. In ch. 4 six different methods for perturbing the initial conditions are presented. Since many of the methods are related to each other, the last section is devoted to an outline of the relationship between the singular vector method and the normal mode, breeding vector and Lyapunov vector method.

Ch. 5 presents the numerical method for solving both the non-linear models and the tangent linear models and their adjoints. The latter two are presented in chs. 2 and 4 and used in some of the perturbation methods (e.g. the singular vector method). The last two sections of the chapter presents the configuration details of the models and perturbation methods; e.g. model parameters, time step etc.

The results of the investigations with the different perturbation methods are presented in ch. 6 and discussed in ch. 7. In those chapters as well as others, we present the results, discussion etc. related to the models separately to improve clarity for the reader. Finally, we sum up the thesis with the main conclusions in ch. 8 together with an outline of possible future work.

1.4 Software

All computational analyses are implemented in `Python 3.9.7` with the following list of third-party packages: `black, matplotlib, numba, numpy pip, pyinstrument, colorama, scipy, seaborn, pyperclip`. Crucial for the optimization of the numerical integration is `numba`, which compiles python code to optimized machine code at runtime. This can give computation times that approach C or FORTRAN, and in the case of especially the Sabra shell model the integration time was lowered by a factor 40.

The main code is available in the GitHub repository `PertMethInEnsWeath Forecasts_src` at `https://github.com/mlf93-physics/PertMethInEns WeathForecasts_src`. The repository is structured into three main folders `lorentz63_experiments`, `shell_model_experiments` and `general`. While the first two folders contain the model specific algorithms, the latter contains all parts common to the models. E.g. the `general/runners` folder contains the scripts that run the perturbation methods individually, while the `comparison_runners` runs a collection of perturbation methods.

Another GitHub repository, `PertMethInEnsWeathForecasts_libutils`, is used for very general utility functions and can be accessed at `https: //github.com/mlf93-physics/PertMethInEnsWeathForecasts_libutils`

2 Chaotic Dynamics and Turbulence

To understand what Lorentz discovered in 1963 and the consequences for predicting the weather, we will in this chapter start out by reviewing fundamental concepts of chaos. To support this an outline of the linear stability analysis of a non-linear model is presented together with a thorough introduction to the Lyapunov vectors and exponents. The stability analysis also forms the basis for the normal mode (sec. 4.1) and singular vector method (sec. 4.2). The Lyapunov vectors are studied as a perturbation method on its own as well as a reference for comparing methods (sec. 4.6). In the last section of this chapter, the theory of turbulence is presented which forms the theoretical basis for introducing the Sabra shell model (sec. 3.2).

If we consider a dynamical system which solutions exist in an M-dimensional phase space, the system is said to be chaotic if the solutions depend *sensitively* on the initial conditions; that is, if infinitesimal perturbations lead to diverging solutions with time. This is shown schematically in fig. 2.1. As will be described in sec. 2.2, this is the case if the system has at least one positive *Lyapunov exponent*. In relation to fig. 2.1, the Lyapunov exponent is given by the average exponential growth rate of the distances between the reference and the perturbed trajectories.

Although embedded in an M-dimensional space, the solutions may be confined to only a subspace due to the attraction from a so-called *attractor*. An attractor is a globally bounded attracting set of points, which can consist of fixpoints, limit cycles and *strange attractors*. A strange attractor is a special kind of attractor, which is characterized by the sensitive dependence on the initial conditions of trajectories starting on the attractor (Kalnay, 2002; Strogatz, 2000).

To characterize the system at a given instant in time, we now study the linearization of a non-linear model.

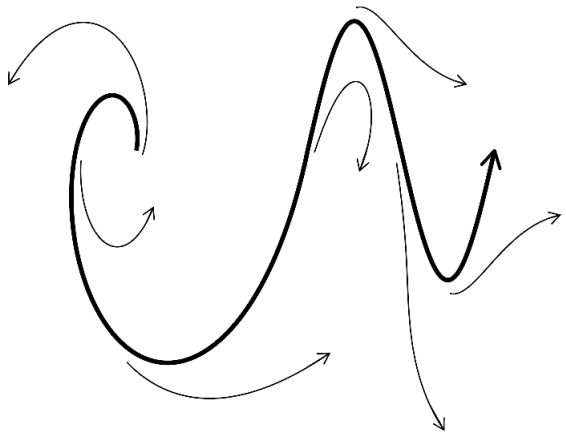


Figure 2.1: Conceptual visualization of a flow in phase space that sensitively depends on the initial conditions. A reference trajectory (thick curve) is perturbed at various positions and the perturbed trajectories (thin curves) diverge exponentially. From (Ditlevsen, 2011).

2.1 Linear Stability Analysis

A system of m first order non-linear differential equation can be written on the form

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) \tag{2.1}$$

from which a trajectory $\mathbf{x}(t)$ can be obtained given an initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$. If perturbing this initial condition, i.e. $\mathbf{x}(t_0) = \mathbf{x}_0 + \mathbf{x}'(t_0)$, a different trajectory $\mathbf{x}(t) + \mathbf{x}'(t)$ is obtained that satisfy

$$\dot{\mathbf{x}}(t) + \dot{\mathbf{x}}'(t) = f(\mathbf{x}(t) + \mathbf{x}'(t)) \tag{2.2}$$

Now we assume that $||\mathbf{x}'(t_0)|| \ll ||\mathbf{x}_0||$ and that t is close to t_0. This enables us to linearize the system around the unperturbed initial point:

$$\dot{x}_i(t) + \dot{x}'_i(t) \approx f_i(\mathbf{x}(t)) + \partial_j f_i\big|_{\mathbf{x}_0} x'_j(t), \tag{2.3}$$

Here we use the tensor notation $\partial_i = \partial/\partial x_i$, $\partial_{ij} = \partial^2/(\partial x_i \partial x_j)$ and the Einstein notation where repeated indices imply summation, i.e. $\partial_i u_i = \partial_1 u_1 + \partial_2 u_2 + \partial_3 u_3$ (Ditlevsen, 2011). We will use this notation in the rest of the thesis. Exceptions to this are marked with a $\star$ when needed.

By subtracting eq. (2.1) from eq. (2.3), we get

$$\dot{x}'_i(t) = \partial_j f_i\big|_{\mathbf{x}_0} x'_j(t) = \mathcal{J}_{ij} x'_j(t), \tag{2.4}$$

where

$$\mathcal{J}_{ij} \equiv \partial_j f_i\big|_{\mathbf{x}_0} \tag{2.5}$$

is the Jacobian of eq. (2.1) at $t = t_0$. The complete set of equations in eq. (2.4) is called the *tangent linear model* (TLM) on differential form (Kalnay, 2002). For later reference, we will also state the TLM on integral form:

$$\mathbf{x}'(t) = L(\mathbf{x}(t_0), \mathbf{x}(t))\mathbf{x}'(t_0), \tag{2.6}$$

where $L(\mathbf{x}(t_0), \mathbf{x}(t))$ is the *forward propagator* of the linearized model; it propagates the initial perturbation from time t_0 to time t with the Jacobian evaluated on the non-linear trajectory. For simplicity, we will from now on refer to it as $L(t_0, t)$, although the time dependence comes through $\mathbf{x}$.

The Jacobian is a square matrix and can be diagonalized to get the eigenvalues and -vectors:

$$\boldsymbol{\mathcal{J}} = \boldsymbol{\Xi}\mathbf{M}\boldsymbol{\Xi}^{-1}, \tag{2.7}$$

where $\mathbf{M}$ is a diagonal matrix with eigenvalues $\mu_1, \mu_2, ..., \mu_m$ on the diagonal sorted in descending order, and the columns of $\boldsymbol{\Xi}$ are the corresponding eigenvectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_m$.

Following (Ditlevsen, 2011) eq. (2.4) can be solved from time t_0 to t:

$$\mathbf{x}'(t) = e^{\boldsymbol{\mathcal{J}}t}\mathbf{x}'(t_0), \tag{2.8}$$

Determining the exponential operators in terms of its Taylor series and substituting $\boldsymbol{\mathcal{J}}$ with eq. (2.7) yields

$$\mathbf{x}'(t) = \sum_{n=0}^{\infty} \frac{(\boldsymbol{\mathcal{J}}t)^n}{n!}\mathbf{x}'(t_0) \tag{2.9a}$$

$$= \boldsymbol{\Xi}\sum_{n=0}^{\infty} \frac{[\mu_1^n, \mu_2^n, \ldots, \mu_m^n]t^n}{n!}\boldsymbol{\Xi}^{-1} \tag{2.9b}$$

$$= \boldsymbol{\Xi}[e^{\mu_1 t}, e^{\mu_2 t}, \ldots, e^{\mu_m t}]\boldsymbol{\Xi}^{-1}\mathbf{x}'(t_0), \tag{2.9c}$$

where $[,,,]$ denotes a diagonal matrix. This can be simplified by introducing the variable $\mathbf{z}(t) = \boldsymbol{\Xi}^{-1}\mathbf{x}'(t)$:

$$\mathbf{z}(t) = [e^{\mu_1 t}, e^{\mu_2 t}, \ldots, e^{\mu_m t}]\mathbf{z}(t_0) \tag{2.10}$$

This shows that the evolution of an initial perturbation is given by a set of eigenmodes $z_i(t) = e^{\mu_i t}z_i(t_0)$ $\star$, where the value of $\Re[\mu_i]$ determines whether

the initial perturbation grows (positive), decays (negative) or stays constant (zero) along the corresponding eigenvector. If the eigenvalue has an imaginary part, the perturbation will experience rotation in phase space in a plane spanned by the corresponding complex-conjugate pair of eigenvectors. This analysis can be used to evaluate the stability of *fixpoints* (points for which $\dot{\mathbf{x}} = 0$) as will be done with the Lorentz-63 model in sec. 3.1. In that case the sign of $\Re[\mu_i]$ indicates if a fixpoint is stable (positive), unstable (negative) or neutral (zero) in the direction of the corresponding eigenvector.

2.2 Lyapunov exponents and vectors

While the above analysis gives knowledge about the instantaneous stability of the flow around a point $\mathbf{x}(t_0)$, the Lyapunov exponents and vectors describe the stability of a larger portion of the flow. One distinguish between the global Lyapunov exponents (often simply referred to as *the Lyapunov exponents*), which are independent of time, and the local Lyapunov exponents (LLE; also known as *finite time Lyapunov exponents* (FTLE)), which are defined for a finite period of time. The Lyapunov vectors are more precisely referred to as local Lyapunov vectors (LLV) since they depend on an initial point, which will be clear in a moment (Kalnay, 2002; Legras et al., 1995).

The global Lyapunov exponents, $\{\lambda_i\}$, describe the average exponential growth or decay rate of the size of a perturbation independently of the position on the attractor. In that way, they are useful to characterize the overall stability of a flow. As described in ch. 1, though, the predictability of the atmosphere is not constant in time. Hence, it is relevant to work with local measures of error growth of perturbations, which the LLEs and LLVs are useful for.

The definition of the global Lyapunov exponents and the LLVs is founded in Oseledec's theorem from which the first two statements will be outlined following Legras et al., 1995. The theorem states that for almost any solution $\mathbf{x}(t)$ to eq. (2.1) and for almost any inner product on $\Re^m$

1. a finite exponent

$$\lambda(\mathbf{r}) = \lim_{t_1 \to \infty} \Big(\frac{1}{t_1 - t_0} \ln \Big[\frac{||L(t_0, t_1)\mathbf{r}||}{||\mathbf{r}||} \Big] \Big), \tag{2.11}$$

exists, for any vector $\mathbf{r}$ in $\Re^m$, and is independent of t_0. The exponent can take maximal m values $\lambda_1, \lambda_2, \ldots, \lambda_m$, which generally are different from each other.

2. the operator

$$S^{\infty}(t_0) = \lim_{t_1 \to \infty} [L(t_0, t_1)^* L(t_0, t_1)]^{\frac{1}{2(t_1 - t_0)}} \tag{2.12}$$

exists, where L^* is the adjoint of L (defined below). The direction of the eigenvectors of $S^{\infty}(t_0)$ depend on the initial point $\mathbf{x}(t_0)$, while the corresponding eigenvalues $e^{2\lambda_i}$ do not.

The adjoint (conjugate transpose) of the forward propagator (in short; *the adjoint*), L^*, is defined by

$$\langle \mathbf{x}; L\mathbf{y} \rangle = \langle L^*\mathbf{x}; \mathbf{y} \rangle. \tag{2.13}$$

From this we have that the global Lyapunov exponents are the $\{\lambda_i\}$ and the LLVs $\{\boldsymbol{\zeta}_i\}$ are the eigenvectors of $S^{\infty}(t_0)$. Put into words, the global Lyapunov exponents are given by the growth rate of any initial perturbation evolved through an infinitely long integration of the TLM. There exists as many global Lyapunov exponents as the dimension of the system, and if at least one λ_i is positive the system is chaotic (Kalnay, 2002).

The LLV $\boldsymbol{\zeta}_1$ is called the dominant or leading LLV and dictates the direction of maximum sustainable growth[1] of an initial perturbation. This holds in general no matter which initial perturbation method is used, given that $t_1 \to \infty$, and is visualized in fig. 2.2 (Kalnay, 2002).

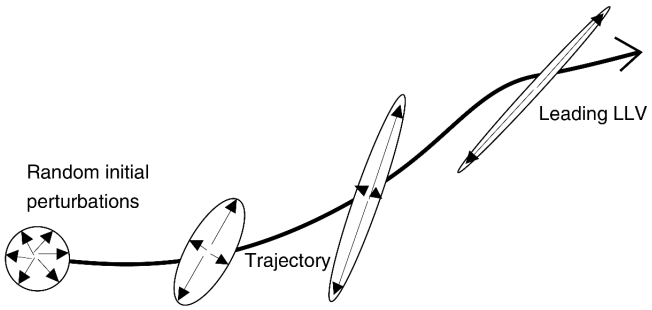


Figure 2.2: Visualization of a random initial perturbation evolving into the direction of the dominant LLV. The expansion rate of the sphere into an ellipsoid is given by the dominant Lyapunov exponent, λ_1. From (Kalnay, 2002).

Having obtained the LLVs, one can estimate the LLEs $\{\pi_i\}$ for some finite time period Δt by evolving the LLVs with the forward propagator into

[1] If $\lambda_1 < 0$ it dictates the direction of minimum sustainable decay.

$\boldsymbol{\zeta}_i(t_0 + \Delta t)$, i.e.

$$\pi_i \approx \frac{1}{\Delta t} \ln\Big[\frac{||\boldsymbol{\zeta}_i(t_0 + \Delta t)||}{||\boldsymbol{\zeta}_i(t_0)||}\Big], \tag{2.14}$$

which are indeed dependent on t_0 (Kalnay, 2002). The largest LLE, π_1, is called the *leading* local Lyapunov exponent, and determines the local degree of predictability through the time $T_{LLE} = 1/\pi_1$ after which predictability is lost (Ditlevsen, 2011).

To align the above definitions with Legras et al. and prepare for the comparison with singular vectors in sec. 4.6.2, a subdivision of the LLVs is needed. If the solution $\mathbf{x}(t)$ to eq. (2.1) lies on the attractor of the system, then it is invariant under the transformation $t \to -t$. In turn, the Oseledec theorem is also valid for such a transformation in which we let $t_0 \to -\infty$ and t_1 be fixed. The global Lyapunov exponents will remain unchanged, but in order to calculate the LLVs we need to find the eigenvectors of

$$S^{-\infty}(t_1) = \lim_{t_0 \to -\infty} [L(t_0, t_1) L(t_0, t_1)^*]^{\frac{1}{2(t_1 - t_0)}}. \tag{2.15}$$

Because we then have two different definitions of the LLVs, we will define $\{\boldsymbol{\zeta}_i^{\infty}\}$ as the LLVs that correspond to the eigenvectors of eq. (2.12) (referred to as *the LLVs*)[2] and $\{\boldsymbol{\zeta}_i^{-\infty}\}$ as the LLVs that correspond to the eigenvectors of eq. (2.15) (referred to as *the adjoint LLVs*)[3].

In practice eq. (2.11) works well to get the dominant Lyapunov exponent, but fails if more than one is needed. Furthermore, it is not possible to calculate and diagonalize the operators in eqs. (2.12) and (2.15) due to numerical under- and overflow. According to Legras et al., 1995, one can overcome those problems by the following: it can be shown that it suffices to integrate an initial random perturbation for a time $t_1 - t_0 \to \infty$ with the $L(t_0, t_1)$, $L(t_0, t_1)^*$ operator to get the dominant LLV and the adjoint LLV, respectively. To get multiple orthogonal LLVs and adjoint LLVs an iterative Gram-Schmidt orthonormalization of the evolved perturbations is made (Legras et al., 1995). The procedure for calculating multiple LLVs $\{\boldsymbol{\zeta}_i^{\infty}\}$ can be outlined as follows.

To get k orthonormal LLVs

1. define an $(m \times k)$ matrix $\mathbf{E}(t_0)$, where the column vectors, $\{\mathbf{e}_i\}$, are random perturbation vectors.

[2]Referred to as the *backward* Lyapunov vectors by Legras et al., 1995

[3]The term adjoint is used, since, as we will see shortly, they are in practice calculated by integrating L^*. Referred to as the *forward* Lyapunov vectors by Legras et al., 1995

2. Evolve this matrix column by column with the forward propagator:

$$\mathbf{e}_i(t_1) = L(t_0, t_1)\mathbf{e}_i(t_0) \tag{2.16}$$

3. Perform a Gram-Schmidt orthonormalization of $\mathbf{E}(t_1)$, which decomposes the matrix into two matrices $\mathbf{Q}$ and $\mathbf{T}$:

$$\mathbf{E}(t_1) = \mathbf{Q}(t_1)\mathbf{T}(t_1). \tag{2.17}$$

   The matrix $\mathbf{Q}$ has size $(m \times k)$, and its columns are orthonormal vectors called *Schmidt vectors*. The matrix $\mathbf{T}$ has size $(k \times k)$ and is upper triangular with positive diagonal entries.

For $t_1 - t_0 \to \infty$ the Schmidt vectors will converge to the LLVs $\{\boldsymbol{\zeta}_i^{\infty}\}$, and from the diagonal entries $T_{i,i}$ one can calculate the global Lyapunov exponents (λ_i). This is done by substituting $T_{i,i}$ for $||L(t_0, t_1)\mathbf{r}||$ in eq. (2.11).

To avoid numerical overflow problems, the orthonormalization is in practice performed at regular intervals, i.e. the items 1 to 3 are repeated. After each iteration, the random perturbations $\{\mathbf{e}_i\}$ in eq. (2.16) are replaced by the Schmidt vectors from the last iteration (Legras et al., 1995). The calculated Lyapunov exponents for each iteration are averaged to get a final estimate of the exponents.

Having introduced the Lyapunov exponents and vectors, we will now continue with the theory of *Turbulence*.

2.3 Introduction to Turbulence

Fully developed turbulence is a phenomenon that is observed in as diverse physical situations and spatial scales as to pour out a glass of wine, atmospheric cyclones or the breaking of a wave on a coast. In that way it is a phenomenon we can observe directly in our daily life, and have an intuitive understanding of, despite its complex nature and unsolved mysteries. More specifically fully developed turbulence can be thought of as the behaviour of any fluid brought into vigorous motion (Ditlevsen, 2011).

It is believed that this motion, no matter what fluid, can be described by the Navier-Stokes equation (NSE) provided a set of initial and/or boundary conditions. Following this path, however, involves heavy numerical simulations since no general analytical solution exists. With considerable advantages, one can instead take another more phenomenological path to establish an understanding of turbulence. This path follows in the footsteps of L. F. Richardson and A. Kolmogorov.

In the proceeding subsections those two paths will be presented starting out with the Navier-Stokes equation.

2.3.1 The Navier-Stokes equation

We will work with a fluid under the continuous approximation (Lautrup, 2011). This approximation holds when the smallest spatial scale

- is large compared to the mean free path of the molecules that constitute the fluid. This ensures that the granularity of the fluid on the molecular level can be ignored.
- is small enough such that variation of a quantity describing the fluid on this scale is small relative to a desired precision.

In that way, the state of the fluid can be completely described by continuous fields of the velocity $u_i(\mathbf{x}, t)$, the temperature $T(\mathbf{x}, t)$, the pressure $p(\mathbf{x}, t)$ and the density $\rho(\mathbf{x}, t)$, where i stands for the ith spatial direction. The evolution of those fields are governed by equations derived from momentum, mass, and energy conservation together with the equation of state. We will assume that the fluid is incompressible and that the buoyancy force can be neglected. The first assumption means that the density of the fluid is constant, while the second assumption makes the evolution of the temperature field independent of the velocity and the pressure fields. We can then focus on the evolution of the velocity and the pressure fields, which is described by the NSE

$$\partial_t u_i + u_j \partial_j u_i = -\partial_i p + \nu \partial_{jj} u_i + f_i, \tag{2.18}$$

where all terms are per unit mass, and the continuity equation

$$\partial_i u_i = 0 \tag{2.19}$$

The NSE describes how the velocity of a fluid parcel changes due to the pressure gradient force, $-\partial_i p$ [4], the viscous force, $\nu \partial_{jj} u_i$, where ν is the kinematic viscosity, and all other forces f_i. This can be viewed from a fixed reference frame, where the second term on the left-hand side (LHS) is then the advection of the velocity field (Eulerian view), or from a reference frame moving with the flow, where the complete LHS is then the material derivative (Lagrangian view). The continuity equation states that there can be no net divergence/convergence in the fluid; in other words the mass of the fluid is conserved.

We can eliminate pressure in eq. (2.18) to get a set of equations in u_i only by assuming that the force term f_i is rotational, $\partial_i f_i = 0$, i.e. with no divergence/convergence. By applying ∂_i on both sides and using eq. (2.19) we get (Ditlevsen, 2011)

[4] The density is normally written in the denominator of the pressure gradient force term, but it is here absorbed into the unit of pressure since the density is assumed constant.

$$\partial_t \partial_i u_i + \partial_i u_j \partial_j u_i + u_j \partial_j \partial_i u_i = -\partial_{ii} p + \nu \partial_{jj} \partial_i u_i \tag{2.20a}$$

$$\partial_{ii} p = -\partial_i u_j \partial_j u_i \tag{2.20b}$$

$$p = -\partial_{kk}^{-1} (\partial_i u_j \partial_j u_i). \tag{2.20c}$$

We will use this equation to derive the spectral form of the NSE in appendix A.1.

The influence of the viscosity on the flow can be seen as to smooth out variations, since it depends on the second derivative of the velocity. On the contrary the non-linear terms (the advection and the pressure gradient force[5]) will make the flow irregular and asymmetric if dominating. What terms dominate in the NSE therefore depicts how regular/irregular the flow will be and is characterized by the Reynolds number

$$\mathsf{R}e \equiv \frac{UL}{\nu} \sim \Big(\frac{\text{advection term}}{\text{viscosity term}}\Big), \tag{2.21}$$

where U is the characteristic velocity at scale L, and L is the length scale of the largest variations of the flow (Ditlevsen, 2011)[6]. In the following we will see visually what happens with the flow of a fluid when increasing the Reynolds number (see fig. 2.3 and refer to (Frisch, 1995; Van Dyke, 1982))

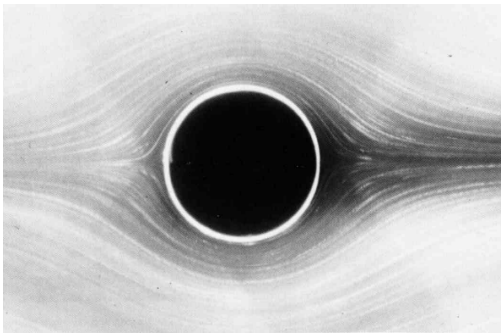
For a small Reynolds number the flow is dominated by the viscous force and appears smooth as seen in fig. 2.3a where the flow of a fluid around a cylindrical obstacle is shown. The flow is symmetric across a horizontal axis and almost symmetric across a vertical axis. As the Reynolds number is increased the latter symmetry breaks, and two eddies start to develop on the right side of the obstacle (fig. 2.3b). A further increase in the Reynolds number yields the Kármán vortex street, where also the symmetry across the horizontal axis is partly broken[7] (fig. 2.3c). The flow shows a periodic pattern in time and can still be said to possess a temporal symmetry. For an even larger Reynolds number the flow is completely dominated by the non-linear forces and appears irregular, asymmetric and chaotic (fig. 2.3d). This behaviour is denoted *fully developed turbulence*, and it generally occurs in the limit of very high Reynolds numbers (Ditlevsen, 2011). The fluid no longer shows temporal and spatial symmetries in their traditional sense, but if one looks at the fluid in a statistical sense symmetries in e.g. time-averaged

[5]The pressure gradient force is non-linear since from eq. (2.20c) it can be seen that it depends on the velocity squared

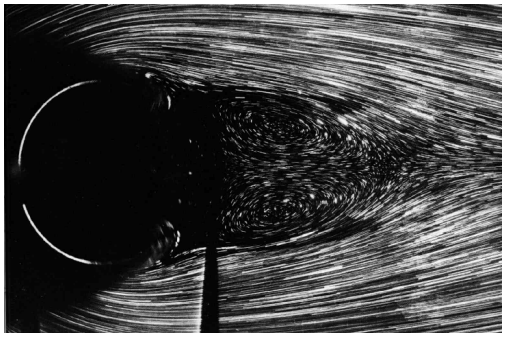
[6]One can derive this quantity by putting the NSE on non-dimensional form as done in appendix A.2

[7]As noted by Frisch, 1995 the symmetry is preserved but offset in time by half a period.

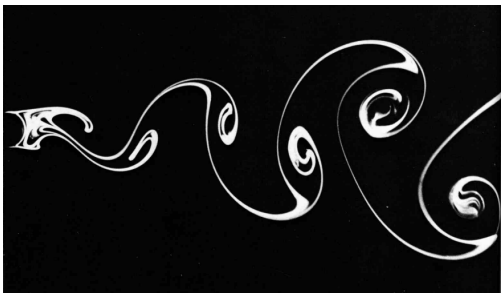
(a) $\mathsf{R}e = 0.16$



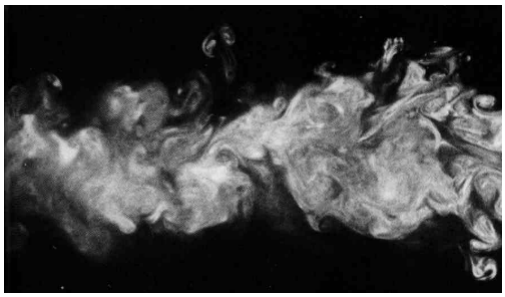
(b) $\mathsf{R}e = 24.3$



(c) $\mathsf{R}e = 140$



(d) $\mathsf{R}e = 1770$



(e) $\mathsf{R}e = 1500$

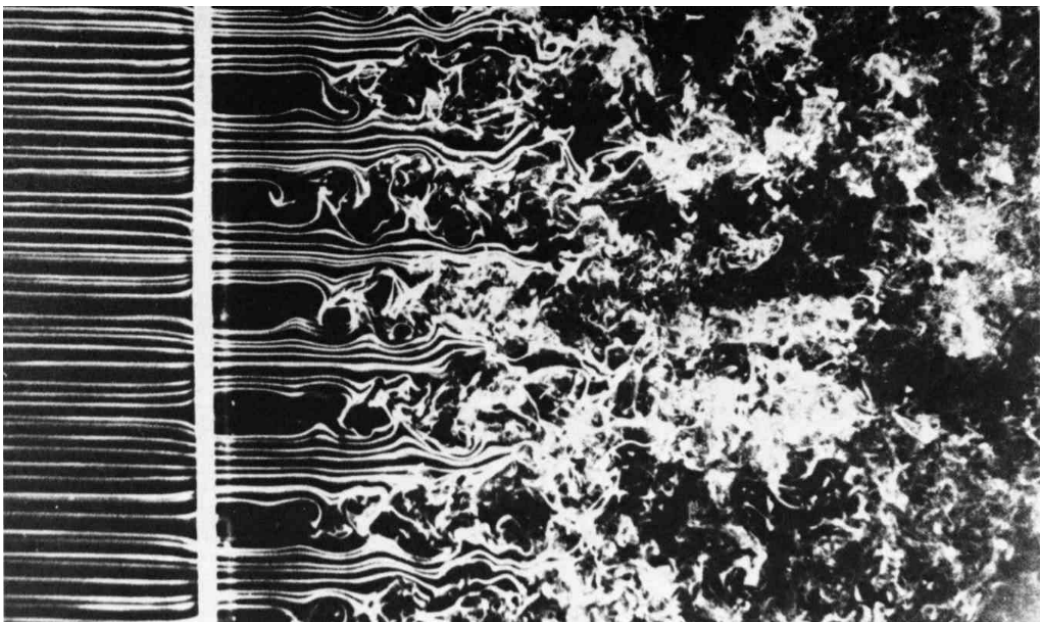


Figure 2.3: The flow of a fluid around a cylindrical obstacle (a-d) and through a grid (e) for an increasing Reynolds number. The Reynolds numbers are estimated from the diameter of the cylindrical obstacle and the distance between grid lines, respectively. Spatial and temporal symmetries are successively broken, but in (e) (in the rightmost part) they are restored in a statistical sense, e.g. averaged quantities appear to be invariant under translation and rotation. The state of the flow in this limit of a very high Reynolds number is called fully developed turbulence. Reference for images: (Van Dyke, 1982).

quantities reappear[8]. This is very clear in the right most part of fig. 2.3e, where the obstacle is replaced by a grid, if imagining taking a time-average of e.g. the velocity of the fluid in a specific point (Frisch, 1995).

To describe fully developed turbulence more quantitatively, one should in principle start from the NSE. However, no such theory exists (yet), and even numerical experiments can be cumbersome and computationally expensive as we will see in sec. 3.2 (Ditlevsen, 2011; Frisch, 1995). Instead, one can follow a statistical and phenomenological path, which is exactly what Kolmogorov did in his theory from 1941. The theory builds upon Richardson's phenomenological picture of turbulence, which we will start out revisiting (Richardson, 1922).

2.3.2 Richardson's picture and Kolmogorov's 1941 theory (K41)

Consider a flow that is stirred at large spatial scales for some time and then let alone. Richardson imagined this flow to initially consist of large eddies that gradually break up into smaller and smaller eddies until dissipated by viscosity as visualized in fig. 2.4. In terms of energy the energy added to the flow at large scales cascades to smaller and smaller scales until it turns into heat by viscosity.

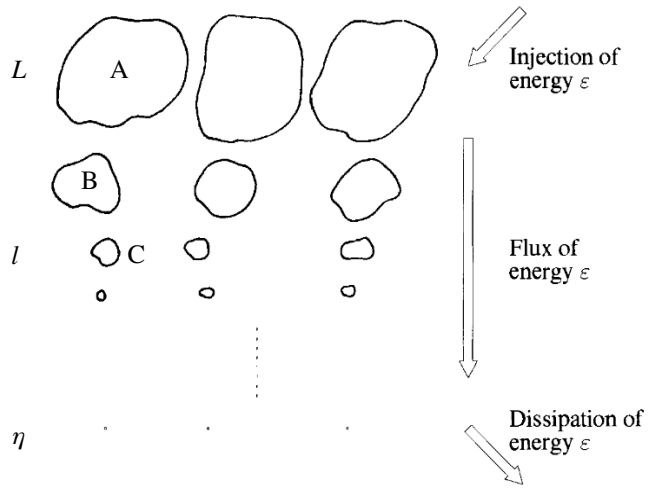


Figure 2.4: Richardson's picture of the energy cascade from large to small scales; from large eddies to small eddies. Three eddies to the left are labelled for later reference. (Frisch, 1995).

[8]If one can assume that the flow is homogeneous and isotropic, which will be explained in sec. 2.3.2, the time-average can equally well be a spatial average or an ensemble average

Kolmogorov developed a scaling theory (the K41 theory) based on this picture ((Kolmogorov, 1941) as cited in (Ditlevsen, 2011)), which has been verified by experiments and observations and still holds today to a large extent (Ditlevsen, 2011). We will derive the essential parts of this theory from a physical perspective by imagining the flow as in Richardson's picture. The forcing is set to continuously act on large scales such that the flow is in statistical equilibrium; i.e. the mean energy dissipation due to viscosity, $\bar{\epsilon}$, is balanced by this forcing. We look at some length scale l for which $\eta \ll l \ll L$, where η is the scale at which viscosity starts to dominate also known as the *Kolmogorov scale*. L is also known as the *integral scale*, while l belongs to the *inertial range*, i.e. the range between L and η where inertia dominates. We assume that the flow is homogeneous and isotropic. To be homogeneous means that the average of some quantity, e.g. the velocity, is invariant under a spatial translation, i.e. there are no gradients of the mean velocity in the flow. Isotropic on the other hand means that there is no preferred direction for averaged quantities, i.e. they are invariant under a rotation (Ditlevsen, 2011).

Now we look at the typical velocity difference between two points separated a distance l:

$$\delta u(l) \equiv |u(r+l) - u(r)|, \tag{2.22}$$

where r is the position of one of the points. We have omitted the vector indices, since it holds for any index and makes the notation simpler. So we can now ask if there is anything special to this scale l? The eddy that corresponds to this scale (eddy B in fig. 2.4) is embedded within a much larger eddy (eddy A), which more or less just moves eddy B around as a rigid body. At the same time an even smaller eddy (eddy C) is embedded within eddy B, which similarly moves eddy C around. This indicates that the answer to the question is *no* and that the flow can be considered self-similar (Ditlevsen, 2011).

With self-similar we mean that for two scales l_1, l_2 for which $\eta \ll l_1 < l_2 \ll L$ the velocity difference at the two scales are related through some function f by

$$\delta u(l_2) = f\Big(\frac{l_1}{l_2}\Big)\delta u(l_1). \tag{2.23}$$

We can determine f from dimensional counting and by the assumption that $\delta u(l)$ only depends on l and $\bar{\epsilon}$, which is also known as Kolmogorov's second universality assumption (as stated in (Frisch, 1995)). The validity of this assumption follows from the above described effect of the larger eddies on

the smaller, which is sometimes called *the sweeping argument*. Due to the large difference in scale between the eddies, no deformation of the smaller eddy by the larger eddy occurs, and they can be considered independent. Furthermore, $\bar{\epsilon}$ characterizes the state of the fluid, since it is a measure of the energy input/output.

The units of l and $\bar{\epsilon}$ are $[l] = \mathrm{m}$ and $[\bar{\epsilon}] = \mathrm{m}^2/\mathrm{s}^3$. For the velocity difference to have correct units ($[\delta u] = \mathrm{m/s}$) they can only be combined in one way[9] which is

$$\delta u(l) \sim (\bar{\epsilon} l)^{1/3}. \tag{2.24}$$

This is the essential relation in the K41 theory. It can be used to derive other important relations e.g. and expression for η:

We look at the scale $l = \eta$ and the point where the non-linear terms and the viscous term in eq. (2.18) will balance[10]:

$$u_j \partial_j u_i \sim \nu \partial_{jj} u_i \tag{2.25a}$$

$$\frac{\delta u(\eta)^2}{\eta} \sim \nu \frac{\delta u(\eta)}{\eta^2} \tag{2.25b}$$

$$\frac{(\bar{\epsilon}\eta)^{2/3}}{\eta} \sim \nu \frac{(\bar{\epsilon}\eta)^{1/3}}{\eta^2} \tag{2.25c}$$

$$(\bar{\epsilon}\eta)^{1/3}\eta \sim \nu \tag{2.25d}$$

$$\eta \sim \left(\frac{\nu^3}{\bar{\epsilon}}\right)^{1/4}, \tag{2.25e}$$

where we have used eq. (2.24) in eq. (2.25c). η can be related to the $\mathsf{R}e$ by keeping the integral quantities U, L and the mean energy dissipation $\bar{\epsilon}$ constant. From eqs. (2.21) and (2.25e) we have

$$\nu \sim 1/\mathsf{R}e \tag{2.26a}$$

$$\eta \sim \nu^{3/4} \sim \mathsf{R}e^{-3/4}. \tag{2.26b}$$

Thus, the higher the Reynolds number, the smaller the Kolmogorov scale and consequently the broader an inertial range. Fully developed turbulence have for that reason a very broad inertial range, in which the effect of dissipation is negligible.

[9] From $\frac{m}{s} = [\delta u(l)] = [\bar{\epsilon}]^\alpha [l]^\beta = \left(\frac{\mathrm{m}^2}{\mathrm{s}^3}\right)^\alpha \mathrm{m}^\beta$ then $\alpha = \beta = 1/3$.

[10] As noted earlier, the pressure gradient force is non-linear as can be seen from eq. (2.20c). From this equation it is also seen that it is of the same order of magnitude as the advection term in eq. (2.18). For that reason we simply compare the advection term with the viscous term.

Another famous relation that can be derived using eq. (2.24) is the scaling relation of the spectral energy density ($E(k)$, where $k \sim 1/l$ is the wave vector at scale l) of the inertial range of a turbulent flow. Here, we derive the relation from dimensional counting. From the Kolmogorov's second universality assumption we have that $E(k)$ can only depend on k and $\bar{\epsilon}$. The relevant units are $[k] = \mathrm{m}^{-1}$ and $[E(k)] = \mathrm{m}^3/\mathrm{s}^2$. We then have

$$\frac{\mathrm{m}^3}{\mathrm{s}^2} = [E(k)] = [\bar{\epsilon}]^\alpha [k]^\beta = \left(\frac{\mathrm{m}^2}{\mathrm{s}^3}\right)^\alpha \left(\frac{1}{\mathrm{m}}\right)^\beta \Rightarrow \tag{2.27}$$

$$\alpha = 2/3, \;\; \beta = -\frac{5}{3} \tag{2.28}$$

which gives

$$E(k) \sim \bar{\epsilon}^{2/3} k^{-5/3}. \tag{2.29}$$

This spectrum has been confirmed both by experiments and observations of the atmospheric boundary layer (Ditlevsen, 2011).

3 Models

The methods for perturbing the initial conditions of an ensemble forecast is studied with two low-dimensional models; the well known Lorentz-63 model introduced by Lorenz, 1963 and the Sabra shell model of turbulence (Ditlevsen, 2000; L'vov et al., 1998). In this chapter we present the models and their main characteristics. The structure of the dynamics in phase space is studied through a linear stability analysis as described in sec. 2.1, which serves as a reference for a later comparison of the different perturbation methods (ch. 6). Before moving on, we will justify the choice of the models.

The Lorentz-63 model is chosen for two reasons: 1) it is a well known and thoroughly studied model, which enables comparison with other studies; 2) it shows qualitative similarities to the large-scale atmospheric flow. Those similarities count e.g. the existence of regimes, multiple timescales and varying local predictability (Palmer, 1993).

In contrast to the Lorentz-63 model, the theory of turbulence deals with all relevant scales of e.g. the atmospheric flow as described in the preceding chapter. Although theoretically interesting, it is computational difficult to handle without applying truncations of some kind. The shell models do exactly that, such that all relevant scales can be contained within at most 30 variables. This makes it possible to study turbulence in a computationally efficient way.

This is the first reason why a shell model is chosen to complement the analyses made with the Lorentz-63 model; to see the difference to the case with all relevant atmospheric scales. The second reason is that to the author's knowledge perturbation methods have never been studied in the context of a shell model and specifically the Sabra shell model.

3.1 Lorentz-63 model

The Lorentz-63 model is a simplified atmospheric model based on convective dynamics (Lorenz, 1963; L. Magnusson et al., 2008). It has been studied extensively in relation to weather prediction (Palmer, 1993; Smith et al., 1999), due to its simplicity and yet non-trivial dynamics.

The model constitutes a three-dimensional system of non-linear first order differential equations given by

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz, \end{aligned} \tag{3.1}$$

where σ, r, b are time-independent parameters of the system (Lorenz, 1963). Written in matrix form the model reads

$$\dot{\mathbf{x}} = \mathbf{M}\mathbf{x}, \tag{3.2}$$

where

$$\mathbf{M} = \begin{bmatrix} -\sigma & \sigma & 0 \\ r & -1 & -x \\ y & 0 & -b \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3.3}$$

The model is dissipative since the divergence of eq. (3.1) is negative if $\sigma + b > -1$ (Palmer, 1993):

$$\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = -(\sigma + b + 1) \tag{3.4}$$

This means in turn that the model has an attractor (Kalnay, 2002). For certain parameter values, the attractor is a strange attractor that takes the form of the famous "butterfly-wings" shape (see fig. 3.1) and has a fractal dimension of ~ 2.05. In this case the model shows chaotic behaviour as evident from the definition of a strange attractor (ch. 2) (L. Magnusson et al., 2008; Palmer, 1993). The most commonly used values that give rise to this attractor are

$$\sigma = 10; \;\; r = 28; \;\; b = 8/3. \tag{3.5}$$

In this thesis, we will use exactly those parameter values because we are interested in studying the predictability in a chaotic system.

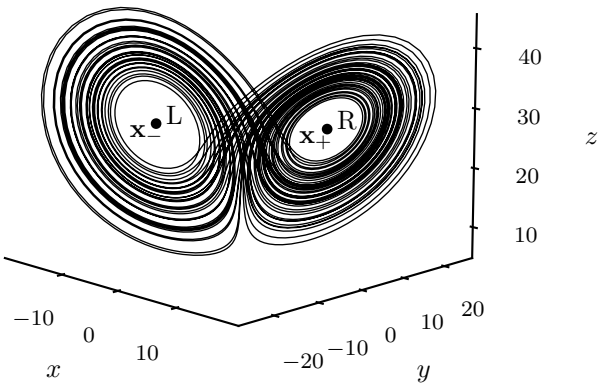


Figure 3.1: The well known attractor of the Lorentz-63 model for $\sigma = 10$, $r = 28$ and $b = 8/3$. The wings are denoted as left (L; around $\mathbf{x}_-$) and right (R; around $\mathbf{x}_+$) wing.

The model has three fixpoints for $r > 1$, which can be located at $\mathbf{x}_\pm = [\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1]$ and $\mathbf{x}_0 = [0,0,0]$ by solving $\dot{\mathbf{x}} = 0$. For $r < 1$ only the last fixpoint is present. This indicates that for $r = 1$ the qualitative structure of the solution of the model changes, which is denoted a *bifurcation* with $r = 1$ being the *bifurcation point* (Strogatz, 2000).

The stability of the fixpoints is crucially determined by the value of the parameters. To analyse the stability lets look at the Jacobian of the system which is derived from eq. (2.4) and eq. (3.2):

$$\boldsymbol{\mathcal{J}} = \begin{bmatrix} -\sigma & \sigma & 0 \\ r-z & -1 & -x \\ y & x & -b \end{bmatrix} \tag{3.6}$$

To evaluate the stability of the point $\mathbf{x}_0$ one can compute the eigenvalues of the Jacobian at that point. By solving $\det(\boldsymbol{\mathcal{J}} - \lambda\mathbf{I})\Big|_{\mathbf{x}=\mathbf{x}_0} = 0$ we get

$$\lambda = \left\{ -b, \frac{-(\sigma+1) \pm \sqrt{(\sigma+1)^2 + 4\sigma(r-1)}}{2} \right\} \tag{3.7}$$

For $r < 1$ the first term in the nominator of the $\pm$ eigenvalue pair is always greater than the last term in absolute sense, meaning that all three eigenvalues are negative and the point $\mathbf{x}_0$ is a stable fixpoint (refer to sec. 2.1). For $r > 1$ one eigenvalue becomes positive making the fixpoint unstable in the direction of the corresponding eigenvector. The same analysis can be made for the other two fixpoints for which the reader is referred to (Kalnay, 2002). For the parameter values that are used in this thesis (eq. (3.5)), all three

fixpoints are unstable. The $\mathbf{x}_\pm$ fixpoints are positioned in the centre of the left/right wing as shown with the L/R marks in fig. 3.1. More precisely the $\mathbf{x}_0$ fixpoint is a saddle point with an unstable direction parallel to the attractor, while the $\mathbf{x}_\pm$ fixpoints are weakly unstable spirals in the plane of the attractor and with a stable direction orthogonal to it (L. Magnusson et al., 2008).

To get an intuition on the flow around those fixpoints, the eigenvectors that correspond to the largest (real part) eigenvalues are plotted on top of the attractor for 5000 randomly chosen locations on the attractor (fig. 3.2). The colouring in (a) is determined by the real part and in (b) by the imaginary part of the eigenvalues. From this we can see that the flow is very divergent around $\mathbf{x}_0$ with no rotation, while in the top right (left) part of the left (right) wing the flow is contracting. We also see that the rotation around the $\mathbf{x}_\pm$ fixpoints is strongest in the top part of the wings.

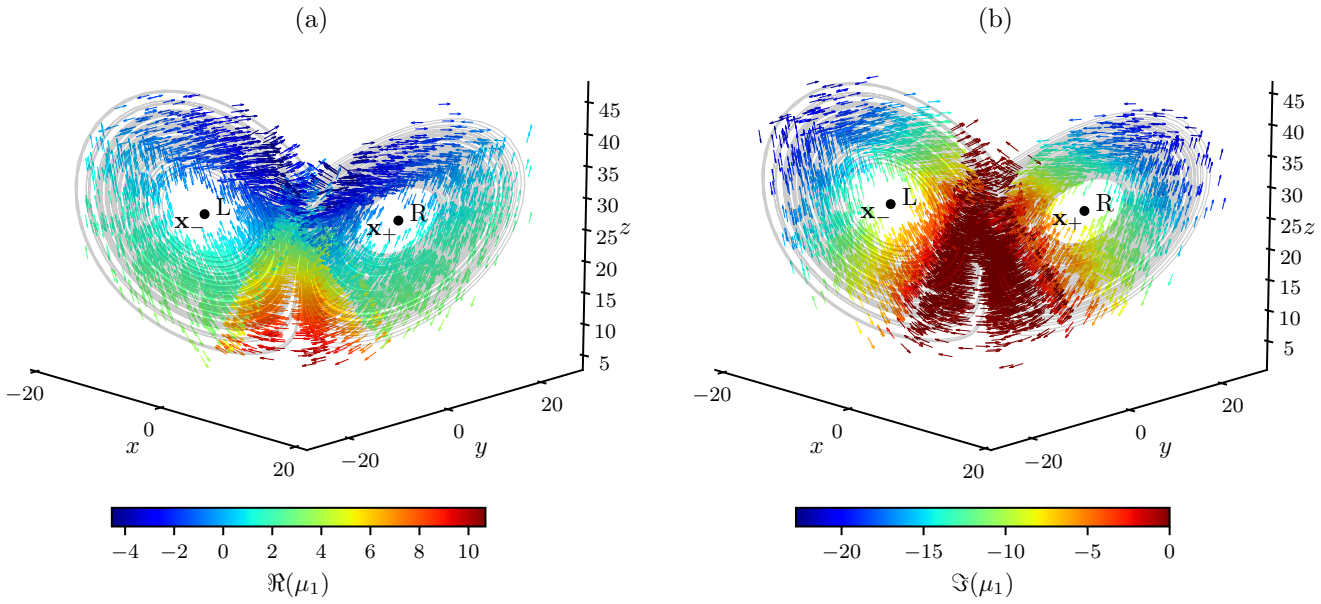


Figure 3.2: The distribution of eigenvectors (arrows) corresponding to the largest (real part) eigenvalue (plotted as the colour of the arrows) of the Lorentz-63 model plotted on top of the attractor. (a) real part of the eigenvalues; the largest real eigenvalues are located nearby the $\mathbf{x}_0$ fixpoint. (b) imaginary part of the eigenvalues; the imaginary part of the eigenvalues is zero at $\mathbf{x}_0$ and otherwise negative.

Despite that the attractor at first glance looks very regular, a trajectory of the system is very irregular and aperiodic (Strogatz, 2000). This is clearly seen when looking at the x component on its own (fig. 3.3) by the irregular

intervals at which the trajectory shifts between the left ($x < 0$) and right ($x > 0$) wing also denoted regime shifts. From this figure, it is also apparent that the system has two different timescales; a short timescale that corresponds to the rotation around the $\mathbf{x}_{\pm}$ fixpoints, and a long timescale that corresponds to the residence time in one wing. The short timescale can be estimated from a power spectrum analysis of the z component to be 0.76tu (see appendix B.1), where tu is an arbitrary time unit. The long timescale can be estimated by measuring the residence time and plotting the distribution (appendix B.1). As seen in fig. B.2, the distribution follows an exponential distribution. A characteristic timescale of such a distribution is the e-folding time, which for the residence time distribution is estimated to 1.35tu. The estimate of the short timescale agree well with (L. Magnusson et al., 2008), while the long timescale do not (L. Magnusson et al. finds 1.8tu). But since the method used by L. Magnusson et al. to estimate the latter is unknown to the author, it is difficult to compare and explain the discrepancy.

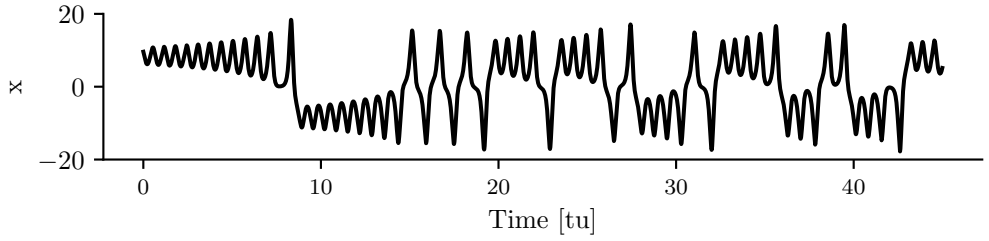


Figure 3.3: The x component of a solution to the Lorentz-63 model. The irregularity of the solution is apparent from aperiodic shifts between left ($x < 0$) and right ($x > 0$) wing.

3.2 Sabra shell model

As described in the introduction to this chapter, it is necessary to truncate the number of wave numbers included in a simulation of (high Reynolds number) turbulence in order to make it computational possible (Ditlevsen, 2011). This can be seen by considering the number of waves or degrees of freedom N_{DOF} needed to resolve the scales larger than the Kolmogorov scale η (eq. (2.25e)). With three spatial dimensions $N_{DOF} \sim \eta^{-3} \sim \mathsf{R}e^{9/4}$ which for an atmospheric Reynolds number of $\sim 10^5$ gives $N_{DOF} \sim 10^{11}$, i.e. way beyond the number of degrees of freedom of modern days NWP models

($\sim 10^7$) [1].

Many methods exist to reduce the number of degrees of freedom. For the case of shell models, the spectral space is divided into N concentric spheres with the radii or equivalently the magnitude of the wave vectors distributed like

$$k_n = k_0 \lambda^n \tag{3.8}$$

for $\lambda > 1$ and n being the shell number. In that way the shell models only retain wave vectors of specific sizes. For $n \in [1; 30]$, and by considering the circumference of the earth as the upper limit, this resolves spatial scales all the way down to 7cm [2], i.e. almost the entire range from small scale eddies produced by the surface drag on the atmosphere up to the largest planetary waves.

The shell models are not derived as approximations to the NSE, but are clever postulates that try to mimic its characteristics by having a similar structure. This holds for the Sabra model in particular. The model is given by the following set of coupled non-linear first order differential equations

$$\dot{u}_n = \iota k_n \big(u^*_{n+1} u_{n+2} - \frac{\epsilon}{\lambda} u^*_{n-1} u_{n+1} - \frac{\epsilon - 1}{\lambda^2} u_{n-2} u_{n-1} \big) - \nu k_n^2 u_n + f_n, \ \star \tag{3.9}$$

where u_n is the complex shell velocity of shell n, ϵ is a constant, ν the kinematic viscosity, and f_n a forcing applied at shell n. More specifically the Sabra shell model mimics the spectral NSE given in appendix A.1, which is clearly seen by comparing the equations term by term (see eq. (A.9b)). The first term on the right-hand side (RHS) of eq. (3.9) approximates the integral in the spectral NSE, since it contains interactions between shells, but only those between nearest and next nearest neighbouring shells. This is the non-linear part of the shell model and in relation to the original NSE (eq. (2.18)) it models the advection and the pressure gradient terms. The last two terms in eq. (3.9) directly correspond to the ones in the spectral NSE.

The Sabra shell model is constructed to have similar inviscid invariants as the NSE by the choice of coefficients in front of the three shell interaction terms $\left\{1, -\frac{\epsilon}{\lambda}, -\frac{\epsilon - 1}{\lambda^2}\right\}$ and by the position of the complex conjugations of the shell velocities (Ditlevsen, 2011).

[1] For ECMWFs HRES model with a horizontal resolution of 0.1 deg and 137 vertical layers $N_{DOF} \sim 6.3 \cdot 10^7$ for each variable (*About our forecasts* 2022).

[2] The spatial scales are given by the wavelengths $l = 2\pi/k_n$. After having normalized to the largest wavelength, $l_{norm} = l/\max(l)$, the wavelengths relative to the circumference of the earth, d_{earth}, is given by $l_{earth} = l_{norm} d_{earth}$

One such inviscid invariant is the total energy

$$\dot{E} = \frac{d}{dt}\sum_{n=1}^{N}\frac{1}{2}u_n u_n^* \tag{3.10a}$$

$$= \frac{1}{2}\sum_{n=1}^{N}(\dot{u}_n u_n^* + u_n \dot{u_n^*}) \tag{3.10b}$$

$$\vdots$$

$$= \frac{1}{2}\sum_{n'=n}^{N}\iota(k_{n'} - \frac{\epsilon}{\lambda}k_{n'+1} + \frac{\epsilon-1}{\lambda^2}k_{n'+2})u_{n'}^* u_{n'+1}^* u_{n'+2} \tag{3.10c}$$

$$= \frac{1}{2}\sum_{n'=n}^{N}\iota k_{n'}(1 - \frac{\epsilon}{\lambda}\lambda + \frac{\epsilon-1}{\lambda^2}\lambda^2)u_{n'}^* u_{n'+1}^* u_{n'+2} = 0, \tag{3.10d}$$

where we used eq. (3.8) in the second to last equal sign. The full derivation is given in appendix C. This result, however, not only shows that the Sabra shell model conserves total energy. It also shows that the local energy in each triad of shells (k_n, k_{n+1}, k_{n+2}) is conserved, since the terms in the parenthesis in eq. (3.10c) cancel. This is a unique property of the Sabra shell model compared to other shell models, e.g. the GOY shell model, and reflects the detailed balance of triads of wave vectors $(\mathbf{k}, \mathbf{k}', \mathbf{k}'')$ in the NSE, where $\mathbf{k} + \mathbf{k}' + \mathbf{k}'' = 0$ (Ditlevsen, 2011).

In fig. 3.4 a solution to the Sabra shell model is shown in three different ways, from which some main characteristics of a turbulent flow can be observed. The detailed model configuration is stated in sec. 5.4.2. First, the average spectral energy density shown in fig. 3.4a follows the Kolmogorov spectrum in eq. (2.29) in the inertial region (approx. shells 3-18). We see the Kolmogorov scale around shell 19 by the drop in energy density at that shell number and above. In fig. 3.4b we see the total energy vs time, which shows energy build-up and dissipation at irregular intervals. This is a very characteristic property of turbulent fluids and is also known as *intermittency* (Ditlevsen, 2011). In fig. 3.4c the build-up and dissipation is seen across the shells in the form of a Hovmöller diagram[3] of the shell energy anomaly. We can see how the energy builds up at large scales (small shell numbers) over long periods of time, and how it gradually is transferred to neighbouring shells. The larger the shell number the shorter the eddy turnover time (i.e. the characteristic timescale of shell n, see appendix B.2), and the transfer

[3]A type of diagram that is suited for showing the temporal evolution of some variable. Invented by Ernest Hovmöller (Liberto, n.d.)

accelerates. The energy ends up being dissipated by the shells around the Kolmogorov scale.

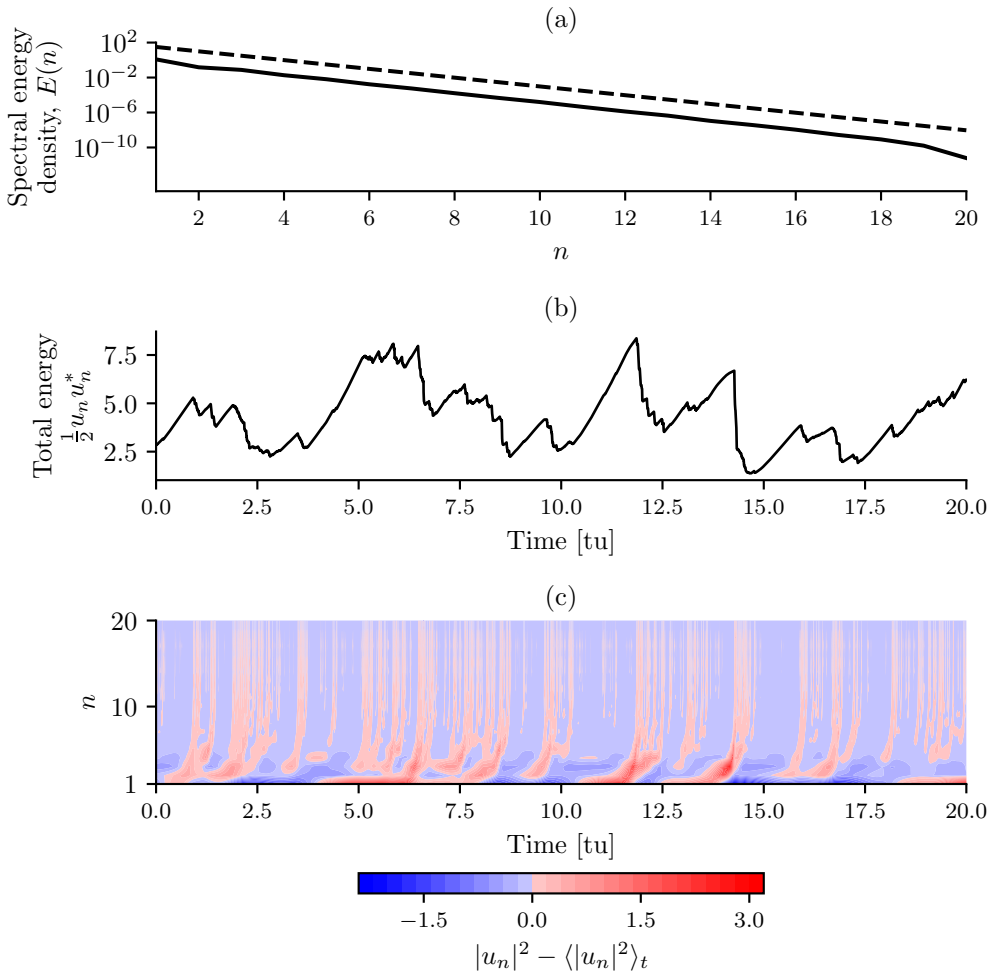


Figure 3.4: The solution to the Sabra shell model presented in three different ways. (a) The average spectral energy density (solid) and the Kolmogorov spectrum (dashed; eq. (2.29)). The average is made over 100tu. (b) The total energy vs time showing the intermittent nature of the flow. (c) Hovmöller diagram for the shell energy anomaly. The energy cascade from large to small scales is evident.

As we will see later, this build-up and dissipation of energy influences the perturbation methods and calls for a separate investigation of the methods in regions dominated by the small and large scales. We define those regions as R_{small} and R_{large}, respectively, and describe in detail an algorithm for separating those regions in appendix D.

As with the Lorentz-63 model, we perform a linear stability analysis of the Sabra shell model to get an intuition on the structure of the flow around the *Kolmogorov fixpoint* of the model. This fixpoint follows the scaling relation $\sim k^{-1/3}$, i.e. equivalent to the Kolmogorov scaling of the velocity (eq. (2.24)). Following sec. 2.1, the (n,m)th component of the Jacobian of eq. (3.9) is given by

$$\begin{aligned}\mathcal{J}_{n,m} = \partial_m \dot{u}_n = \iota k_n \Big(& u^*_{n+1}\delta_{m,n+2} + u_{n+2}\delta_{m,n+1} \\ & - \frac{\epsilon}{\lambda}\big(u^*_{n-1}\delta_{m,n+1} + u_{n+1}\delta_{m,n-1}\big) \\ & - \frac{\epsilon - 1}{\lambda^2}\big(u_{n-2}\delta_{m,n-1} + u_{n-1}\delta_{m,n-2}\big)\Big) - \nu k_n^2, \ \star \end{aligned} \tag{3.11}$$

which is a five-diagonal matrix. Evaluated for 1000 randomly selected velocity profiles u_n of a reference run with the same model configuration as above, we get the eigenvalue spectrum (real part) and eigenvector distribution presented in fig. 3.5. From this figure it is seen that the model has 4 positive eigenvalues that correspond to eigenvectors located in the inertial range; this indicates divergence of the phase space in this range. A correspondence between shell number, n, and eigenvector index, m is seen. The eigenvector with the largest eigenvalue ($m = 1$) is located mainly around the 13th-15th shell. This makes sense because the eddy turnover times of those shells are short, and at the same time the diffusion is negligible; i.e. the fast non-linear dynamics give rise to large eigenvalues.

The eigenvalues for $m \in [5; 12]$ are approximately zero indicating no expansion/contraction in phase space. The corresponding eigenvectors are concentrated at the integral scale and the dynamics thus take place at a long timescale (large eddy turnover time). Thus, the vanished eigenvalues are a natural consequence of the very slowly evolving dynamics.

For $m > 12$ the eigenvalues turn negative which indicates contraction in phase space. This makes sense because the corresponding eigenvectors are located around the Kolmogorov scale. A clear correspondence between n and m is again seen, which becomes very strong for $m \geq 17$ due to the dominance of the diffusion.

The (slightly skewed) symmetry of the eigenvectors around $m \approx 9$ is noteworthy (denoted the V-shape for later reference). Looking specifically at the inertial range and the corresponding eigenvalues, this indicates that the phase space, locally, is characterized by both expansion and contraction for a given shell number. To investigate this further, the projectibility (real part) of pairs of vectors symmetrically located around $m = 9$ is plotted in fig. 3.6. We see that on average, the vectors are approximately orthogonal to each other.

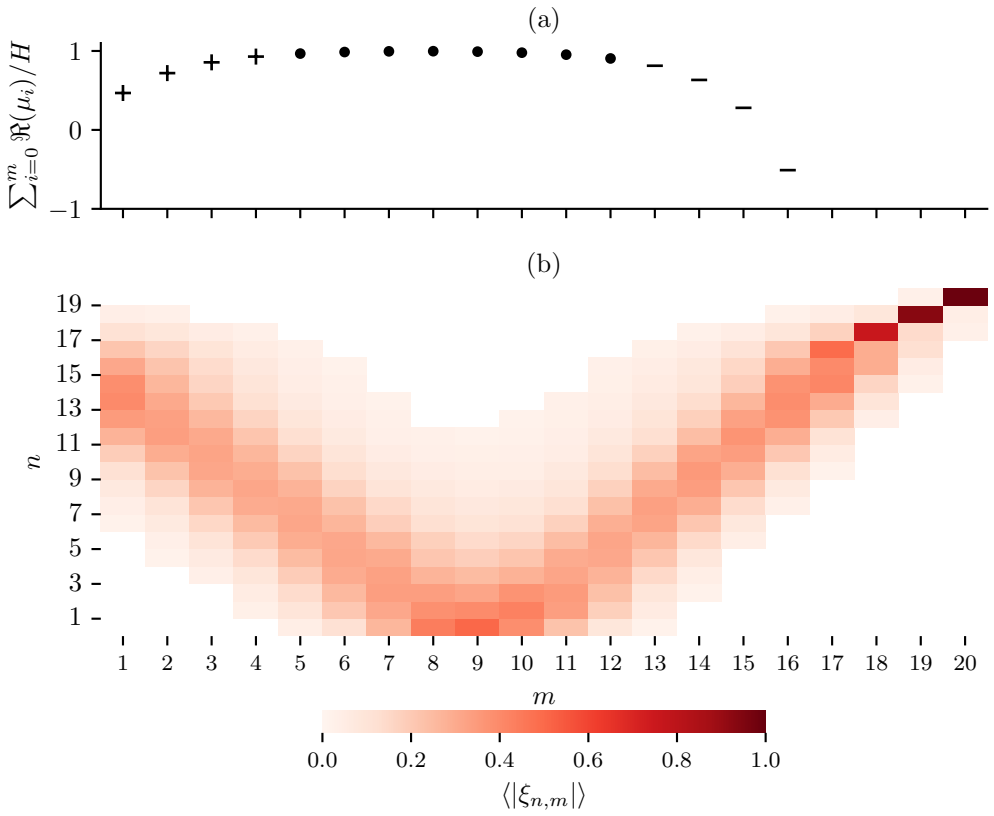


Figure 3.5: The average spectrum of eigenvalues (real part) and distribution of eigenvectors of the Sabra shell model calculated from 1000 randomly chosen positions in time. Sorted in descending order according to the eigenvalues (a) The cumulative eigenvalue spectrum normalized with the Kolmogorov-Sinai entropy $H = \sum_{\Re(\mu_m)>0} \Re(\mu_m)$ (Ditlevsen, 2011). Only eigenvalues up to $m = 16$ are shown to avoid the most negative eigenvalues to dominate the plot completely. The sign of $\Re(\mu_m)$ is shown with $+$, $\cdot$ and $-$, where $\cdot$ means approximately zero. (b) The squared absolute vector components of the eigenvectors. Values below 0.02 is not plotted in order to highlight the relevant components of the vectors.

This behaviour can be understood from Liouville's theorem which says that for a dynamical system governed by Hamiltonian equations, the phase space is incompressible, i.e. eq. (2.19) holds with $\partial_i = \partial/\partial u_i$ being the derivative of the phase space coordinates (Ditlevsen, 2011). In the inertial range, the diffusion term of the Sabra shell model is negligible and the forcing term is zero. For that reason the model approximately fulfils the theorem since the non-linear terms on the RHS of eq. (3.9) are independent of u_n. So the local contraction and expansion of phase space for shells in the inertial range can be seen as a natural consequence of the phase space being incompressible

in this range.

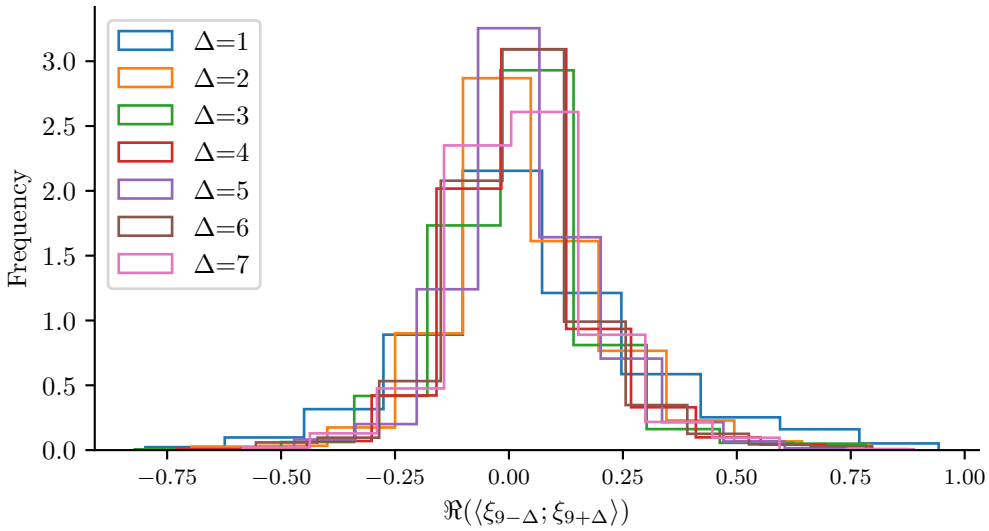


Figure 3.6: The projectibility (real part) of eigenvectors symmetrically located around $m = 9$.

4 Perturbation Methods

In this chapter, we will present all perturbation methods studied except the Lyapunov vector method the theory of which has been presented in sec. 2.2. In the last section, we outline the theoretical relationship between the singular vectors and the normal modes, breeding vectors and Lyapunov vectors.

As described in the introduction (ch. 1), the perturbations of the initial conditions should be made in order to 1) obtain an ensemble spread comparable to the RMSE and 2) to sample the subspace of the phase space that spans the directions associated with dynamical instabilities. Since the number of degrees of freedom of typical NWP models is very large (see sec. 3.2), it is very ineffective to sample this subspace without doing some pre-analysis of the phase space[1] (Molteni et al., 1996).

Methods that identify the subspace are defined as constrained methods in contrary to unconstrained methods such as random sampling (RD method) of the complete phase space (L. Magnusson et al., 2008). In addition to sampling the subspace ineffectively, unconstrained methods produce vectors that most likely will not be in line with the underlying dynamics; in other words, the vector points away from the attractor of the dynamical system. A perturbation initiated along such a vector will die out rapidly, which in the Lorentz-63 model is related to the strong attraction towards the strange attractor, while in real NWP models it can be explained by the generation of gravity waves (Magnusson et al., 2009, 2008).

We now continue with the presentation of the methods. While the normal mode method uses the eigenvectors of the Jacobian of the model, the singular vector method integrates the tangent linear model and its adjoint. In contrast, the breeding vector method performs several cycles of integrations of the full non-linear model followed by rescaling in order to *breed* the optimal perturbation vectors. Finally, the random field method takes a very simple approach and defines perturbations according to scaled differences between

[1] In other words the probability of sampling a direction associated with a dynamical instability is low

historical state vectors.

4.1 Normal mode method ($\mathrm{NM_{mth}}$)

The normal mode method determines the directions of the fastest error growth by performing an eigenvalue and -vector analysis at the initial time of an ensemble run (see sec. 2.1). The eigenvector with the largest eigenvalue (real part) is used to initiate a perturbation, since by construction the normal mode that corresponds to this eigenvector, -value pair dominates the error growth (see eq. (2.10)). If the eigenvalue is complex, the complex-conjugate eigenvector pair spans a plane in which the fastest growing error will lie. In that case multiple perturbations are initiated at random in this plane (L. Magnusson et al., 2008).

For the Lorentz-63 model the state vector $\mathbf{x}$ is real, and for that reason the perturbations need also to be real even though the eigenvalue and -vector are complex. We can ensure this by exploiting that for any real matrix, complex eigenvalues arise in conjugate pairs together with a conjugate pair of eigenvectors, i.e.

$$\mu_1 = \mu_2^*, \;\; \boldsymbol{\xi}_1 = \boldsymbol{\xi}_2^*. \tag{4.1}$$

Writing out the contribution from those two eigenmodes to eq. (2.9) reads

$$\alpha_1 \boldsymbol{\xi}_1 e^{\mu_1 t} + \alpha_2 \boldsymbol{\xi}_2 e^{\mu_2 t} = \alpha_1 \boldsymbol{\xi}_1 e^{\mu_1 t} + \alpha_2 \boldsymbol{\xi}_1^* e^{\mu_1^* t}. \tag{4.2}$$

Define $\boldsymbol{\xi}_1 \equiv \mathbf{u} + \iota\mathbf{w}$, $\mu_1 \equiv \rho + \iota\sigma$. By letting $\alpha_1 = \alpha_2^* \equiv \beta + \iota\gamma$, it is seen that the contribution from the eigenmodes is ensured to be real:

$$\begin{aligned}
\alpha_1 \boldsymbol{\xi}_1 e^{\mu_1 t} + \alpha_2 \boldsymbol{\xi}_1^* e^{\mu_1^* t} = \;& (\beta + \iota\gamma)(\mathbf{u} + \iota\mathbf{w}) e^{(\rho + \iota\sigma)t} + (\beta - \iota\gamma)(\mathbf{u} - \iota\mathbf{w}) e^{(\rho - \iota\sigma)t} && \text{(4.3a)} \\
=\; & \big[(\beta + \iota\gamma)(\mathbf{u} + \iota\mathbf{w})(\cos(\sigma t) + \iota\sin(\sigma t)) \\
& \quad + (\beta - \iota\gamma)(\mathbf{u} - \iota\mathbf{w})(\cos(\sigma t) - \iota\sin(\sigma t))\big] e^{\rho t} && \text{(4.3b)} \\
=\; & 2e^{\rho t}(\beta\cos(\sigma t) - \gamma\sin(\sigma t))\mathbf{u} \\
& \quad - 2e^{\rho t}(\beta\sin(\sigma t) + \gamma\cos(\sigma t))\mathbf{w} && \text{(4.3c)}
\end{aligned}$$

Since the Jacobian of the Lorentz-63 model is real, we initiate a perturbation in the plane of a pair of complex-conjugate eigenvectors by choosing the pair of coefficients α_i to also be complex-conjugate. Beside this constraint the α_i are chosen at random.

4.2 Singular vector method ($\mathrm{SV_{mth}}$)

We present the formalism behind the singular vector method by following R. Buizza et al., 1995 and start off from eq. (2.6).

The forward propagator can be split up into successive integrations of smaller time steps:

$$L(t_0, t_n) = L(t_{n-1}, t_n) \cdots L(t_1, t_2) L(t_0, t_1). \tag{4.4}$$

In simple cases L will take the form of an $m \times m$ matrix, while for more complicated cases it comprises an algorithm as part of e.g. a NWP system (Kalnay, 2002).

We now look at the perturbation norm[2] at time t of the perturbation vector $\mathbf{x}'(t)$:

$$||\mathbf{x}'(t)||^2 \equiv \langle \mathbf{x}'(t); \mathbf{x}'(t) \rangle = \langle L\mathbf{x}'(t_0); L\mathbf{x}'(t_0) \rangle = \langle \mathbf{x}'(t_0); L^* L\mathbf{x}'(t_0) \rangle, \tag{4.5}$$

where L^* is the adjoint of the forward propagator (see eq. (2.13)). The adjoint propagates a perturbation backwards in time, which can best be understood by taking the adjoint of eq. (4.4) and using the property of the adjoint of a product:

$$L^*(t_0, t_n) = L^*(t_0, t_1) \cdots L^*(t_{n-2}, t_{n-1}) L^*(t_{n-1}, t_n) \tag{4.6}$$

Acted on from the right to left, a perturbation is integrated from t_n to t_{n-1}, t_{n-1} to t_{n-2} etc., i.e. backwards in time.

Using eq. (4.5) the error growth of the perturbation from t_0 to t can be measured by the *amplification factor* (Diaconescu et al., 2012)

$$\sigma^2 = \frac{||\mathbf{x}'(t)||^2}{||\mathbf{x}'(t_0)||^2} = \frac{\langle \mathbf{x}'(t_0); L^* L\mathbf{x}'(t_0) \rangle}{||\mathbf{x}'(t_0)||^2}. \tag{4.7}$$

Thus, to maximize the error growth one needs to maximize this quantity.

Since the operator L^*L is normal, that is $(L^*L)^* L^* L = L^* L (L^* L)^*$, while L and L^* on their own are not, the quotient in eq. (4.7) takes the form of a Rayleigh quotient[3] for L^*L. The maximum of such a quotient is the

[2]The choice of norm influences the resulting SVs and makes the vectors orthogonal under this norm. R. Buizza et al., 1995 uses the *energy norm* defined with the vorticity and divergence components of the vector being normed as well as the temperature field. In this thesis, we follow (L. Magnusson et al., 2008) and use a simple $\mathrm{L_2}$ norm to be able to compare.

[3]A quotient on the form $\frac{\mathbf{x}^* \mathbf{M} \mathbf{x}}{\mathbf{x}^* \mathbf{x}}$, where M is a complex Hermitian matrix and $\mathbf{x}$ is a non-zero vector

maximum eigenvalue of L^*L (Ghojogh et al., 2019), which is found by solving the eigenvalue problem

$$(L^*L)\mathbf{v}_i(t_0) = \sigma_i^2 \mathbf{v}_i(t_0). \ \star \tag{4.8}$$

Here $\mathbf{v}_i$ is the set of orthonormal eigenvectors with real eigenvalues $\sigma_i^2 \geq 0$ (R. Buizza et al., 1995; Diaconescu et al., 2012). The eigenvectors at time t_0 evolve into

$$\mathbf{v}_i(t) = L\mathbf{v}_i(t_0) \tag{4.9}$$

at the future time t, which satisfy the eigenvalue problem

$$(LL^*)\mathbf{v}_i(t) = \sigma_i^2 \mathbf{v}_i(t). \ \star \tag{4.10}$$

The σ_i are the singular values of the operator L, while the vectors $\mathbf{v}_i(t_0)$, $\mathbf{v}_i(t)$ are the singular vectors of L at *initial* and *final* (or *optimization*) time, respectively. If sorting the singular values in descending order, both the initial and final singular vector $\mathbf{v}_1(t_0)$, $\mathbf{v}_1(t)$ that correspond to σ_1 are associated with the maximum perturbation growth over the interval $t - t_0$. Both vectors are also referred to as leading singular vectors (R. Buizza et al., 1995).

In fig. 4.1 this is visualized by applying the propagators to a unit sphere initially spanned by the first two normalized singular vectors. Applying L to

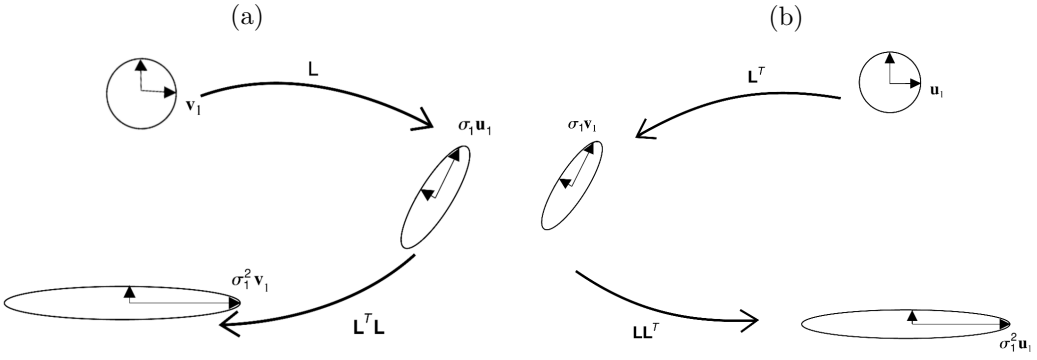


Figure 4.1: The figure visualizes the effect of applying the (a) L^*L and (b) LL^* operators on a unit sphere initially spanned by the first two normalized (a) initial and (b) final singular vectors. From (Kalnay, 2002), in which only the case of real L is considered. The transpose, T, shown in the figure generalizes to the conjugate transpose for complex L. Also, $\mathbf{v}$ corresponds to $\mathbf{v}(t_0)$ and $\mathbf{u}$ to $\mathbf{v}(t)$.

the sphere in fig. 4.1a spanned by the initial singular vectors stretches and rotates it into a tilted ellipsoid with semi-major axis σ_1. Applying L^* to this

ellipsoid further stretches the semi-major axis to σ_1^2 and rewind the rotation. Similarly, the effect of L^* followed by L applied on a sphere initially spanned by the final singular vectors is shown in fig. 4.1b, which stretches the sphere by σ_1^2 as well.

4.3 Breeding vector method ($\mathrm{BV}_{\mathrm{mth}}$)

The breeding method is a simple method to generate perturbation vectors that lead to large error growth, since it doesn't imply any analysis of the tangent linear model or its adjoint. Instead, it works by performing a sequence of integrations of the full non-linear weather prediction model (see fig. 4.2) (R. Buizza et al., 1995; Toth et al., 1993).

In each integration cycle (also called *breeding cycle*), an unperturbed control forecast is made together with one or more forecasts with perturbed initial conditions. For the first cycle a random perturbation is used, e.g. at t_{-n} in fig. 4.2. After each cycle, the difference between the perturbed and the control forecast is rescaled to have the same norm as the initial perturbation. The rescaled difference is then used as the perturbation of the initial conditions of the control forecast for the next integration cycle starting at t_{-n+1} (R. Buizza et al., 1995). After n cycles, the final rescaled perturbation vectors are the *breed vectors* valid at t_0.

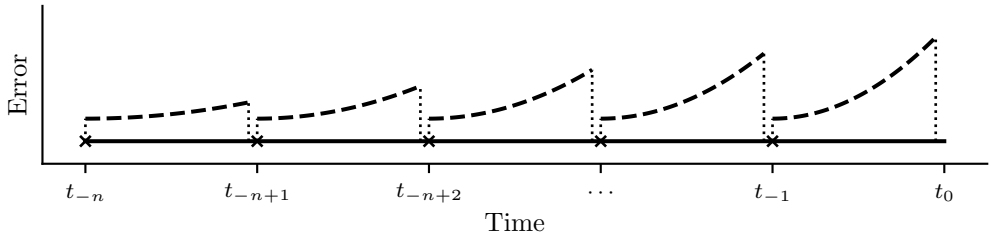


Figure 4.2: Conceptual visualization of the breeding method. An initial random perturbation (vertical dotted line at t_{-n}) is integrated by the non-linear model (dashed line). At regular intervals (crosses), the integrated perturbation is rescaled and added to an unperturbed control forecast (solid line) visualized by the vertical dotted lines at $t_{-n+1}, t_{-n+2}, \ldots, t_{-1}$. Each integration and subsequent rescaling is a breeding cycle. The final evolved and rescaled perturbation is the breeding vector valid at time t_0.

As will be explained in sec. 4.6.3, the breed vectors converge to the leading LLV for sufficiently many breeding cycles. This implies that all the breed vectors that are optimized for the same time t_0 to some extent will be linearly

dependent (L. Magnusson et al., 2008). As we will see through numerical simulations (ch. 6), this has consequences for the method's ability to produce large ensemble spreads.

4.4 Orthogonal breeding vector method (BV-EOF$_{\text{mth}}$)

One can overcome the problem with linear dependency by calculating an orthogonal complement to the breed vectors (L. Magnusson et al., 2008). Following L. Magnusson et al., this is done by calculating the *empirical orthogonal functions* (EOF) of the breed vectors.

The EOF analysis is a principal component analysis that works by identifying the (orthogonal) directions accounting for the largest variances in some dataset. It is commonly used to reduce the dimensionality of e.g. a high dimensional atmospheric phase space, which makes it possible to visualize structures in the phase space, to validate the performance of an ensemble forecast (Franco Molteni et al., 1999) or in this context to calculate an orthogonal complement to a set of vectors. In the following, the mathematical framework for calculating the BV-EOFs is outlined.

The EOF analysis constitutes a singular value decomposition of the space-covariance matrix of the $(n \times m)$ matrix $\mathbf{B}$, where the columns of $\mathbf{B}$ in this context are taken as the n-dimensional breed vectors, i.e.

$$\mathbf{B} = [\mathbf{x}_1 - \mathbf{x}_c, \mathbf{x}_2 - \mathbf{x}_c, \ldots, \mathbf{x}_m - \mathbf{x}_c], \tag{4.11}$$

where $\mathbf{x}_i$ is the state vector of the ith perturbed member of an ensemble with size m, and $\mathbf{x}_c$ is the state vector of the control member (L. Magnusson et al., 2008). The singular value decomposition can be written as

$$\mathbf{B} = \mathbf{ESP^T}, \tag{4.12}$$

where the columns of the $(n \times m)$ matrix $\mathbf{E}$ are the BV-EOFs $\mathbf{e}_i$; the columns of the $(m \times m)$ matrix $\mathbf{P}$ are the principal components (PCs), which describe how much the individual BV-EOFs project onto the breed vectors; and the $(m \times m)$ diagonal matrix $\mathbf{S}$ contains the standard deviations s_i explained by the $\mathbf{e}_i$. We normalize the $\mathbf{e}_i$ and $\mathbf{p}_i$ in the following way

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} \tag{4.13a}$$

$$\langle \mathbf{p}_i, \mathbf{p}_j \rangle = m\delta_{ij}, \tag{4.13b}$$

We can now turn eq. (4.12) into an eigenvalue problem for the PCs by multiplying with $\mathbf{B}^T$ from the left on both sides:

$$\mathbf{B}^T\mathbf{B} = (\mathbf{ESP}^T)^T\mathbf{ESP}^T \tag{4.14a}$$

$$= (\mathbf{SP}^T)^T(\mathbf{E}^T\mathbf{E})\mathbf{SP}^T \tag{4.14b}$$
$$= \mathbf{PS}^T(\mathbf{E}^T\mathbf{E})\mathbf{SP}^T \tag{4.14c}$$
$$= \mathbf{PS}^T\mathbf{ISP}^T, \tag{4.14d}$$

where $\mathbf{I}$ is the identity matrix and eq. (4.13a) is used from third to fourth equal sign. The eigenvalue problem is now obtained by multiplying each side with $\mathbf{p}_i$ and by using eq. (4.13b):

$$\mathbf{B}^T\mathbf{B}\mathbf{p}_i = \mathbf{PS}^T\mathbf{SP}^T\mathbf{p}_i \tag{4.15a}$$
$$= ms_i^2\mathbf{p}_i \ \star \tag{4.15b}$$

From the $\mathbf{p}_i$ we can get the $\mathbf{e}_i$ from the expression

$$\mathbf{e}_i = (ms_i)^{-1}\mathbf{B}\mathbf{p}_i, \ \star \tag{4.16}$$

which is obtained from eq. (4.12) by multiplying with $\mathbf{P}$ from the right and using eq. (4.13b):

$$\mathbf{BP} = \mathbf{ESP}^T\mathbf{P} \tag{4.17a}$$
$$\mathbf{BPS}^T = \mathbf{ESS}^T m \tag{4.17b}$$
$$(m\mathbf{SS}^T)^{-1}\mathbf{BPS}^T = \mathbf{E}, \tag{4.17c}$$

When perturbing a new forecast based on the BV-EOFs, they are first sorted in descending order according to the variance s_i^2, such that $\mathbf{e}_1$ is the BV-EOF that corresponds to the direction of the largest variance in the BV subspace.

4.5 Random field method ($\mathrm{RF_{mth}}$)

As mentioned in the beginning of this chapter, unconstrained perturbation methods, such as random sampling of the phase space, are typically not in line with the underlying dynamics and fail to sample the directions of unstable dynamics effectively. In (Magnusson et al., 2009) the authors introduce the random field method to produce random perturbations of NWP models that have realistic constraints. It constructs a perturbation by rescaling the difference between two randomly chosen historical state vectors of the system to an appropriate amplitude. I.e. a perturbation vector $\mathbf{x}'(t)$ is calculated as

$$\mathbf{x}'(t) = \frac{\mathbf{s}_{t_1} - \mathbf{s}_{t_2}}{|\mathbf{s}_{t_1} + \mathbf{s}_{t_2}|}, \tag{4.18}$$

where $\mathbf{s}_{t_1}, \mathbf{s}_{t_2}$ are the state vectors of the system at times t_1, t_2, respectively[4]. Magnusson et al. constraints the choice of state vectors in the context of NWP models by requiring that they belong to the same season but different years. In that way they are ensured to be uncorrelated, while still preserving some characteristics of the atmospheric variability for that season.

In the context of the Lorentz-63 model, L. Magnusson, 2009 defines the constraint such that the state vectors belong to the same wing of the attractor of the system as the initial point. To ensure that they are also uncorrelated, we impose that they should be separated at least 13 times the average residence time within one wing (see sec. 3.1). The wings of the attractor are separated by a simple split at $x = 0$ as seen in fig. 4.3.

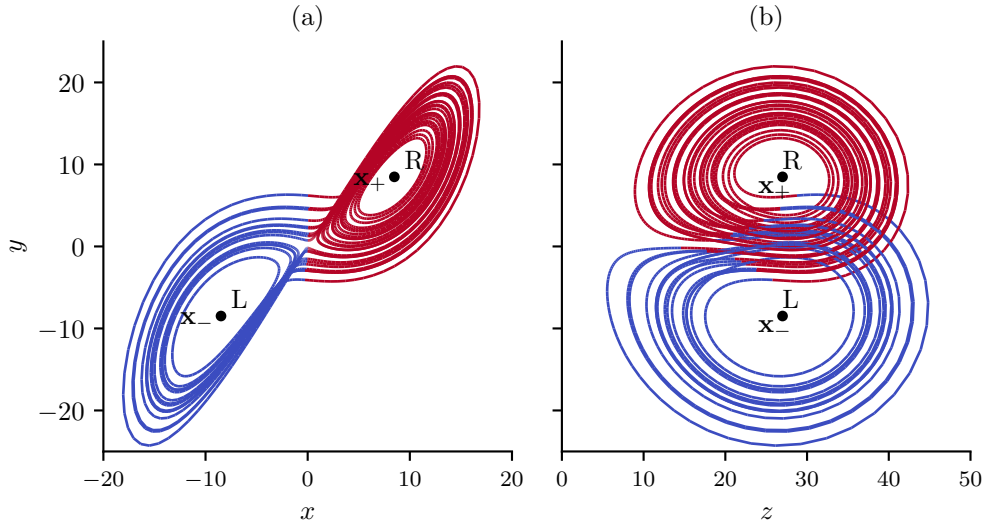


Figure 4.3: The split of the attractor of the Lorentz-63 model at $x = 0$ seen in (a) the x-y plane and (b) the y-z plane.

For the Sabra shell model we use the region analysis presented in appendix D to define the constraints: the chosen state vectors have to belong to the same type of region as the initial point. First, two specific regions to sample from are chosen at random. Then, the state vectors are chosen at random within those regions. Since the minimum duration of a region

[4]In (Magnusson et al., 2009) a tuning parameter α is multiplied on the RHS in order to tune the amplitude of the perturbations such that the ensemble spread is comparable for the different perturbation methods studied. In that way Magnusson et al. avoids differences in various performance scores of the ensembles due to differences in ensemble spread. We do not analyse the ensemble spread and ensemble performance scores in this thesis and thus do not work with this parameter.

(0.05tu) is an order of magnitude larger than the valid time of the TLM (see sec. 5.2.1), we put no effort into ensuring the state vectors to be uncorrelated to each other and the initial point, because the non-linear dynamics makes the chance of choosing correlated state vectors very low.

Since the difference between the state vectors on average will follow the Kolmogorov spectrum, $\mathbf{x}'(t)$ will be dominated by the large scales (i.e. slowly evolving dynamics) and obtain very small error growth. To prevent this, we normalize $\mathbf{x}'(t)$ to the Kolmogorov scaling of the velocity (eq. (2.24)).

4.6 Relationship between methods

To understand how the above methods relate and why some methods (theoretically) lead to larger error growth of an initial perturbation than others, the relationship between singular vectors and normal modes, local Lyapunov vectors and breeding vectors is outlined in this section. Based on the mathematical formalism of the singular vectors, we will see how the use of the adjoint of the forward propagator (eq. (4.6)) makes the singular vector calculation more optimal than the breeding method, and how the singular vectors, optimized for a sufficiently long time, converge to the local Lyapunov vectors. To begin with, we will explore how the initial and final singular vectors can be related to the normal and adjoint modes.

4.6.1 Singular vectors vs normal and adjoint modes

The relation between singular vectors and normal modes can be understood by performing the linearization in eq. (2.3) about a stationary trajectory $\mathbf{x}(t) = \mathbf{x}_0$. This results in a time independent Jacobian (eq. (2.7)) with normalized time independent eigenvectors $\boldsymbol{\xi}_i$ and -values μ_i, and the solution of eq. (2.4) is given by the normal modes (eq. (2.10)).

We can now relate eq. (2.6) to eq. (2.8) by writing

$$L(t_n, t_0) = e^{\boldsymbol{\mathcal{J}}(t_n - t_0)} \tag{4.19}$$

and noting that the requirement that the time t_n should be close to t_0 is no longer necessary. As noted by R. Buizza et al. $\boldsymbol{\mathcal{J}}$ is in general not normal for realistic stationary atmospheric flows meaning that the eigenvectors $\boldsymbol{\xi}_i$ will in general not be orthogonal. The eigenvectors, however, still exist and are also eigenvectors of L due to eq. (2.9). The corresponding eigenvalues of L are given by $e^{\mu_i(t_n - t_0)}$.

Now define the adjoint operator L^* of eq. (4.19) to have (possibly non-orthogonal) normalized eigenvectors $\boldsymbol{\eta}_i$ and -values θ_i. Using eq. (2.13) we

can derive the following: Let $\mathbf{x} = \boldsymbol{\eta}_i$, $\mathbf{y} = \boldsymbol{\xi}_j$, then

$$\langle \boldsymbol{\eta}_i; L(t_n, t_0)\boldsymbol{\xi}_j \rangle = \langle L^*(t_n, t_0)\boldsymbol{\eta}_i; \boldsymbol{\xi}_j \rangle \tag{4.20a}$$

$$\langle \boldsymbol{\eta}_i; \mu_j \boldsymbol{\xi}_j \rangle = \langle \theta_i \boldsymbol{\eta}_i; \boldsymbol{\xi}_j \rangle \ \star \tag{4.20b}$$

$$(\mu_j - \overline{\theta}_i)\langle \boldsymbol{\eta}_i; \boldsymbol{\xi}_j \rangle = 0 \tag{4.20c}$$

where the complex conjugation on θ_i arises due to the definition of the inner product (see Definitions). For $i = j$ eq. (4.20c) is called the *biorthogonality condition* of the eigenvector/adjoint eigenvector pair $\{\boldsymbol{\eta}_i, \boldsymbol{\xi}_i\}$. It holds, however, for $i \neq j$ too and tells that if the pair of eigenvectors are not orthogonal, the corresponding eigenvalues form a complex conjugate pair, i.e. $\mu_j = \overline{\theta}_i$. The projection of $\boldsymbol{\xi}_i$ onto $\boldsymbol{\eta}_i$ is given by $\cos(\alpha_i)$, where α_i is the angle between the vectors in phase space.

Even though the eigenvectors $\boldsymbol{\xi}_i$ may be non-orthogonal, we follow R. Buizza et al. and assumes that the set of normal modes (eq. (2.9)) can be considered complete in a finite dimensional system. We also rewrite eq. (2.9) as a sum of normal mode contributions to the evolution of an initial perturbation:

$$\mathbf{x}'(t_n) = \sum_i c_i \boldsymbol{\xi}_i e^{\mu_i (t_n - t_0)}, \tag{4.21}$$

where the coefficients $\{c_i\}$ are given as $\boldsymbol{\Xi}^{-1}\mathbf{x}'(t_0)$ (refer to eq. (2.9)). We can, however, obtain another expression for $\{c_i\}$ from eq. (4.20c) and eq. (4.21). By projecting $\boldsymbol{\eta}_i$ onto $\mathbf{x}'$ at time $t_n = t_0$ and assuming that $\mu_j \neq \overline{\theta}_i$ for $i \neq j$ we get:

$$\langle \boldsymbol{\eta}_i; \mathbf{x}'(t_0) \rangle = \langle \boldsymbol{\eta}_i; \sum_j c_j \boldsymbol{\xi}_j \rangle = c_i \langle \boldsymbol{\eta}_i; \boldsymbol{\xi}_i \rangle \tag{4.22a}$$

$$c_i = \frac{\langle \boldsymbol{\eta}_i; \mathbf{x}'(t_0) \rangle}{\langle \boldsymbol{\eta}_i; \boldsymbol{\xi}_i \rangle} \tag{4.22b}$$

From the sum in eq. (4.21) it is seen, that the evolution of the initial perturbation will be dominated by the normal mode with the largest (real part) eigenvalue as $t_n \to \infty$. This means also that the dominating singular vector optimized for a very long time will correspond to this normal mode.

The coefficient c_i that corresponds to the dominating normal mode at optimization time can be maximized through maximizing eq. (4.22b). Noteworthy, this is not achieved by letting $\mathbf{x}(t_0) = \boldsymbol{\xi}_1$, i.e. the eigenvector corresponding to the largest (real part) eigenvalue, for which $c_1 = 1$. Instead, if $\mathbf{x}(t_0) = \boldsymbol{\eta}_1$ then $c_1 = 1/\cos(\alpha_1)$, which is called the *projectibility factor* of the normal mode $\boldsymbol{\xi}_1$ and can be much larger than one[5]. This means that

[5] As noted in (R. Buizza et al., 1995), Zhang, 1988 shows that it can take a value of 20 for time averaged wintertime barotropic flows.

the dominating singular vector at initial time will correspond to the adjoint eigenvector with the largest eigenvalue.

The difference between the growth of an initial perturbation, when it is aligned along the dominating normal or adjoint mode at initial time, is illustrated in fig. 4.4. A pair of non-orthogonal normal $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ and adjoint $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$ modes are drawn to the right in the figure. The normal modes are decaying, i.e. they have negative eigenvalues, with $\boldsymbol{\xi}_2$ decaying the fastest. Furthermore, they are orthogonal to the adjoint modes for $i \neq j$ by the assumption $\mu_j \neq \overline{\theta}_i$ and eq. (4.20c).

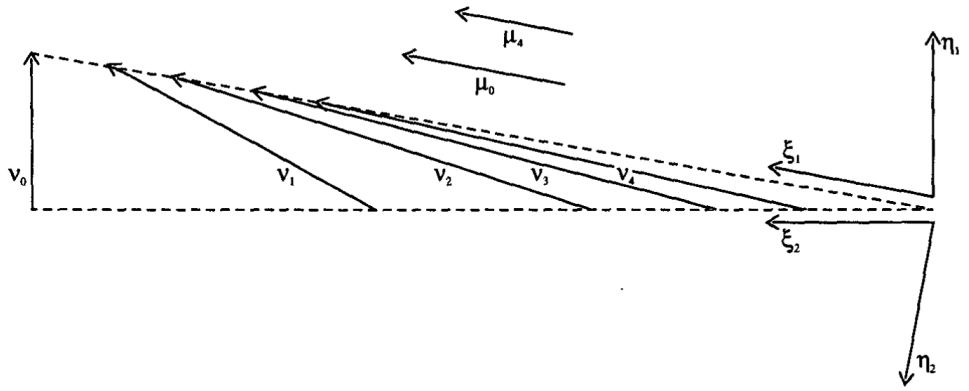


Figure 4.4: Diagram that shows the difference between growth of an initial perturbation along the dominating normal mode and the dominating singular vector (R. Buizza et al., 1995). © American Meteorological Society. Used with permission.

To represent an initial perturbation, the vectors $\boldsymbol{\nu}_0$ and $\boldsymbol{\mu}_0$ are drawn parallel to the adjoint and normal mode, respectively. As time passes by, the vector $\boldsymbol{\nu}_0$ evolve into $\boldsymbol{\nu}_n$ for $n = 1, 2, \ldots, N$, for some finite N. Despite $\boldsymbol{\nu}_0$ being parallel to the adjoint mode, the time evolution is governed by eq. (4.21), which is shown in the figure by evolving the tip and tail of the vector along $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively. In the same way, the vector $\boldsymbol{\mu}_0$ is evolved. It is clear to see how the amplitude of $\boldsymbol{\nu}$ increases while that of $\boldsymbol{\mu}$ decreases, and that the projection of $\boldsymbol{\nu}_n$ onto $\boldsymbol{\xi}_1$ is much larger than that of $\boldsymbol{\mu}_n$. This reflects how a perturbation aligned with the initial singular vector evolve from initial to final time and produce maximal error growth.

The corresponding singular value will in this example depend on the eigenvalue and the projectibility factor of the dominating normal mode. The latter can be understood visually by imagining $\boldsymbol{\xi}_1$ being rotated to project more onto $\boldsymbol{\eta}_1$. The triangle with sides equal to $\boldsymbol{\nu}_0$ and the two dashed lines will then become smaller. This means that the growth of the amplitude of $\boldsymbol{\nu}$ will

be smaller; that is smaller error growth along the dominating singular vector meaning a smaller singular value.

As a final remark, if returning to the case of finite optimization times, the dominating singular vector will not only project onto the dominating normal and adjoint mode at final and initial time, respectively. Instead, it will project onto multiple modes. Likewise, the singular value that corresponds to the dominating singular vector will not be bounded by the projectibility of the dominating normal mode alone (R. Buizza et al., 1995).

4.6.2 Singular vectors vs local Lyapunov vectors

The correspondence between the local Lyapunov vectors and the singular vectors is rather clear, since they both are defined through the L^*L or LL^* operator. The initial SVs thus converge to the adjoint LLVs ($\boldsymbol{\zeta}_i^{-\infty}$) in the limit of $t_n \to \infty$ and fixed t_0 in eq. (4.8) (see also sec. 2.2); the final SVs converge to the LLVs ($\boldsymbol{\zeta}_i^{\infty}$) in the limit of $t_0 \to -\infty$ and fixed t_n in eq. (4.10).

This behaviour is observed in simulations of the Lorentz-63 model (see fig. 4.5), the details of which will be outlined in ch. 5. The dominant initial

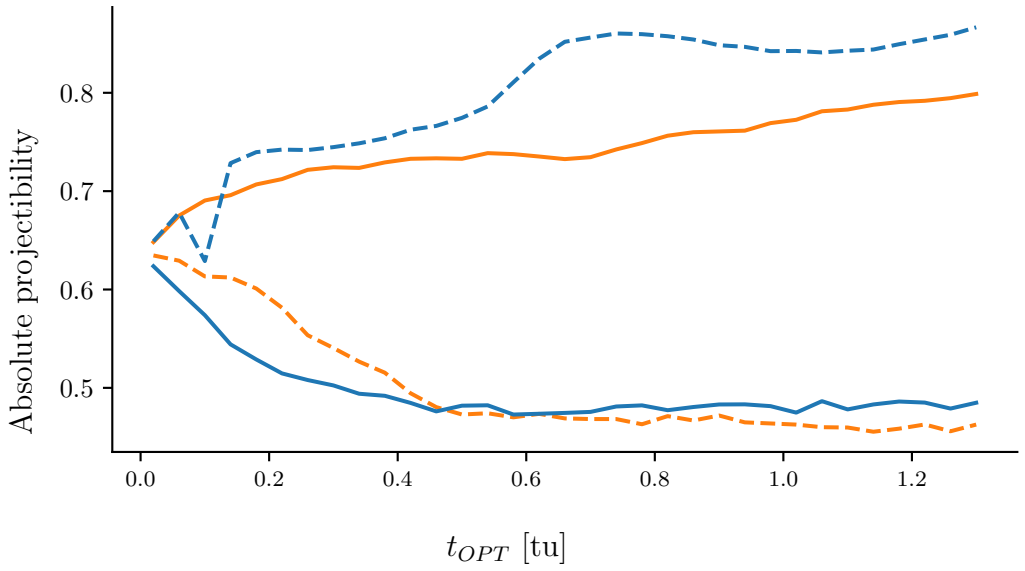


Figure 4.5: The average projection of the dominant initial (blue) and final (orange) singular vector onto the dominant LLV (solid) and adjoint LLV (dashed) as function of the optimization time. The optimization time for the LLV and adjoint LLV is kept constant (see sec. 5.5) The projection is calculated as the inner product defined in Definitions. The average is made from 5000 ensemble runs for each optimization time.

and final SV is calculated for varying optimization time, t_{OPT}, and projected onto the dominant LLV and adjoint LLV (t_{OPT} is kept constant for the LLV and adjoint LLV (see sec. 5.5)).

As we see from the figure the projection of the initial SV onto the LLV decreases with t_{OPT}, whereas the projection onto the adjoint LLV increases. The opposite is seen for the dominant final SV, where the projection onto the LLV increases with t_{OPT}, while the projection onto the adjoint LLV decreases.

4.6.3 Singular vectors vs breeding vectors

To relate the singular vectors to the breeding vectors, we will note that the effect of the rescaling after each breeding cycle is to linearize the perturbation growth. This can be explained by the fact that the non-linear model can be approximated by the forward propagator of the linearized model (eq. (4.4)) for a small enough integration time of one breeding cycle.

For that reason, we can think of each breeding cycle as applying a pseudo-linear propagator $L^{BV}(t_m, t_{m-1})$ to the difference between the perturbed and the control forecast, where $\delta t = t_m - t_{m-1}$ (for $n \geq m \geq 1$) is the length of one cycle. For n cycles the sequence can be written as

$$L^{BV}(t_n, t_{n-1})L^{BV}(t_{n-1}, t_{n-2}) \dots L^{BV}(t_1, t_0) \tag{4.23}$$

Acted upon an initial random perturbation (as explained in sec. 4.3), the resulting breed vector will converge to the dominant LLV for a sufficiently large number of breeding cycles (R. Buizza et al., 1995). This is due to the approximate similarity between $L^{BV}(t_m, t_{m-1})$ and the forward propagator of the linearized model[6].

For any finite time interval δt, however, the growth rate of a breed vector evolved until a given time t_{n-1} will be smaller than or equal to the growth rate of the dominant initial singular vector optimized for the same time. This is seen through the following:

$$\frac{||L(t_n, t_{n-1})\mathbf{x'}^{BV}(t_{n-1})||}{||\mathbf{x'}^{BV}(t_{n-1})||} \leq \max_{\mathbf{x'}(t_{n-1})\neq\mathbf{0}}\Big(\frac{||L(t_n, t_{n-1})\mathbf{x'}(t_{n-1})||}{||\mathbf{x'}(t_{n-1})||}\Big), \tag{4.24}$$

where

$$\mathbf{x'}^{BV}(t_{n-1}) = L^{BV}(t_{n-1}, t_{n-2}) \dots L^{BV}(t_1, t_0)\mathbf{x'}^{BV}(t_0). \tag{4.25}$$

[6]As explained in sec. 2.2 the forward propagator of the linearized model propagates an initial random perturbation into the dominant LLV $\boldsymbol{\zeta}_1^{\infty}$ for sufficiently long integration time.

Here $\mathbf{x}'^{BV}(t_0)$ is the initial perturbation before starting the first breed cycle, i.e. a random perturbation; $\mathbf{x}'^{BV}(t_{n-1})$ is the breed vector that has converged to the LLV; $\mathbf{x}'(t_{n-1})$ is the optimized initial singular vector.

To understand why eq. (4.24) is true lets look at the stationary case as in sec. 4.6.1. For a sufficiently long sequence in eq. (4.25) $\mathbf{x}'^{BV}(t_{n-1})$ will converge to the dominant normal mode, since the LLV for a time-independent Jacobian reduces to the dominant eigenvector of $\boldsymbol{\mathcal{J}}$. This is seen from the forward propagator (eq. (2.12)), which becomes time independent and consequently $S^{\infty}(t_0)$ reduces to the Jacobian. On the contrary $\mathbf{x}'(t_{n-1})$ will converge to the dominant adjoint mode for sufficiently long optimization time as explained in sec. 4.6.1. In that way the LHS and RHS in eq. (4.24) differ by the projectibility factor of the normal mode.

Summing up, eq. (4.24) means that the breeding method produces vectors that are less optimal for perturbing the initial conditions than the singular vectors, optimized for the same period of time, in terms of exponential growth rate. The main reason for this is that the breeding method does not make use of the adjoint equations, which on the contrary improve its simplicity.

The convergence of the BVs to the dominant LLV is observed in simulations of the Lorentz-63 model (see fig. 4.6). By varying the length of

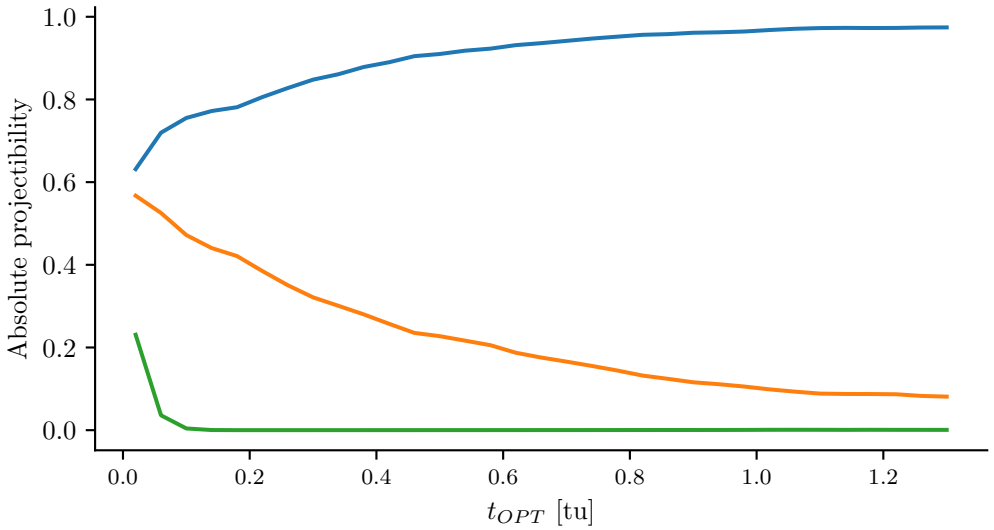


Figure 4.6: As in fig. 4.5 but for breed vectors. The colours denote the projection of the BVs onto the first (blue), second (orange) and third (green) LLV. The number of breeding cycles is kept constant at a value of 4.

the breeding cycle and keeping the number of cycles fixed at a value of 4 we see how the average projection of the BVs onto the dominant LLV converges towards a value close to one. At the same time the projection onto the second and third LLV decreases.

5 Numerical Setup

In this chapter, we first present the numerical setup for solving the two non-linear models as well as the TLM and its adjoint. Then the Lanczos algorithm is introduced, which is used to find the singular vectors and values. In the last two sections, we outline the configuration details of the numerical setup relevant for all models and perturbation methods.

5.1 Non-linear models

In order to solve the Lorentz-63 model (eq. (3.1)) and the Sabra shell model (eq. (3.9)) numerically, we use the fourth order explicit Runge-Kutta (RK4) scheme. We use the same numerical scheme for the two models, despite that such a sophisticated scheme is not necessary for the Lorentz-63 model. Lower order schemes will suffice, provided the time step is small enough to ensure stability. For the Sabra shell model, however, a high order scheme is required in order for the solution to be stable without being forced to use an impractically short time step. This choice enables faster development and debugging of the code, since it can first be tested on the Lorentz-63 model, which is computationally inexpensive to run, and hereafter applied to the Sabra shell model.

The RK4 scheme evaluates the $n+1$ time step of a first order differential equation (eq. (2.1)) as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{1}{6}dt(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) \tag{5.1a}$$

$$t_{n+1} = t_n + dt \tag{5.1b}$$

where

$$\mathbf{k}_1 = f(t_n, \mathbf{x}_n) \tag{5.1c}$$

$$\mathbf{k}_2 = f(t_n + \frac{1}{2}dt, \mathbf{x}_n + \frac{1}{2}dt\mathbf{k}_1) \tag{5.1d}$$

$$\mathbf{k}_3 = f(t_n + \frac{1}{2}dt, \mathbf{x}_n + \frac{1}{2}dt\mathbf{k}_2) \tag{5.1e}$$

$$\mathbf{k}_4 = f(t_n + dt, \mathbf{x}_n + dt\mathbf{k}_3) \tag{5.1f}$$

are the slopes of the solution evaluated at the start-, mid- and endpoint of the time interval $[t_n; t_n+dt]$ (Sandu et al., 2010)[1]. Applied to the two models, we note that f do not explicitly depend on time.

In practice the RK4 scheme requires a time step of order $\sim 10^{-7}$ for the numerical solution to the Sabra shell model to remain stable. This is because the Sabra shell model can be characterized as being a set of stiff equations. Although the stiffness property do not have any precise definition, it generally means that the equations have some terms that give rise to rapid variations in the solution and consequently that some numerical methods require a very short time step. For the Sabra shell model this is attributed to the smallest eddy turnover times found for shells 17-18 (see eq. (B.1b)) and to the quadratic factor in the diffusion term.

We can partly counteract this requirement by using the explicit analytical solution to the diffusion term[2]. Ignore for a moment the non-linear terms in eq. (3.9). Then

$$\dot{u}_n = -\nu k_n^\alpha u_n \Rightarrow \tag{5.2a}$$

$$u_n(t) = Ae^{-\nu k_n^\alpha t}, \tag{5.2b}$$

where A is an arbitrary constant. Now we define $U_n^{NL}(t)$ to be the solution to the non-linear and forcing terms using RK4. The complete numerical solution is then given by

$$u_n(t + dt) = U_n^{NL}(t + dt)e^{-\nu k_n^2(t+dt)} \tag{5.3}$$

Since the quadratic term now appears in the exponential, it does not influence the stability of the numerical solution anymore. So now only the shortest eddy turnover time limits the choice of the time step.

5.2 Tangent linear model and its adjoint

Integration of the TLM on differential from (eq. (2.4)) into integral form (eq. (2.6)) is done using the RK4 scheme. In that way, $L(t_0, t_n)$ really comprises all steps of the RK4 scheme applied successively from time t_0 to time t_n.

[1]Sandu et al. present the implicit Runge-Kutta method, which differs from the explicit Runge-Kutta method by including dependence of k_i on k_j for all j and not only $j < i$. This leads to improved stability, but also to a more complex scheme to implement.

[2]Reference: supervision meeting 12.10.2021 with Peter D. Ditlevsen

The scheme works on the Jacobian of the model (eq. (2.5)) evaluated at the non-linear trajectory. This means that the non-linear model needs to be integrated concurrently with the TLM. Written explicitly the scheme approximates the next time step as (Sandu et al., 2010)

$$\mathbf{x}'_{n+1} = \mathbf{x}'_n + \frac{1}{6}dt(\mathbf{l}_1 + 2\mathbf{l}_2 + 2\mathbf{l}_3 + \mathbf{l}_4) \tag{5.4a}$$

$$t_{n+1} = t_n + dt, \tag{5.4b}$$

where

$$\mathbf{l}_1 = \boldsymbol{\mathcal{J}}|_{\mathbf{X}_1} \cdot (\mathbf{x}'_n) \tag{5.4c}$$

$$\mathbf{l}_2 = \boldsymbol{\mathcal{J}}|_{\mathbf{X}_2} \cdot (\mathbf{x}'_n + \frac{1}{2}dt\mathbf{l}_1) \tag{5.4d}$$

$$\mathbf{l}_3 = \boldsymbol{\mathcal{J}}|_{\mathbf{X}_3} \cdot (\mathbf{x}'_n + \frac{1}{2}dt\mathbf{l}_2) \tag{5.4e}$$

$$\mathbf{l}_4 = \boldsymbol{\mathcal{J}}|_{\mathbf{X}_4} \cdot (\mathbf{x}'_n + dt\mathbf{l}_3), \tag{5.4f}$$

The $\mathbf{X}_i$ are state vectors of the integrated non-linear model, evaluated at the sub grid points of the RK4 method, and are given by

$$\mathbf{X}_1 = \mathbf{x}_n \tag{5.5a}$$

$$\mathbf{X}_2 = \mathbf{x}_n + \frac{1}{2}h\mathbf{k}_1 \tag{5.5b}$$

$$\mathbf{X}_3 = \mathbf{x}_n + \frac{1}{2}h\mathbf{k}_2 \tag{5.5c}$$

$$\mathbf{X}_4 = \mathbf{x}_n + h\mathbf{k}_3, \tag{5.5d}$$

where the $\mathbf{k}_i$'s are defined in eqs. (5.1c) to (5.1f)

Loosely speaking, to integrate the adjoint of the TLM on differential form into eq. (4.6) we have to reverse time in eq. (5.4) and use the adjoint Jacobian, $\boldsymbol{\mathcal{J}}^*$. For that reason the intermediate slopes are calculated in reverse order. Following Sandu et al., 2010, the adjoint of the RK4 scheme for the TLM approximates the adjoint variables $\mathbf{x}'^*_n$ by

$$\mathbf{x}'^*_n = \mathbf{x}'^*_{n+1} + \frac{1}{6}dt(\mathbf{j}_1 + 2\mathbf{j}_2 + 2\mathbf{j}_3 + \mathbf{j}_4) \tag{5.6a}$$

$$t_n = t_{n+1} - dt, \tag{5.6b}$$

where

$$\mathbf{j}_4 = \boldsymbol{\mathcal{J}}^*|_{\mathbf{x}_4} \cdot (\mathbf{x}'^*_{n+1}), \tag{5.6c}$$

$$\mathbf{j}_3 = \boldsymbol{\mathcal{J}}^*|_{\mathbf{x}_3} \cdot (\mathbf{x}'^*_{n+1} + \frac{1}{2} dt \mathbf{l}_4) \tag{5.6d}$$

$$\mathbf{j}_2 = \boldsymbol{\mathcal{J}}^*|_{\mathbf{x}_2} \cdot (\mathbf{x}'^*_{n+1} + \frac{1}{2} dt \mathbf{l}_3) \tag{5.6e}$$

$$\mathbf{j}_1 = \boldsymbol{\mathcal{J}}^*|_{\mathbf{x}_1} \cdot (\mathbf{x}'^*_{n+1} + dt \mathbf{l}_2) \tag{5.6f}$$

and the $\mathbf{X}_i$ is given by eq. (5.5). As for the integration of the TLM, this means that the non-linear model is integrated concurrently with the integration of the adjoint of the TLM.

5.2.1 Verification of the TLM and its adjoint

The calculation of e.g. the local Lyapunov vectors or the singular vectors crucially depends on the implementation of the above numerical schemes to be correct. For that reason a verification is needed.

To verify the integration of the TLM one can assess the error made relative to an integration of the non-linear model. We consider a small initial perturbation at time t_0, $\mathbf{x}'(t_0)$, and integrate it with the TLM. A perturbed and unperturbed reference state vector, $\mathbf{x}(t_0)$, is integrated with the non-linear model. The linear model should then reproduce the difference between the integrations of the non-linear model with error $\mathcal{O}(||\mathbf{x}'||^2)$, i.e.

$$\mathbf{x}'(t) = L(t_0, t)\mathbf{x}'(t_0) = M(\mathbf{x}(t_0) + \mathbf{x}'(t_0)) - M(\mathbf{x}(t_0)) + \mathcal{O}(||\mathbf{x}'||^2) \tag{5.7a}$$

$$E(t) \equiv L(t_0, t)\mathbf{x}'(t_0) - \Big(M(\mathbf{x}(t_0) + \mathbf{x}'(t_0)) - M(\mathbf{x}(t_0))\Big) = \mathcal{O}(||\mathbf{x}'||^2) \tag{5.7b}$$

To verify the integration of the adjoint of the TLM the identity

$$(L(t_0, t)\mathbf{x}'(t_0))^*(L(t_0, t)\mathbf{x}'(t_0)) = \mathbf{x}'(t_0)^* L^*(t_0, t)(L(t_0, t)\mathbf{x}'(t_0)), \tag{5.8}$$

is tested ((Navon et al., 1992) as cited in (Kalnay, 2002))[3]. While the LHS only involves integration of the TLM, the RHS involves both integration of the TLM and its adjoint. According to Kalnay, the integration of the adjoint will be correct if this identity holds up to *machine precision*. It is unclear what Kalnay means precisely with machine precision, and it will probably be different from today's machine precision. We will just take it as a quantity $\ll 1$ and present the numerical results.

[3]In (Kalnay, 2002) the identity is stated for real matrices/operators only with the transpose instead of the conjugate transpose

Since calculations of the LHS and RHS for both models turned out to be *very* small, we used the difference relative to the mean of the LHS and RHS as the verification measure, i.e.

$$\mathcal{V}_{ATLM} = \frac{|(L\mathbf{x}'(t_0))^*(L\mathbf{x}'(t_0)) - \mathbf{x}'(t_0)^*L^*(L\mathbf{x}'(t_0))|}{|\frac{1}{2}\Big((L\mathbf{x}'(t_0))^*(L\mathbf{x}'(t_0)) + \mathbf{x}'(t_0)^*L^*(L\mathbf{x}'(t_0))\Big)|}. \tag{5.9}$$

The results of the verification of the linear models through eqs. (5.7b) and (5.9) is presented in the following. Refer to secs. 5.4 and 5.5 for the parameters used.

Lorentz-63 model

For the Lorentz-63 model the average $E(t)$ (eq. (5.7b)) is shown in fig. 5.1a for a random initial perturbation. The average is based on 1000 independent integrations. For comparison a logarithmic and a linear function is fitted to the error data in two different time ranges (see figure caption for the functional form of the fits). For the logarithmic fit $]0; 0.5]$tu is used, while for the linear fit $[1.0; 3.0]$tu is used. The parameter a of the logarithmic fit is estimated to 1.4. Although other functions could fit the $]0; 0.5]$tu as well as the logarithmic, a transition from this range, possibly governed by polynomial error growth, to the $[1.0; 3.0]$tu range, governed by exponential error growth, is clearly seen. The point where the logarithmic fit and the error diverges gives an estimate of ~ 0.3tu for the time up until which the TLM is valid.

Concerning the adjoint of the TLM, the implementation seems to be correct, since an integration interval $t - t_0 > 1$tu of the forward/backward propagators is needed in order to even observe a finite value for eq. (5.9). The average of 1000 independent estimates of the logarithm of $\mathcal{V}_{ATLM}$ with $t - t_0 = 5$tu gives $\ln(\mathcal{V}_{ATLM}) \approx -29$.

Sabra shell model

The results of similar calculations with the TLM of the Sabra shell model are shown in figs. 5.1b and 5.1c calculated in R_{large} and R_{small}, respectively. The time ranges for the logarithmic fits are $]0; 0.004]$tu and $]0; 6 \cdot 10^{-4}]$tu, respectively, while those for the linear fit are $[0.008; 0.02]$tu and $[0.002; 0.02]$tu. The parameter a of the logarithmic fits are estimated to 0.7 and 1.3, respectively. Again we see a clear transition to an exponential error growth, and thus get an estimate of ~ 0.005tu and ~ 0.002tu, respectively, for the time up until which the TLM is valid.

The verification of the adjoint of the TLM yields an average value of $\ln(\mathcal{V}_{ATLM}) \approx -7$ based on 1000 independent estimates and $t - t_0 = dt$ (see sec. 5.4). Although not as obvious as for the Lorentz-63 model, the adjoint code is still considered to be implemented correct based on this verification. The difference in integration length and value for $\mathcal{V}_{ATLM}$ between the models is probably due to the larger dimension and number of distinct timescales of the Sabra shell model compared to the Lorentz-63 model.

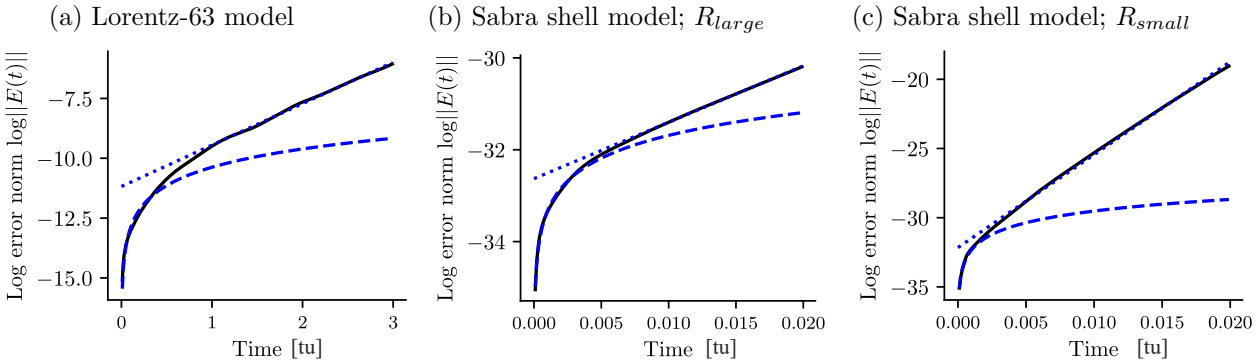


Figure 5.1: The average error vs time of the TLM integration of a small perturbation compared to the difference between a perturbed and unperturbed integration of the non-linear model (black) as stated in eq. (5.7b). Lines in blue represent fits to a logarithmic (dashed, $f(x) = a\ln(x) + b$) and linear (dotted, $f(x) = ax + b$) function in specific time ranges. (a) Lorentz-63 model; logarithmic fit in $]0; 0.5]$tu with $a = 1.1$, $b = -10.4$; linear fit in $[1.0; 3.0]$tu with $a = 1.7\text{tu}^{-1}$, $b = -11.2$. (b) Sabra shell model in R_{large}; logarithmic fit in $]0; 0.004]$tu with $a = 0.7$, $b = -28.4$; linear fit in $[0.008; 0.02]$tu with $a = 122.9\text{tu}^{-1}$, $b = -32.6$. (c) as in (b) but in R_{small}; logarithmic fit in $]0; 6 \cdot 10^{-4}]$tu with $a = 1.2$, $b = -24$; linear fit in $[0.002; 0.02]$tu with $a = 671.7\text{tu}^{-1}$, $b = -32.2$.

5.3 Lanczos algorithm

As noted in sec. 4.2 L can for simple models be written as a matrix, and the eigenvalue problem in eq. (4.8) can be solved by diagonalizing L^*L. In more complex problems, where this is not possible or where the number of degrees of freedom is too large for standard matrix algorithms to work, an iterative procedure can be followed to calculate the most dominant singular values and vectors (R. Buizza et al., 1995). One such procedure is the Lanczos algorithm, for which the detailed procedure is outlined in appendix E.

Overall, the algorithm works by constructing an orthonormal basis $\{\mathbf{p}_i\}$ (called the *Lanczos vectors*) of an N-dimensional subspace of the M-dimensional

model space from $N \ll M$ iterations of L^*L acting on a random initial vector. From this basis a tridiagonal matrix is then constructed and diagonalized using standard diagonalization libraries. The resulting eigenvectors are the approximate dominant singular vectors of L and the square root of the corresponding eigenvalues are the approximate singular values (R. Buizza et al., 1995). In that way, the algorithm turns the problem of finding the eigenvalues and -vectors of an $(M \times M)$ matrix/operator into the problem of diagonalizing an $(N \times N)$ tridiagonal matrix.

Unfortunately, the Lanczos algorithm is sensitive to numerical errors due to finite precision arithmetic and does not necessarily preserve the orthogonality between the Lanczos vectors. This means that the tridiagonal matrix is not a proper projection of L^*L onto the $\{\mathbf{p}_i\}$ basis (see eq. (E.1)), and consequently the estimated eigenvalues and -vectors of L^*L will not be accurate. Some eigenvalues may also become very similar, and it is hard to determine if it reflects actually similar eigenvalues of L^*L or if it is a consequence of the finite precision (Meurant et al., 2006)[4].

To suppress the loss of orthogonality various methods have been proposed to re-orthogonalize the Lanczos vectors periodically (Demmel, 1997; Meurant et al., 2006). The simplest but also computationally most expensive way is to do a *full* re-orthogonalisation of the new $\mathbf{p}_{i+1}$[5] with respect to all existing $\mathbf{p}_i$'s. We will follow this approach, since the computational expense is small for the low-dimensional models in use. The re-orthogonalisation is performed in item 3f in appendix E.

5.4 Model configurations

We present in this section the configurations of the numerical setup for the two models and how the models are equilibrated before saving a reference run. The reference run functions as the true state of the model from which perturbations are made as will be described in ch. 6.

5.4.1 Lorentz-63 model

Model parameters:

$$\sigma = 10, \;\; r = 28, \;\; b = 8/3 \tag{5.10}$$

[4]R. Buizza et al. does not mention explicitly that the finite precision arithmetic is the reason, but they consider only the $N/3$ most dominant singular vectors to be estimated accurately.

[5]In practice, it is actually done on the intermediate Lanczos vectors (the $\mathbf{w}_j$'s), which differ from the $\mathbf{p}_{i+1}$'s only by a normalization

Time step:

$$dt = 0.01 \tag{5.11}$$

The time step is chosen to properly resolve the characteristic timescales of the system (see sec. 3.1).

Model equilibration and reference run
The model is run for 10^3tu starting from $\mathbf{x} = [1, 1, 1]$ after which data for 10^5tu is saved as the reference run. This ensures that the reference run is completely independent of the initial state and that the reference trajectory lies on the attractor of the system.

5.4.2 Sabra shell model

Model parameters:

$$\epsilon = 1/2, \;\; k_0 = 1, \;\; \lambda = 2, \;\; N = 20, \;\; f_n = \delta_{n,0} \tag{5.12}$$

Time step:

$$dt = 10^{-5} \tag{5.13}$$

The time step is chosen to properly resolve the smallest timescales of the system and thus ensure the RK4 scheme to be stable (see sec. 5.1).

Model equilibration and reference run
The model is run for 10^3tu starting from $u_n = k_n^{-1/3}$, i.e. the Kolmogorov fixpoint. This corresponds to ~ 425 eddy turnovers of the largest eddy ($n = 0$) (see appendix B.2). This ensures that the model has equilibrated, such that statistical measures can be made without influence from the initial state and the transient period. After this the model is run for $3 \cdot 10^3$tu which is saved with a sampling rate of $1/10$ as the reference run.

Kinematic Viscosity
The kinematic viscosity is determined by f_0 and by requiring the time average energy content to be constant. The energy input per time is $f_0 u_0$ while the energy output per time is $\nu k_K^2 u_K^2$, where K is the shell number that corresponds to the Kolmogorov scale in equilibrium. These two quantities will equal the mean energy flux across scales, which can be quantified by $\bar{\epsilon}$, i.e.

$$\nu k_K^2 u_K^2 \sim \bar{\epsilon} \sim f_0 u_0, \tag{5.14}$$

Using eq. (2.24), this can be manipulated as follows:

$$\nu k_K^2 \bar{\epsilon}^{2/3} k_K^{-2/3} \sim \bar{\epsilon} \sim f_0 \bar{\epsilon}^{1/3} k_0^{-1/3} \tag{5.15a}$$

$$\nu k_K^{4/3} \bar{\epsilon}^{1/3} \sim \bar{\epsilon}^{2/3} \sim f_0 k_0^{-1/3} \tag{5.15b}$$

Inserting eq. (2.25e) into the first and the last term we get a relation between ν, f_0 and K:

$$\nu k_K^{4/3} \sim f_0 k_0^{-1/3} k_K^{-4/3} \nu^{-1} \tag{5.16a}$$

$$\nu^2 k_K^{8/3} \sim f_0 k_0^{-1/3} \tag{5.16b}$$

$$\nu \sim \sqrt{\frac{f_0}{k_0^{1/3} k_K^{8/3}}} = \sqrt{\frac{f_0}{k_0^3 \lambda^{8K/3}}}, \tag{5.16c}$$

where in the last step we have used eq. (3.8). The model is configured to have

$$K = 19 \tag{5.17a}$$

$$\nu \approx 2.37 \cdot 10^{-8} \mathrm{m}^2/\mathrm{s} \tag{5.17b}$$

5.5 Method configuration

In this section we present the numerical configuration of the different perturbation methods. In this and future chapters, we will refer to the cycle length in the BV method, the integration time in the LLV calculation, and the optimization time in the SV method collectively as the optimization time, t_{OPT}.

The choice of t_{OPT} in the BV method has a large impact on what dynamical structures that are bred upon; a short t_{OPT} will not breed on dynamical instabilities that evolve on long time scales and vice versa. The t_{OPT} influences the resulting singular vectors in the SV method similarly, despite the optimization is made with the TLM and its adjoint.

For the SV method, the choice of t_{OPT} is made on the basis of the verification of the TLM presented in sec. 5.2.1 to ensure that the optimization time is within the linear range. To be able to compare the BV method with the SV method the same t_{OPT} is used. The t_{OPT} in the LLV method is chosen to be considerably larger than a characteristic timescale of the models while still within a computational practical range.

5.5.1 Lorentz-63 model

The t_{OPT} in the LLV method is chosen based on the mean residence time in a wing (see sec. 3.1); for $t_{OPT} = 20$tu this corresponds to ~ 15 regime shifts. The norm of the perturbations is set to $||\mathbf{x}'|| = 0.01$ as in (L. Magnusson et al., 2008), which is small compared to the norm of a state vector on the attractor.

For the RF method, we choose historical states from the first 10^4tu of the reference run. For that reason, all investigations with the perturbation methods are started after this time.

Table 5.1: Parameters for the configuration of the BV_{mth}, LLV_{mth} and SV_{mth} in the Lorentz-63 model

$\mathbf{BV_{mth}}$		$\mathbf{LLV_{mth}}$		$\mathbf{SV_{mth}}$	
Number of cycles	4	Integration time	20.0tu	Optimization time	0.1tu
Cycle length	0.1tu	Rescale interval	4.0tu		
Number of vectors	8				

5.5.2 Sabra shell model

We investigate the perturbation methods separately in regions dominated by the small and large scales, and use for that reason two different optimization times in the BV and SV methods.

The relevant timescale to choose t_{OPT} for the LLV method upon is considered to be the smallest eddy turnover time before viscosity sets in. This corresponds to the shell range in which the most unstable dynamic take place; for $t_{OPT} = 0.1$tu this corresponds to ~ 17 eddy turnovers of shell 17 (see appendix B.2).

The norm of the perturbations is set to $||\mathbf{x}'|| = 10^{-14}$, which is small compared to the norm of the average velocity state vector ($||\langle \mathbf{u} \rangle_t|| \sim 1$m/s), and also small compared to the smallest individual components of this vector ($\sim 10^{-6}$m/s). The latter is needed in order for the perturbations not to dominate completely the smaller scales compared to the underlying dynamics.

Table 5.2: Parameters for the configuration of the BV_{mth}, LLV_{mth} and SV_{mth} in the Sabra shell model

	R_{small}	R_{large}		
$\mathbf{BV_{mth}}$			$\mathbf{LLV_{mth}}$	
Number of cycles	10		Integration time	0.1tu
Cycle length	$5 \cdot 10^{-5}$tu	0.004tu	Rescale interval	0.01tu
$\mathbf{SV_{mth}}$				
Optimization time	$5 \cdot 10^{-5}$tu	0.004tu		

6 Results

In this chapter we present the results of the investigations with the different perturbation methods. We start out with some preliminary remarks about the investigations together with the definition of a terminology used when describing the results.

6.1 Preliminary remarks

In outline, we examine the perturbations on their own e.g. their spatial/spectral distributions, and through ensemble forecasts started from those perturbations. In the latter case we are interested in the ability of the methods to produce large error growth relative to a reference. The investigations made are greatly inspired by L. Magnusson et al., 2008, and this study will work as a reference for comparison of the results for the Lorentz-63 model.

We will consider the reference run as the "true" solution of the model, and perform perturbations and ensemble forecasts of this solution. In relation to a true ensemble system, this compares to perform perturbations and forecasts from the current best estimate for the state of the atmosphere obtained through data assimilation. The main difference is that such an analysis comes with an error, while in this study we have no such error due to perfect knowledge of the models. We do not attempt to simulate this error, since it would simply add an offset to the initial perturbations, which is irrelevant for the comparison of the perturbation methods.

The initial conditions of the members of an ensemble are perturbed according to the various perturbation methods presented in ch. 4 and by the LLVs presented in sec. 2.2. We therefore denote an ensemble by the perturbation method used e.g. the SV ensemble (in short: $\mathrm{SV}_{\mathrm{ens}}$) is perturbed by the SV method. When a perturbation method produces perturbations with a specific order (e.g. the first, second, etc. SV in the SV method), the ensemble members are perturbed in that order e.g. the first SV ensemble member (in short: $\mathrm{SV}^1_{\mathrm{mbr}}$) is perturbed by the first SV (in short: SV^1) etc.

We follow the structure in the rest of the thesis and present the results for the two models separately. Throughout the chapter the following colour-code is used for the different perturbation methods: **RD**, **NM**, **BV**, **BV-EOF**, **SV**, **LLV**, **RF**. Concerning the SV method, only the initial SVs (from now on simply the singular vectors, SV) are investigated to limit the number of methods to compare.

6.2 Lorentz-63 model

In fig. 6.1 the various perturbations are shown on top of the reference run as vectors. The norm of the vectors are scaled up for visualization. From this figure we can see many of the characteristics of the different methods.

The RD method results in perturbations uniformly distributed around the perturbed reference point. This means that the majority of perturbations point off the attractor and are not in-line with the underlying dynamics. Contrarily, the NMs lie completely in the plane of the attractor, as well as most RF perturbations and BVs. The distribution of RF perturbations are tilted slightly compared to the NMs. This is because they are constructed from reference states positioned all over the given attractor wing, which is not a perfect plane. The BVs concentrate in the direction parallel to the LLV^1. The LLV method gives by construction two perturbations (first and second) which lie in the attractor plane, while the third perturbation then is orthogonal to the attractor plane. This is also seen to some degree for the BV-EOF and SV method, but in general the first member perturbations of those methods point slightly off the plane. To see this, the supplementary figs. F.1 and F.2 show the perturbations along a trajectory in phase space for the three methods. Note the difference between the index one vectors: while the LLV^1s systematically lie approximately tangential to the trajectory, the $BV\text{-}EOF^1$s, and SV^1s lie more perpendicular to the trajectory.

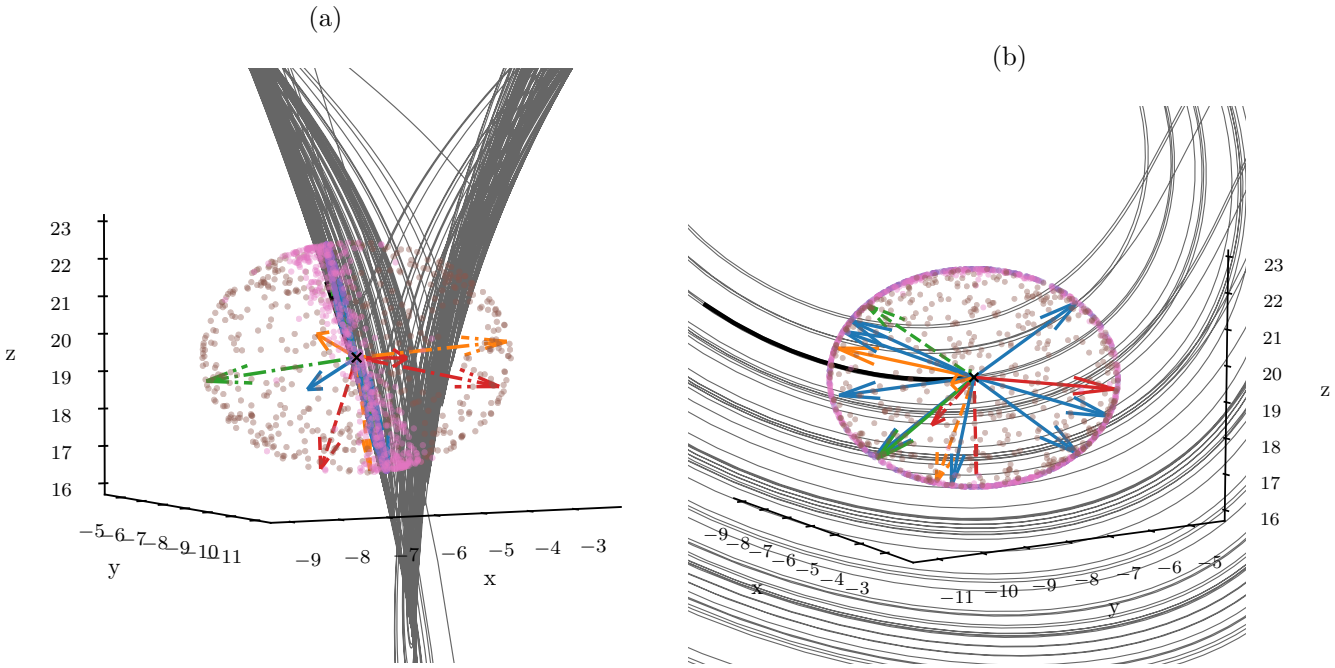


Figure 6.1: An example of the perturbations optimized for a specific point in time (black cross) shown as vectors from two different views; (a) in the plane of the left wing, (b) orthogonal to the plane of the left wing. The colouring follows the convention **RD**, **NM**, **BV**, **BV-EOF**, **SV**, **LLV**, **RF**. The first, second and third vector of the BV-EOF, LLV and SV method are plotted with solid, dashed and dash-dotted vectors, respectively. For the RD, NM and RF methods the vectors are shown as dots; i.e. the perturbation vectors start at the black cross and points towards the dots. The black curve highlights the optimization time of the SVs.

In fig. 6.2 the error growth relative to the reference run of ensemble forecasts started using the various perturbation methods are shown. The chosen region of the reference run shows a regime shift from right to left wing. Starting with fig. 6.2b we can see how the $\mathrm{RD}_{\mathrm{ens}}$ has a larger spread than the $\mathrm{NM}_{\mathrm{ens}}$ and $\mathrm{RF}_{\mathrm{ens}}$ and that the spread of the latter two mainly covers the top part of the $\mathrm{RD}_{\mathrm{ens}}$ spread. The third and fourth $\mathrm{NM}_{\mathrm{ens}}$ from the left shows no spread at all, which is because the reference trajectory is in the region where the eigenvalues are real (see fig. 3.2b and sec. 4.1). Note how the maximal error growth of the ensembles increases around the regime transition.

Comparing with fig. 6.2a the $\mathrm{BV\text{-}EOF}_{\mathrm{ens}}$, $\mathrm{SV}_{\mathrm{ens}}$ and $\mathrm{LLV}_{\mathrm{ens}}$ spread approximately as much as the $\mathrm{RD}_{\mathrm{ens}}$, but with only three members. On the contrary the $\mathrm{BV}_{\mathrm{ens}}$ most often "group together"; sometimes in the direction of large error growth (e.g. fifth ensemble from right), sometimes not (e.g.

fourth ensemble from right). The SV^1 generally finds the direction of maximum error growth, while the BV-EOF^1 less often do. In e.g. the first three ensembles the BV finds either a direction of suboptimal error growth (the first two) or results in a large spread of the ensemble (third), which makes the BV-EOF^1 fail to find the optimal direction. Comparing with the LLV_{ens}, the BV_{ens} generally evolve similarly to the LLV^1_{mbr}, while the BV-EOF^2_{mbr} and BV-EOF^3_{mbr} evolve similarly to the LLV^2_{mbr} and LLV^3_{mbr}, respectively.

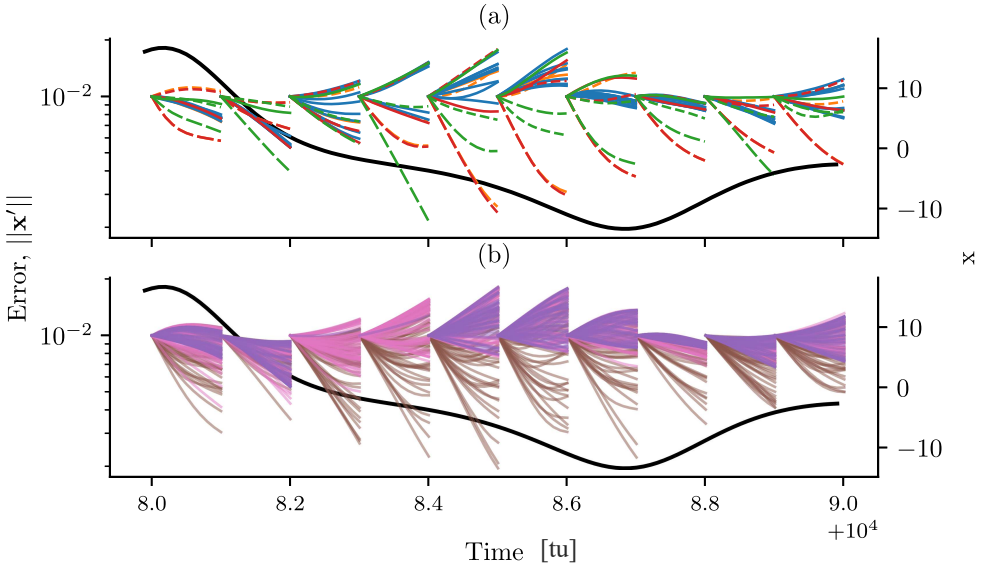


Figure 6.2: An example of ensemble forecasts perturbed by the various perturbation methods shown as the error growth relative to the reference run (left ordinate). The colouring follows the convention **RD**, **NM**, **BV**, **BV-EOF**, **SV**, **LLV**, **RF**. The first, second and third member of the BV-EOF, LLV and SV ensembles are plotted with solid, dashed and dash-dotted lines, respectively. The forecasts are started sequentially each 0.1tu. In the background the x component of the reference run is shown (black curve, right ordinate), which indicates that a regime shift occurs. For the RD, NM and RF ensembles they consist of 50 members, for the BV ensemble 8 members, and for the rest 3 members. Compare with fig. 1 in (L. Magnusson et al., 2008) included in fig. H.1.

To assess the performance of the various methods in terms of error growth, we investigate the exponential growth rate of the ensemble forecasts defined as

$$\kappa(t) = \frac{1}{dt}\ln\left(\frac{||\mathbf{x}'(t+dt)||}{||\mathbf{x}'(t)||}\right), \tag{6.1}$$

where $\mathbf{x}'(t)$ is the error of an ensemble forecast relative to the reference (L. Magnusson et al., 2008). More precisely we denote eq. (6.1) the *instantaneous*

exponential growth rate at time t when $dt \to 0$. The *mean* exponential growth rate $\kappa_{mean}(t_0)$ is given by eq. (6.1) with a finite $dt = \Delta t$ and constant $t = t_0$.

In fig. 6.3 the average instantaneous exponential growth rates are plotted for the various methods with dt taken as the time step of the model integration (sec. 5.4). The average is made from 5000 ensemble runs for each method, which are initiated sequentially with 1tu separation. Each ensemble consists of three members.

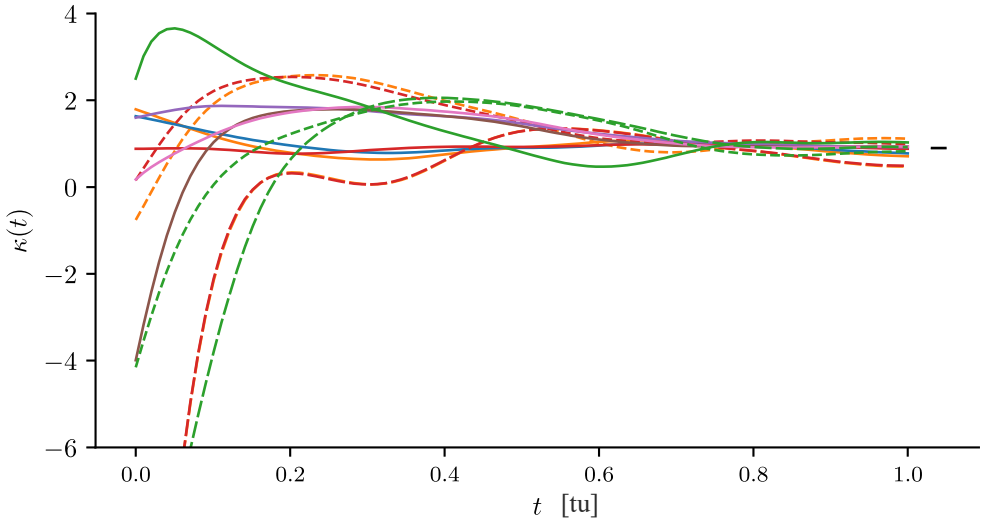


Figure 6.3: The instantaneous exponential growth rate, $\kappa(t)$, for the various methods. The colouring follows the convention **RD**, **NM**, **BV**, **BV-EOF**, **SV**, **LLV**, **RF**. The growth rates that correspond to the first, second and third member of the BV-EOF, LLV and SV ensembles are plotted with solid, dashed and dash-dotted lines, respectively. The growth rates are averaged over 5000 ensemble runs for each method, which are initiated sequentially with 1tu separation. Each ensemble consists of three members and all forecasts are run for 1tu. To the right the black horizontal line shows the dominant global Lyapunov exponent given by $\lambda_1 = 0.9\text{tu}^{-1}$. Compare the performance of the RD, NM, BV, BV-EOF and SV methods with fig. 3 in (L. Magnusson et al., 2008) and the RF method with fig. 5.2 in (L. Magnusson, 2009). The figures are included in appendix H.

During the first ~ 0.4tu, the methods differ significantly, while they all tend to converge to the dominant global Lyapunov exponent for $t \to 1$tu shown as the black line ($\lambda_1 = 0.9\text{tu}^{-1}$). Focusing on the first ~ 0.4tu, the RD_{ens}, $\text{BV-EOF}^3_{\text{mbr}}$, SV^2_{mbr}, SV^3_{mbr} and $\text{LLV}^3_{\text{mbr}}$ show the most negative growth rates, while the SV^1_{mbr} shows the most positive. The BV_{ens} and $\text{BV-EOF}^1_{\text{mbr}}$ perform very similarly as well as the $\text{LLV}^2_{\text{mbr}}$, $\text{BV-EOF}^2_{\text{mbr}}$ and

$\text{LLV}^3_{\text{mbr}}$, $\text{BV-EOF}^3_{\text{mbr}}$ respectively. Note that the $\text{LLV}^2_{\text{mbr}}$ and $\text{BV-EOF}^2_{\text{mbr}}$ actually shows larger growth rates than the $\text{LLV}^1_{\text{mbr}}$ and $\text{BV-EOF}^1_{\text{mbr}}$ respectively for $t > 0.05$tu. Compared with the RD_{ens}, the RF_{ens} shows positive growth rates at all times and converges to the RD_{ens} at $t \sim 0.1$tu. The NM_{ens} also converges to RD_{ens} but at $t \sim 0.2$tu and the growth rates start off larger than the BV_{ens} and $\text{BV-EOF}_{\text{ens}}$.

The distribution of $\kappa_{mean}(t_0)$ in phase space is shown in fig. 6.4 based on 1000 ensembles started sequentially in time with 1tu separation. Each ensemble consists of three members.

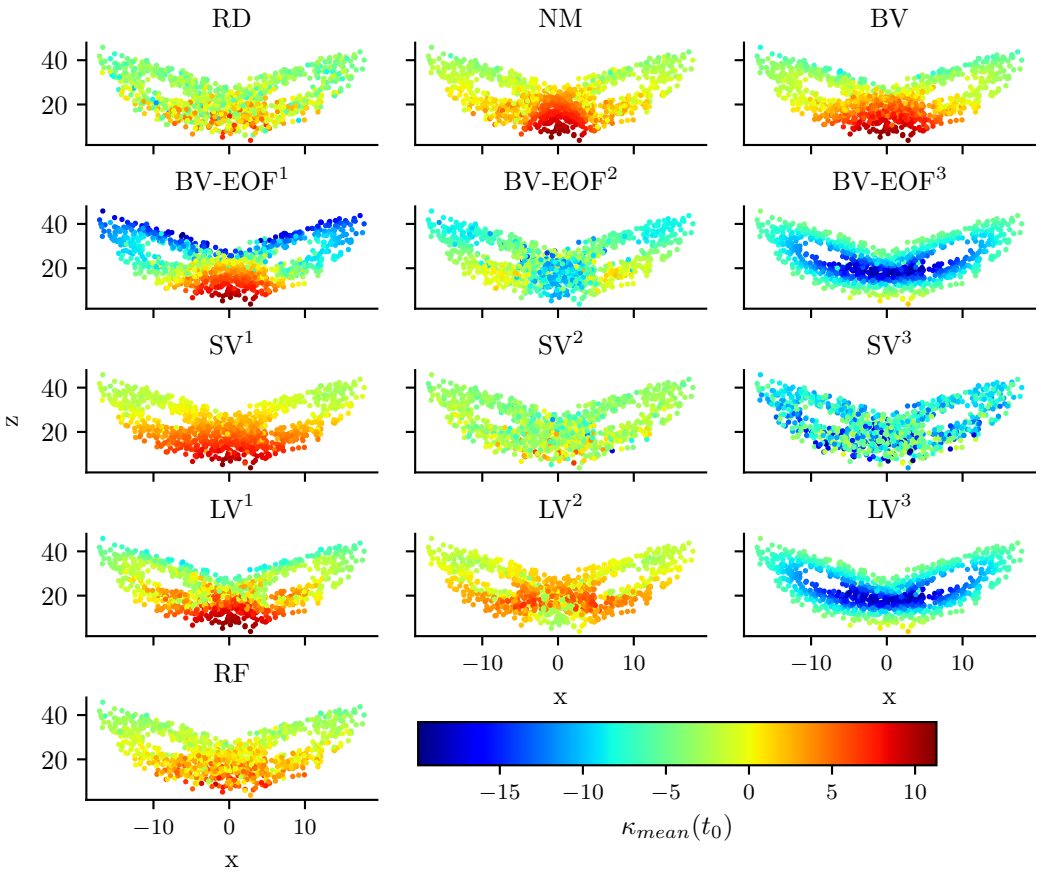


Figure 6.4: The phase space distribution of mean exponential growth rates, $\kappa_{mean}(t_0)$, of the various ensembles. Plotted in the z-x plane as dots positioned on the reference trajectory at $t = t_0$. The dots represent ensemble averages except for the $\text{BV-EOF}_{\text{ens}}$, SV_{ens} and LLV_{ens} which are split into separate member plots. For each method 1000 ensemble runs are made positioned sequentially in time with 1tu separation. Each ensemble consists of three members. The forecast time and $dt = \Delta t$ in eq. (6.1) is 0.1tu.

An immediate similarity between the $\mathrm{NM}_{\mathrm{ens}}$, $\mathrm{BV}_{\mathrm{ens}}$ and $\mathrm{LLV}^1_{\mathrm{mbr}}$ is observed, although the region with large positive growth rates nearby the $\mathbf{x}_0$ fixpoint seem more narrow for the $\mathrm{NM}_{\mathrm{ens}}$. The $\mathrm{SV}^1_{\mathrm{mbr}}$ also appear similar to those three ensembles, but with this region being more broad. The top of the wings show also larger growth rates than the before mentioned ensembles.

Comparing the $\mathrm{BV}_{\mathrm{ens}}$ with the $\mathrm{BV\text{-}EOF}_{\mathrm{ens}}$ we see how the $\kappa_{mean}(t_0)$ distribution of the former is split into three different distribution; the $\mathrm{BV\text{-}EOF}^1_{\mathrm{mbr}}$ covers the region nearby the $\mathbf{x}_0$ fixpoint, the $\mathrm{BV\text{-}EOF}^2_{\mathrm{mbr}}$ the regions around the $\mathbf{x}_\pm$ fixpoints, and the $\mathrm{BV\text{-}EOF}^3_{\mathrm{mbr}}$ the region near the edge of the attractor. Those distributions compare well with the ones for $\mathrm{LLV}_{\mathrm{ens}}$ especially $\mathrm{LLV}^3_{\mathrm{mbr}}$. The $\mathrm{LLV}^1_{\mathrm{mbr}}$ shows less negative growth rates around the top of the wings, while the $\mathrm{LLV}^2_{\mathrm{mbr}}$ shows larger growth rates in the central part of the attractor. Similar clear patterns are not seen in $\mathrm{SV}^2_{\mathrm{mbr}}$ and $\mathrm{SV}^3_{\mathrm{mbr}}$, which merely shows almost uniformly distributed growth rates, which resembles the $\mathrm{RD}_{\mathrm{ens}}$. Comparing the $\mathrm{RF}_{\mathrm{ens}}$ with $\mathrm{RD}_{\mathrm{ens}}$ we observe that the former results in larger growth rates nearby the $\mathbf{x}_0$ fixpoint than the latter, but to a lesser extent than any other method.

6.3 Sabra shell model

In fig. 6.5 the equivalent to fig. 6.2 is shown for the Sabra shell model. The ensembles are started sequentially separated by 0.01tu. The chosen section of the reference run shows both regions dominated by the smaller scales (e.g. fourth and sixth ensemble from the left) and by the larger scales (e.g. first three ensembles from the left). Despite this, the t_{OPT} valid for R_{small} is used for all ensembles. Notice the different behaviour of the error growth in the two different regions; the error growth in regions dominated by the small scales is much faster than in regions dominated by the large scales, but the influence from viscosity, observed as dampening of the error growth, is also much more apparent. The forecast length of 0.01tu is chosen to show the behaviour of the forecasts outside the linear range, where the influence from the non-linear dynamics is particularly clear in regions dominated by the small scales.

In fig. 6.5b we see that the members of the $\mathrm{NM}_{\mathrm{ens}}$ on average give slightly larger error growth than the $\mathrm{RD}_{\mathrm{ens}}$, since the members are confined to the part of the $\mathrm{RD}_{\mathrm{ens}}$ with the largest error growths (e.g. the first two ensembles from the left). In contrary, the $\mathrm{RF}_{\mathrm{ens}}$ shows a large spread of the members and especially a large portion of members with negative error growth (e.g. the first three ensembles from the left). As with the Lorentz-63 model we also see how the members of the $\mathrm{BV}_{\mathrm{ens}}$ evolve very alike (e.g. first, second

and sixth ensemble from the left), whereas the BV-EOF$_{\text{ens}}$, SV$_{\text{ens}}$ and LLV$_{\text{ens}}$ are capable of spreading more (fig. 6.5a).

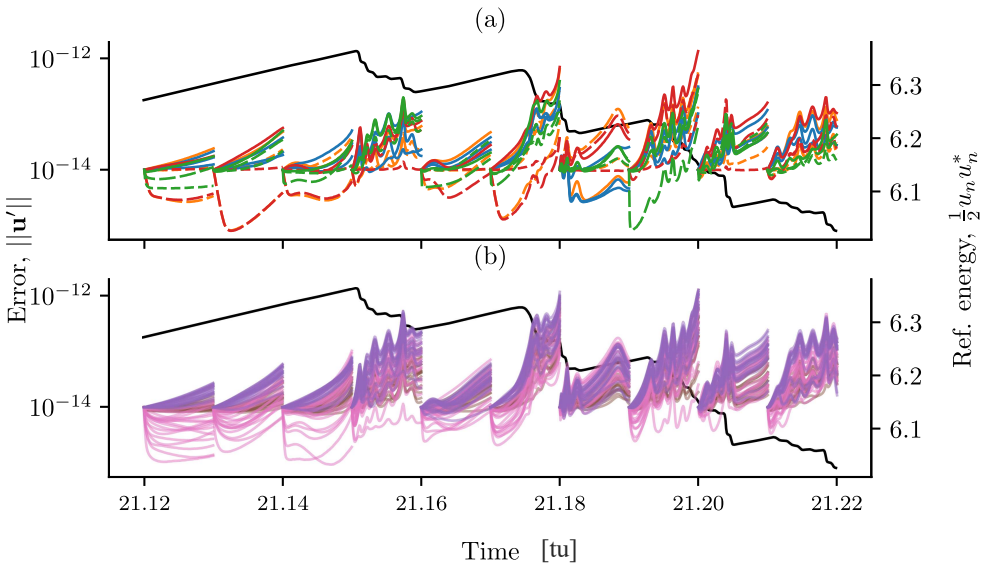


Figure 6.5: An example of ensemble forecasts perturbed by the various perturbation methods shown as the error growth relative to the reference run (left ordinate). The colouring follows the convention **RD**, **NM**, **BV**, **BV-EOF**, **SV**, **LLV**, **RF**. The forecasts are started sequentially each 0.01tu. In the background the total energy of the reference run is shown (black curve, right ordinate), which indicates periods dominated by the small scales (e.g. fourth and sixth ensemble from the left) and the large scales (e.g. first three ensembles from the left). Despite this, the t_{OPT} is set to the one valid for R_{small} for all ensembles. All ensembles consist of 20 members, but only 3 members are shown for BV$_{\text{ens}}$ (randomly chosen), BV-EOF$_{\text{ens}}$ (mbr. 1, 10, 18), LLV$_{\text{ens}}$ (mbr. 2, 10, 18) and SV$_{\text{ens}}$ (mbr. 1, 10, 18). For the latter three ensembles the error growth of the specific members is plotted as solid, dashed and dash-dotted lines, respectively.

In fig. 6.6, the average instantaneous exponential growth rates given by eq. (6.1) are plotted. The average is made from 1000 ensemble runs for each method in (a) R_{small} and (b) R_{large}. Since some methods perform very alike, supplementary plots of the results for the individual methods are given in figs. G.1 to G.3 and the results for all members of the SV$_{\text{ens}}$ is given in fig. G.4.

Starting with fig. 6.6a we see that the SV, RD and RF methods, together with BV-EOF$^{18}_{\text{mbr}}$ and LLV$^{18}_{\text{mbr}}$, start off with negative growth rates. After the first 0.002tu a clear convergence of the methods is seen. Since the growth rates are evaluated in distinct regions, the convergence cannot be directly compared to the dominant global Lyapunov exponent as with the Lorentz-63

model. Instead, the convergence is compared to the dominant LLE ($\pi_2 = 206\text{tu}^{-1}$ and $\pi_2 = 64\text{tu}^{-1}$ for R_{small} and R_{large}, respectively)[1] shown as the black line to the right in the plot.

The $\text{BV-EOF}^1_{\text{mbr}}$ shows the largest growth rates followed by the BV_{ens} and $\text{LLV}^2_{\text{mbr}}$ in the first $5 \cdot 10^{-4}$tu. The latter two show almost identical growth rates for all t, and also the members of the $\text{BV-EOF}_{\text{ens}}$ and LLV_{ens} show very similar behaviour. The BV_{ens}, $\text{BV-EOF}^1_{\text{mbr}}$ and $\text{LLV}^2_{\text{mbr}}$ show a characteristic drop to negative growth rates in $[5 \cdot 10^{-4}; 1.5 \cdot 10^{-3}]$tu, which is due to the viscosity that dampens the error growth as noted in relation to fig. 6.5. The same is seen for all other ensembles, but less pronounced. In

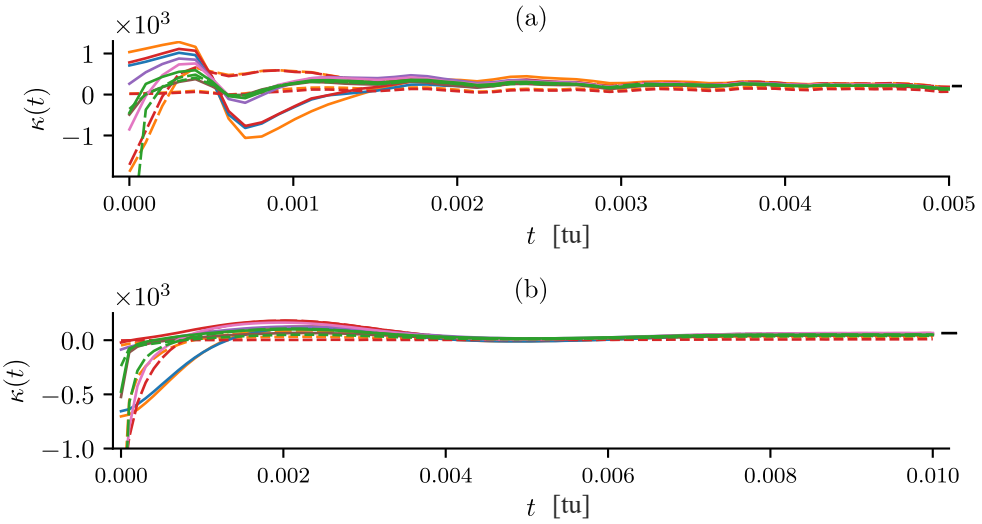


Figure 6.6: The instantaneous exponential growth rate, $\kappa(t)$, for the various methods. The growth rates are averaged over 1000 ensembles for each method, which are initiated in (a) R_{small} and (b) R_{large}. All ensembles consist of 20 members, but for the $\text{BV-EOF}_{\text{ens}}$, LLV_{ens} and SV_{ens} the growth rates are shown for the individual members [1, 10, 18], [2, 10, 18] and [1, 10, 18], respectively. To distinguish the growth rates of the individual members they are plotted with solid, dashed and dash-dotted lines, respectively. For comparison, the second LLE is plotted to the right (black line; $\pi_2 = 206\text{tu}^{-1}$ and $\pi_2 = 64\text{tu}^{-1}$ for R_{small} and R_{large}, respectively) averaged over 1000 LLV_{ens}. All forecasts are run for 0.01tu, which is also the Δt used in the LLE calculation (eq. (2.14)). Notice the different time range on the abscissas.

the beginning $5 \cdot 10^{-4}$tu we also see that the NM_{ens} shows larger growth rates than the SV_{ens} and the RF_{ens}, although the RF_{ens} reaches the same growth

[1] As described later, the first LLV shows very different behaviour than the rest of the LLVs and is treated specially. For that reason, the second LLE is used as the dominant LLE.

rates in the end of this interval.

In fig. 6.6b we see that the methods show both similar and different behaviour compared to fig. 6.6a. The time before the methods converge is longer, ~ 0.006tu. A characteristic drop in growth rates for all methods is seen again, but this time around ~ 0.005tu and much less pronounced. In the beginning ~ 0.001tu all methods show to some degree negative growth rates; a clear difference to fig. 6.6a being the very negative growth rates of the BV method and BV-EOF$^1_{\text{mbr}}$. The maximum growth rate is seen in LLV$^2_{\text{mbr}}$ closely followed by the RF method and the NM method.

In fig. 6.7 we present the average spectral distribution of the perturbation vectors for the BV, BV-EOF, SV, LLV and RF methods calculated in R_{small}. Similar figures for R_{large} is given in appendix G.4. All vectors are normalized to have norm one before averaging.

As seen in fig. 6.7d, both the BV and RF vectors mainly occupy the shells in the range $n \in [14; 18]$. While the BV vectors in the lower part of the spectrum show equal component amplitudes of ~ 0.1, the RF vector components decrease in amplitude for $n \to 1$. The orthogonalisation of the breed vectors through EOF analysis gives the BV-EOF distribution in fig. 6.7c (bottom panel). In the top panel the variance explained by the ith BV-EOF normalized to the total variance is shown; the plot is truncated at 10^{-4} for better visualization of the largest variances. The orthogonalisation clearly pulls out information from the linearly dependent BVs and gives a spectral structure for $i < 5$ and $i > 15$ with resemblance to the LLV spectral distribution in fig. 6.7b (bottom panel). For $i \in [5; 15]$ on the contrary, the spectrum shows little to no structure.

The LLV distribution shows a clear resemblance to the eigenvector distribution in fig. 3.5 with the characteristic V-shape. A distinct difference is seen for $i = 1$ for which reason the average LLV$^1_{\text{mbr}}$ is plotted in fig. 6.7d (red curve). For comparison the Kolmogorov scaling of the velocity is plotted (eq. (2.24), black dashed line). The Lyapunov exponents are shown in fig. 6.7b (top panel) except λ_1, which is a factor 10 larger than the next largest exponent and would dominate the plot. Furthermore, the plot is truncated at -200tu^{-1} for similar reasons concerning the most negative exponents.

The distribution of SVs is shown in fig. 6.7a (bottom panel) together with the exponential growth rates calculated from the singular values as $\frac{1}{t_{OPT}}\ln(\sigma_i)$ (top panel). Despite being relatively blurred out, the distribution show some spectral structure with the first five singular vectors concentrated around $n \in [15; 18]$. The first eight singular values give positive growth rates, while the rest give negative. For $i > 15$ some singular values turned out to be negative contrary to what is expected from theory (sec. 4.2), for which reason the growth rates are undefined and not shown.

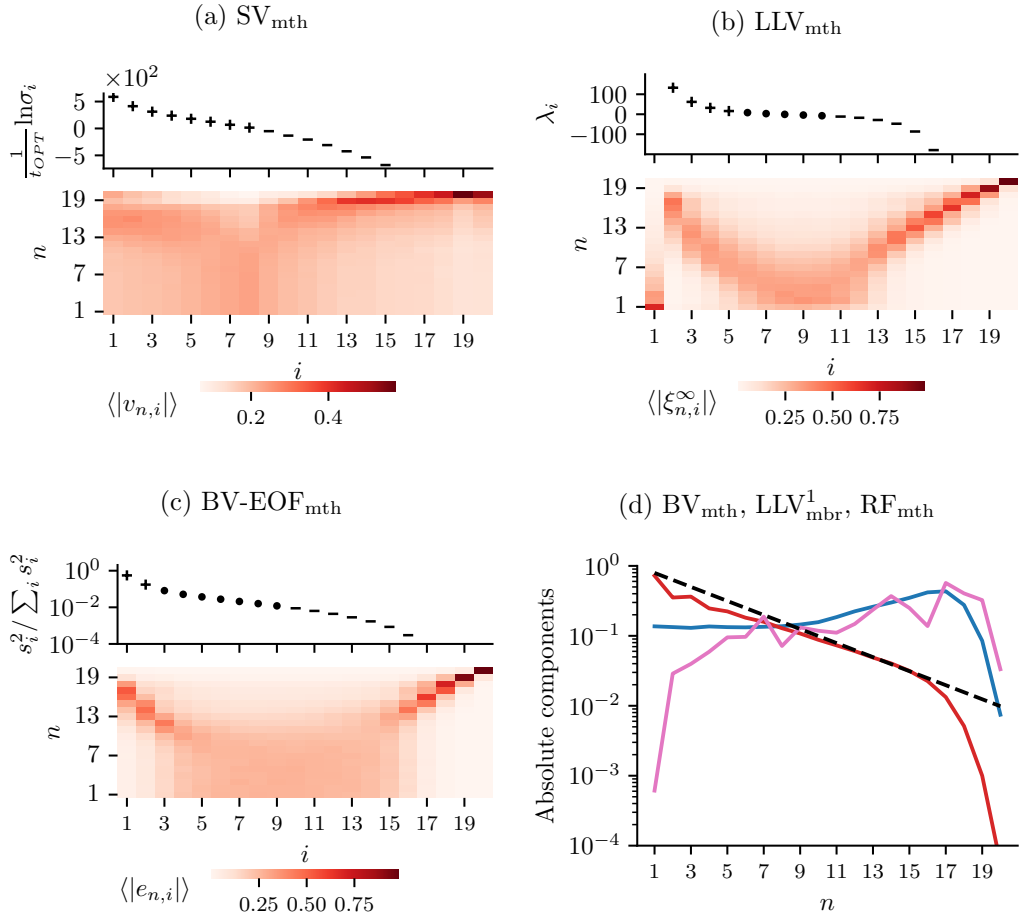


Figure 6.7: (a-c) The average spectral distribution of perturbations for three different methods (bottom panels) together with the mean exponential growth rates calculated from the singular values, the global Lyapunov exponents and the BV-EOF variances normalized to the total variance, respectively (top panels). The average is made over 1000 ensembles initiated in R_{small}. (d) The mean spectrum of the RF and BV perturbations together with the mean LLV$^1_{\mathrm{mbr}}$. The Kolmogorov scaling of the velocity eq. (2.24) is shown for comparison (black dashed). In (a; top) the sign of the growth rates is shown with $+$, $-$. The singular values for $i > 15$ turned out negative and the growth rates are not shown. In (b; top) the sign of λ_i is shown with $+$, $\cdot$ and $-$, where $\cdot$ means approximately zero. The exponent corresponding to LLV^1 is not shown since it is a factor 10 larger than the next largest exponent and would dominate the plot. Furthermore, the plot is truncated at $-200\mathrm{tu}^{-1}$ for similar reasons concerning the most negative exponents. In (c; top) the size of the relative variance is shown with $+$, $\cdot$ and $-$ signifying above 0.1, between 0.01 and 0.1, and below 0.01. The plot is truncated at 10^{-4} for better visualization of the largest variances.

7 Discussion

In this thesis we have studied various methods to produce perturbations of the initial conditions of an ensemble system using two low dimensional non-linear models. This section is devoted to a discussion of the results.

7.1 Lorentz-63 model

Concerning the results of the RD, NM, BV, BV-EOF and SV methods for the Lorentz-63 model we generally see a good agreement of figs. 6.2 and 6.3 with figs. 1 and 3 in (L. Magnusson et al., 2008) (shown in figs. H.1 and H.2), with some deviations. For an example, L. Magnusson et al. found a bit larger values for the maximal instant exponential growth rate of SV^1_{mbr} and $\text{BV-EOF}^1_{\text{mbr}}$, and a bit smaller values for the initial growth rates of the BV_{ens}. The $\text{BV-EOF}^1_{\text{mbr}}$ and BV_{ens} are thus reported by L. Magnusson et al. to perform differently contrary to what is observed in fig. 6.3. This discrepancy may be due to the number of breeding cycles and breed vectors used[1], which affects the co-variance in the EOF analysis and the resulting BV-EOFs.

The superiority of the $\text{BV-EOF}^2_{\text{mbr}}$ to the $\text{BV-EOF}^1_{\text{mbr}}$ after $\sim$ 0.05tu (also observed by L. Magnusson et al.) is surprising, since the first BV-EOF reflects the largest portion of the variance of the BVs, which one could expect to contain the direction of the largest error growth rate. L. Magnusson et al. proposes that it can be a result of the underlying eigenvector structure being non-orthogonal, as investigated by Smith et al., 1999, which is the case for the Lorentz-63 model since the Jacobian is asymmetric (eq. (3.6)). The average angles between the eigenvectors evaluated at 5000 randomly chosen points on the attractor are 68.3 deg ($\arccos(\langle \boldsymbol{\xi}_1; \boldsymbol{\xi}_2 \rangle)$), 78.5 deg ($\arccos(\langle \boldsymbol{\xi}_1; \boldsymbol{\xi}_3 \rangle)$), 79.6 deg ($\arccos(\langle \boldsymbol{\xi}_2; \boldsymbol{\xi}_3 \rangle)$).

Smith et al. calculates the instantaneous exponential growth rates in a

[1] The values are not stated in (L. Magnusson et al., 2008) and could for that reason not be reproduced.

two-dimensional *linear* system with a *constant* Jacobian constructed to have eigenvectors separated by an angle δ. By varying δ Smith et al. observes how the growth rates of an initial error exceeds the maximum eigenvalue in the transient period for $\delta \neq \pm 90\,\text{deg}$; the less orthogonal, the larger growth rates (refer to fig. 3 in (Smith et al., 1999)). The direction that corresponds to the largest growth rate is off by $-1/2\text{arctan}(\text{cotan}(\delta))$ relative to the first eigenvector (Smith et al., 1999), which corresponds to $-10.9\,\text{deg}$ for the Lorentz-63 model. This reasoning thus suggests, that the BV-EOF^2 should project on the first eigenvector to a high degree ($\cos(-10.9\,\text{deg}) \approx 0.98$). The average projection is however 0.5 ($\langle \text{BV-EOF}^2; \boldsymbol{\xi}_1 \rangle$ averaged over 5000 ensembles) and not even significantly different from the same calculation for $\langle \text{BV-EOF}^1; \boldsymbol{\xi}_1 \rangle$.

The application of this reasoning to the non-linear Lorentz-63 model with non-constant Jacobian can be questioned, but may have some relevance for the understanding of the observed behaviour of the BV-EOF$_{\text{ens}}$, as long as the breed vectors are optimized for a time within the validity of the linear approximation. The calculated average projections, however, do not seem to support it.

Concerning the RF method, we can compare with (L. Magnusson, 2009; included in fig. H.3) which reveals a large disagreement in the beginning 0.2tu. While L. Magnusson find growth rates starting out at $\sim 2\text{tu}^{-1}$ and increasing to a maximum of $\sim 3\text{tu}^{-1}$, we find growth rates starting slightly above zero and peaking around $\sim 2\text{tu}^{-1}$.

As described in sec. 4.5, L. Magnusson defines as a constraint on the sampling of historical states, that the states shall belong to the same wing as the initial point of an ensemble. No further information is given on how the wings are separated. The crude split at $x = 0$ used in this thesis may be different from that used by L. Magnusson. No matter what, it is certainly not the most accurate split, since as seen in fig. 4.3 portions of the wings are faulty classified near the transition between the wings. Even with more elaborate splitting techniques, however, the classification will still have some error since the transition occurs gradually. The error made will produce some RF perturbations that are not in line with the underlying dynamics and thus contribute with negative growth rates to the mean.

Another difference to (L. Magnusson, 2009) is that we use a minimum separation time between the randomly chosen historical states, while L. Magnusson does not. Since, however, the chance of choosing correlated states in the 10^4tu long period used in the RF method (sec. 5.5.1) is very small, the effect of not having this minimum separation time is negligible.

7.2 Sabra shell model

The results for the Sabra shell model raises several questions to be discussed. Comparing the performance of the SV method with all other methods, it is surprising that both the BV, BV-EOF, LLV, NM, and to some degree the RF method produce larger exponential growth rates as opposed to what is expected from theory (sec. 4.6) and what is seen with the Lorentz-63 model. The reason may be related to the performance of the Lanczos algorithm, which, as noted in sec. 5.3, suffers from being sensitive to numerical errors. Those errors may develop more/faster in the Sabra shell model than in the Lorentz-63 model due to the larger number of (especially short) timescales and lead to suboptimal SV estimation.

Focusing on the SV method, it is also surprising that the SV that gives rise to the maximal error growth rate within the optimization interval is not SV^1. From fig. G.4 it is seen that this is obtained for $\mathrm{SV}^2_{\mathrm{mbr}}$, $\mathrm{SV}^4_{\mathrm{mbr}}$ in R_{small}, R_{large}, respectively. The difference is, however, small and may be an effect of the influence of the non-linear dynamics on the SV perturbations, which are optimized with the linear models.

Another reason may be related to the orthogonality of the SVs. From theory, it is expected that the SVs are completely orthogonal (sec. 4.2), but in practice the average $\langle \mathbf{v}_i; \mathbf{v}_j \rangle$ for $j \neq i$ is ~ 0.26. In that way, the direction of the largest error growth depicted by SV^1 is partly contained in other SVs, which then may produce large growth rates.

The non-orthogonality of the SVs could be a result of the sensitivity to numerical errors of the Lanczos algorithm as noted above. Although the re-orthogonalisation of the Lanczos vectors solved the problem of loss of orthogonality between those vectors, it cannot be ruled out that such numerical errors have had an impact on the orthogonality of the SVs.

This moderate non-orthogonality is also reflected by the distribution shown in fig. 6.7a. Since the SVs converge towards the adjoint LLVs as $t_{OPT} \to \infty$ (sec. 4.6.2), which shows a spectrum similar to the LLVs (fig. G.5), we expect to see a somewhat similar spectrum for the SVs. The finite value for t_{OPT} of course has a large impact on the observed discrepancies, but we also believe that the non-orthogonality of the SVs plays a role. The impact from the t_{OPT} was investigated by running the SV method for $t_{OPT} = 0.001$tu and $t_{OPT} = 0.005$tu too. The corresponding distributions are shown in appendices G.5 and G.6. Compared to fig. 6.7a we see that the range of vectors, that concentrate around $n \in [15; 18]$ and correspond to positive growth rates, is extended from $i \in [1; 5]$ to $i \in [1; 7]$ when t_{OPT} is increased by a factor 10; i.e. more vectors are concentrated in the same shell range. From this

investigation, the similarity to the distribution of the adjoint LLV is not seen to improve.

In regard to the distribution of the BV-EOFs and the corresponding variances, the clear spectral structure for $i < 5$ is expected because of the unstable dynamics around $n \in [15; 18]$; this results in a large variance of the BVs, which is captured in the EOF analysis. More surprising is the clear structure for $i > 15$. Since this range is dominated by contraction of the phase space due to diffusion, the variance of the BV components is very low. One could expect this to make the EOF algorithm struggle to pull out enough information to recover the spectral structure. However, the distinct orthogonal structure of the phase space in this range, as observed in fig. 3.5, may be the explanation; it simply forces the BVs and consequently the BV-EOFs to reflect this structure.

Another way to understand these observations, is through the theoretical knowledge that the BVs converge towards the leading LLV for $t_{opt} \to \infty$ as described in sec. 4.6.3. Since a Gram-Schmidt orthogonalisation procedure is used both to obtain multiple LLVs and to orthogonalize the BVs into BV-EOFs, it seems natural that the distribution of BV-EOFs looks like the distribution of LLV (fig. 6.7b). The blurry spectrum for $5 < i < 15$ is an exception to this. Since this range corresponds to the shells $n \in [1; 9]$ with eddy turnover times between 2.07tu and 0.1tu, the optimization time of $5 \cdot 10^{-4}$tu is too short to properly breed on the dynamics of those shells. This is also reflected in the flat spectrum of the average BVs seen in fig. 6.7d.

The distribution of LLVs resembles very much the eigenvector spectrum in fig. 3.5 except for LLV^1. Although the LLVs depend on the Jacobian evaluated along a trajectory and the eigenvectors only depend on the Jacobian evaluated at a specific point in time, the LLVs reduces to the eigenvectors if the flow is stationary as noted in sec. 4.6.3. The averaging of the eigenvectors and LLVs imitates this, since the local characteristics of the eigenvectors valid at different times, and likewise for Lyapunov vectors, are averaged out. In that way, the very alike distributions of LLVs and eigenvectors is to be expected. It is noted that the LLV spectrum except LLV^1 corresponds well with the results of Yamada et al., 1988, which studied the LLV spectrum of the very similar GOY shell model.

Regarding the RF method, the highly negative growth rates during the transient period in fig. 6.6, and the fact that the RD_{mth} initially performs better, indicates that the configuration and/or implementation of the method is not optimal.

One explanation lies in the normalization of eq. (4.18) with the Kolmogorov spectrum of the shell velocities. Despite that it leads to perturbations dominated by the small scales, the spectrum of the mean perturbation

appears relative "bumpy", and continues farther into the dissipative range than e.g. the spectrum of the BVs (see fig. 6.7d). Better ways to manipulate eq. (4.18) to give large growth rates may exist.

Another explanation lies in the constraints on how to choose the state vectors. This is done on the basis of the region analysis presented in appendix D. The random choice of a state vector within a given region, however, is not ensured to be a state vector that reflects the overall trend of the region (i.e. if the small or large scales dominate). This is due the size of the kernel, which makes the regions a coarse-grained characteristic of the actual dynamics. In that way, the choice of state vectors may produce perturbation vectors that do not reflect the dynamics of the perturbation point and consequently result in suboptimal error growth of the ensemble.

8 Conclusion and Outlook

In this final chapter, we present a short sum up of the thesis together with the main conclusions. As a final remark we describe the next steps to be made and possible future work in an outlook section.

8.1 Conclusion

Predicting the weather has been and is of great importance, both in an academic and a practical context, and is a fundamentally difficult task due to the chaotic nature of the atmosphere. Focusing on ensemble weather forecasts, this work set out to investigate and review the crucial choice of method used to perturb the initial conditions of a forecast model. Seven different perturbation methods were studied with two low dimensional non-linear and chaotic models; the Lorentz-63 model and the Sabra shell model.

Through a theoretical outline of the methods, we saw how the normal mode, breeding vector and Lyapunov vector methods were related to the singular vector method. This was confirmed by numerical investigations.

The perturbation methods were investigated to obtain knowledge of the spatial/spectral distribution of the resulting perturbations, and the ability of the methods to produce large error growth relative to a reference.

In the investigations with the Lorentz-63 model, we first saw how the perturbations were distributed in phase space. While the RD method produced perturbations uniformly distributed around an initial point, all other methods produced perturbations that were more in-line with the underlying dynamics. The first SV of the SV method produced perturbations that led to the largest error growth compared to all other methods. This was expected from theory, and it also confirmed the findings of L. Magnusson et al., 2008. In general, the methods performed very alike to what was reported in this article, despite some deviations in the BV and BV-EOF methods, which we attributed to a difference in method parameters. The RF method was compared to (Magnusson et al., 2009) and did not produce as large exponential

growth rates as reported by Magnusson et al. The discrepancy was ascribed to a possible difference in the way the attractor of the Lorentz-63 model was split in order to sample historical states of the same wing as the initial point.

Concerning the investigations with the Sabra shell model, we evaluated the perturbation methods separately in two different regions; one region dominated by the large scales and one dominated by the small scales. This was done because the model showed very different behaviour in those regions which led to significantly different behaviour of the perturbation methods.

Surprisingly, the SV method did not produce the largest error growth rates; this was instead observed in the first (second) member of the BV-EOF (LLV) method in R_{small} (R_{large}). Also, the SVs were not observed to be completely orthogonal as opposed to theory. This unexpected behaviour was attributed to the Lanczos algorithm, which was used to find the singular vectors. The algorithm is known to be sensitive to numerical errors the effect of which was sought diminished by re-orthogonalizing the Lanczos vectors. The number of (especially short) timescales of the Sabra shell model, however, probably led to fast growth of the numerical errors and thus suboptimal SV estimation.

Regarding the RF method, we observed very suboptimal error growth rates especially in R_{large}. Even the RD method produced larger growth rates in the transient period. This behaviour was attributed partly to the normalization of the RF vectors relative to the Kolmogorov spectrum of shell velocities and partly to the constraints put on the choice of historical samples.

8.2 Outlook

In relation to the conclusions about the SV method and the Lanczos algorithm, an apparent next step would be to investigate ways to further counteract its sensitivity to numerical errors and search for alternative algorithms.

Another path to follow would be to investigate alternative perturbation methods to the ones presented in this thesis. We have mainly focused on the more traditional methods, while research into e.g. non-linear LLVs (Feng et al., 2014) and non-linear SVs (Winkler et al., 2020) shows interesting generalizations of the linear methods that could be worth examining.

Although suitable for reviewing fundamental concepts of the perturbation methods, the Lorentz-63 model and Sabra shell model are after all low-dimensional and more of theoretical relevance. Therefore, another very interesting research topic would be to apply the analyses to a real NWP model.

This has been done for global models in e.g. (L. Magnusson, 2009; Molteni et al., 1996; Toth et al., 1993), but similar studies for regional models do, to the author's impression, not exist to the same extend. At least conversations with Henrik Feddersen, PhD at DMI have revealed that some need and interest exists for investigating the perturbation methods in the context of the COMEPS model.

Bibliography

About our forecasts (2022). URL: https://www.ecmwf.int/en/forecasts/documentation-and-support (visited on 05/05/2022).

Bjerknes, Vilhelm (Dec. 1904 [2009]). "Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik [The problem of weather prediction, considered from the viewpoints of mechanics and physics]". Trans. from the German by E. Volken and S. Brönnimann. In: *Meteorologische Zeitschrift* 21, [18.6], 1–7, [663–667]. DOI: 10.1127/0941-2948/2009/416. URL: http://dx.doi.org/10.1127/0941-2948/2009/416.

Buizza, R. and T. N. Palmer (1995). "The Singular-Vector Structure of the Atmospheric Global Circulation". In: *Journal of Atmospheric Sciences* 52.9, pp. 1434–1456. DOI: 10.1175/1520-0469(1995)052<1434:TSVSOT>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/52/9/1520-0469_1995_052_1434_tsvsot_2_0_co_2.xml.

Buizza, Roberto et al. (2010). "Combined use of EDA- and SV-based perturbations in the EPS". In: (123), pp. 22–28. DOI: 10.21957/codf7bpp. URL: https://www.ecmwf.int/node/17467.

Charney, J. G. (1949). "On A Physical Basis For Numerical Prediction Of Large-Scale Motions In The Atmosphere". In: *Journal of Atmospheric Sciences* 6.6, pp. 372–385. DOI: 10.1175/1520-0469(1949)006<0372:OAPBFN>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/6/6/1520-0469_1949_006_0372_oapbfn_2_0_co_2.xml.

Charney, J. G., R. Fjörtoft, and J. Von Neumann (1950). "Numerical Integration of the Barotropic Vorticity Equation". In: *Tellus* 2.4, pp. 237–254. DOI: https://doi.org/10.1111/j.2153-3490.1950.tb00336.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2153-3490.1950.tb00336.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1950.tb00336.x.

Christiansen, Bo (Jan. 2021). "The blessing of dimensionality for the analysis of climate data". In: *Nonlinear Processes in Geophysics* 28, pp. 409–422. DOI: 10.5194/npg-2021-2.

Demmel, J. W. (1997). "Iterative Methods for Eigenvalue Problems". In: *Applied Numerical Linear Algebra*. Chap. 7, pp. 361–387. DOI: 10.1137/1.9781611971446.ch7. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611971446.ch7. URL: https://epubs.siam.org/doi/abs/10.1137/1.9781611971446.ch7.

Diaconescu, Emilia Paula and René Laprise (2012). "Singular vectors in atmospheric sciences: A review". eng. In: *Earth-science reviews* 113.3-4, pp. 161–175. ISSN: 0012-8252.

Ditlevsen (2000). "Symmetries, invariants, and cascades in a shell model of turbulence". eng. In: *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* 62.1 Pt A, pp. 484–489. ISSN: 1063-651X.

— (2011). *Turbulence and shell models*. eng. Cambridge: Cambridge University Press. ISBN: 0521190363.

Eliassen, A. (1949). *The Quasi-static Equations of Motion with Pressure as Independent Variable*. Geofysiske Publikasjoner. Grøndahl.

Epstein, E. S. (1969). "Stochastic dynamic prediction1". In: *Tellus* 21.6, pp. 739–759. DOI: https://doi.org/10.1111/j.2153-3490.1969.tb00483.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2153-3490.1969.tb00483.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1969.tb00483.x.

Feng, Jie et al. (2014). "The Application of Nonlinear Local Lyapunov Vectors to Ensemble Predictions in Lorenz Systems". eng. In: *Journal of the atmospheric sciences* 71.9, pp. 3554–3567. ISSN: 0022-4928.

Fortin, V. et al. (2014). "Why Should Ensemble Spread Match the RMSE of the Ensemble Mean?" eng. In: *Journal of hydrometeorology* 15.4, pp. 1708–1713. ISSN: 1525-755X.

Frisch, Uriel (1995). *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press. DOI: 10.1017/CBO9781139170666.

Frogner, Inger-Lise et al. (2019). "HarmonEPS—The HARMONIE Ensemble Prediction System". eng. In: *Weather and forecasting* 34.6, pp. 1909–1937. ISSN: 0882-8156.

Geostrophic adjustment (Jan. 2012). URL: https://glossary.ametsoc.org/wiki/Geostrophic_adjustment (visited on 05/12/2022).

Ghojogh, Benyamin, Fakhri Karray, and Mark Crowley (2019). *Eigenvalue and Generalized Eigenvalue Problems: Tutorial*. DOI: 10.48550/ARXIV.1903.11240. URL: https://arxiv.org/abs/1903.11240 (visited on 05/12/2022).

Holton, James R. and Gregory J. Hakim (2013). "Atmospheric Oscillations: Linear Perturbation Theory". In: *An Introduction to Dynamic Meteorology (Fifth Edition)*. Ed. by James R. Holton and Gregory J. Hakim. Fifth Edition. Boston: Academic Press. Chap. 5, pp. 127–170. ISBN: 978-0-12-384866-6. DOI: https://doi.org/10.1016/B978-0-12-384866-6.00005-2. URL: https://www.sciencedirect.com/science/article/pii/B9780123848666000052.

Isaksen, L. et al. (2010). "The new Ensemble of Data Assimilations". In: (123), pp. 17–21. DOI: 10.21957/vzbtkbf8. URL: https://www.ecmwf.int/node/17466.

Kalnay, E. (2002). "Atmospheric predictability and ensemble forecasting". In: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, pp. 205–260, 264–275. DOI: 10.1017/CBO9780511802270.007.

Kolmogorov, A. N. (1941). "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers". In: *Dokl. Akad. Nauk SSSR*.

L'vov, Victor S et al. (1998). "Improved shell model of turbulence". eng. In: *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* 58.2, pp. 1811–1822. ISSN: 1063-651X.

Lautrup, B. (Benny). (2011). "Continuous matter". eng. In: *Physics of continuous matter : exotic and everyday phenomena in the macroscopic world*. 2nd ed. Boca Raton, Fla: CRC Press, pp. 1–15. ISBN: 9781420077001.

Legras, Bernard and R. Vautard (Jan. 1995). "A guide to Lyapunov vectors". In: *Seminar on Predictability, Reading* 1, pp. 143–156.

Liberto, Tom Di (n.d.). *Hovmöller diagram: A climate scientist's best friend*. URL: https://www.climate.gov/news-features/understanding-climate/hovm%C3%B6ller-diagram-climate-scientist%E2%80%99s-best-friend.

Lorenz, E. N. (1963). "Deterministic Nonperiodic Flow". In: *Journal of Atmospheric Sciences* 20.2, pp. 130–141. DOI: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml.

— (1965). "A study of the predictability of a 28-variable atmospheric model". eng. In: *Tellus* 17.3, pp. 321–333. ISSN: 0040-2826.

Magnusson, L, J Nycander, and E Källén (2009). "Flow-dependent versus flow-independent initial perturbations for ensemble prediction". eng. In: *Tellus. Series A, Dynamic meteorology and oceanography* 61.2, pp. 194–209. ISSN: 0280-6495.

Magnusson, L. (2009). "Sampling uncertainties in ensemble weather forecasting". PhD thesis. Stockholm University, Department of Meteorology, p. 52. ISBN: 978-91-7155-865-7.

Magnusson, L., E. Källén, and J. Nycander (2008). "Initial state perturbations in ensemble forecasting". In: *Nonlinear Processes in Geophysics* 15.5, pp. 751–759. DOI: 10.5194/npg-15-751-2008. URL: https://npg.copernicus.org/articles/15/751/2008/.

Meurant, Gérard and Zdeněk Strakoš (2006). "The Lanczos and conjugate gradient algorithms in finite precision arithmetic". In: *Acta Numerica* 15, pp. 471–542. DOI: 10.1017/S096249290626001X.

Molteni, F et al. (1996). "The ECMWF Ensemble Prediction System: Methodology and validation". eng. In: *Quarterly journal of the Royal Meteorological Society* 122.529, pp. 73–119. ISSN: 0035-9009.

Molteni, Franco and Roberto Buizza (1999). "Validation of the ECMWF Ensemble Prediction System Using Empirical Orthogonal Functions". In: *Monthly Weather Review* 127.10, pp. 2346–2358. DOI: 10.1175/1520-0493(1999)127<2346:VOTEEP>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/127/10/1520-0493_1999_127_2346_voteep_2.0.co_2.xml.

Navon, I. M. et al. (1992). "Variational Data Assimilation with an Adiabatic Version of the NMC Spectral Model". In: *Monthly Weather Review* 120.7, pp. 1433–1446. DOI: 10.1175/1520-0493(1992)120<1433:VDAWAA>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/120/7/1520-0493_1992_120_1433_vdawaa_2_0_co_2.xml.

Palmer, T. N (1993). "Extended-Range Atmospheric Prediction and the Lorenz Model". eng. In: *Bulletin of the American Meteorological Society* 74.1, pp. 49–65. ISSN: 0003-0007.

Richardson, L. F. (1922). "Weather prediction by numerical process". In: *Quarterly Journal of the Royal Meteorological Society* 48.203, pp. 282–284. DOI: https://doi.org/10.1002/qj.49704820311. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49704820311. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49704820311.

Sandu, Adrian and Philipp Miehe (2010). "Forward, tangent linear, and adjoint Runge-Kutta methods for stiff chemical kinetic simulations". eng. In: *International journal of computer mathematics* 87.11, pp. 2458–2479. ISSN: 0020-7160.

Smith, L. A, C Ziehmann, and K Fraedrich (1999). "Uncertainty dynamics and predictability in chaotic systems". eng. In: *Quarterly journal of the Royal Meteorological Society* 125.560, pp. 2855–2886. ISSN: 0035-9009.

Strogatz, Steven H (Steven Henry). (2000). "Bifurcations". eng. In: *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering /*. Studies in nonlinearity. Cambridge, MA: Westview, pp. 44–92. ISBN: 0738204536.

Toth, Z and E Kalnay (1993). "Ensemble Forecasting at NMC: The Generation of Perturbations". eng. In: *Bulletin of the American Meteorological Society* 74.12, pp. 2317–2330. ISSN: 0003-0007.

Van Dyke, M. (1982). *An Album of Fluid Motion*. An Album of Fluid Motion del 2. Parabolic Press. ISBN: 9780915760022. URL: https://books.google.dk/books?id=cWoeAQAAIAAJ.

Winkler, Jens, Michael Denhard, and Bernhard A Schmitt (2020). "Krylov methods for adjoint-free singular vector based perturbations in dynamical systems". eng. In: *Quarterly journal of the Royal Meteorological Society* 146.726, pp. 225–239. ISSN: 0035-9009.

Yamada, Michio and Koji Ohkitani (June 1988). "The Inertial Subrange and Non-Positive Lyapunov Exponents in Fully-Developed Turbulence". In: *Progress of Theoretical Physics* 79.6, pp. 1265–1268. ISSN: 0033-068X. DOI: 10.1143/PTP.79.1265. eprint: https://academic.oup.com/ptp/article-pdf/79/6/1265/5295968/79-6-1265.pdf. URL: https://doi.org/10.1143/PTP.79.1265.

Zhang, Zuojun (Jan. 1988). "The linear study of zonally asymmetric barotropic flows". PhD thesis. Reading Univ. (England).

Appendices

A Alternative forms of the NSE

In this appendix we derive the NSE on spectral and non-dimensional form, respectively. The derivations follow Ditlevsen, 2011. The spectral form is very useful in many theoretical calculations, and it makes the structural resemblance of the Sabra shell model to the NSE very clear. The non-dimensional form is used to define the Reynolds number.

A.1 Spectral form

We start by presenting the Fourier transform of the velocity field $u_i(\mathbf{x})$:

$$\mathcal{F}_- : \quad u_i(\mathbf{k}) = \frac{1}{(2\pi)^3} \int e^{\iota \mathbf{k}\mathbf{x}} u_i(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{A.1}$$

$$\mathcal{F}_+ : \quad u_i(\mathbf{x}) = \int e^{\iota \mathbf{k}\mathbf{x}} u_i(\mathbf{k}) \mathrm{d}\mathbf{k}, \tag{A.2}$$

where $\mathbf{k}$ is the wave vector corresponding to $\mathbf{x}$. We calculate the Fourier transform of each term in eq. (2.18) by using those transformations:

$$\mathcal{F}_-[u_j \partial_j u_i] = \frac{1}{(2\pi)^3} \int u_j(\mathbf{x}) \partial_j u_i(\mathbf{x}) e^{-\iota \mathbf{k}\mathbf{x}} d\mathbf{x} \tag{A.3a}$$

$$= \frac{1}{(2\pi)^3} \iiint u_j(\mathbf{k}') \iota k_j'' u_i(\mathbf{k}'') e^{\iota(\mathbf{k}'+\mathbf{k}''-\mathbf{k})\mathbf{x}} d\mathbf{k}' d\mathbf{k}'' d\mathbf{x} \tag{A.3b}$$

$$= \iint \iota k_j'' u_j(\mathbf{k}') u_i(\mathbf{k}'') \delta(\mathbf{k}'+\mathbf{k}''-\mathbf{k}) d\mathbf{k}' d\mathbf{k}'' \tag{A.3c}$$

$$= \int \iota (k_j - k_j') u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}' \tag{A.3d}$$

$$= \int \iota k_j u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}' \tag{A.3e}$$

where in the last equation we have used the spectral incompressibility condition $k_j' u_j(\mathbf{k}') = 0$ (the Fourier transform of eq. (2.19)). Using the same procedure we get

$$\mathcal{F}_-[\partial_{jj} u_i] = \frac{1}{(2\pi)^3} \int \partial_{jj} u_i(\mathbf{x}) e^{-\iota \mathbf{k}\mathbf{x}} d\mathbf{x} \tag{A.4a}$$

$$= -\frac{1}{(2\pi)^3} \iint k_j' k_j' u_i(\mathbf{k}') e^{\iota(\mathbf{k}'-\mathbf{k})\mathbf{x}} d\mathbf{k}' d\mathbf{x} \tag{A.4b}$$

$$= -k^2 \int u_i(\mathbf{k}') \delta(\mathbf{k}'-\mathbf{k}) d\mathbf{k}' \tag{A.4c}$$

$$= -k^2 u_i(\mathbf{k}) \tag{A.4d}$$

and

$$\mathcal{F}_-[\partial_i p] = \frac{1}{(2\pi)^3}\int \partial_i p(\mathbf{x}) e^{-\iota \mathbf{kx}} d\mathbf{x} \tag{A.5a}$$

$$= \frac{1}{(2\pi)^3}\iint \iota k_i' p(\mathbf{k}') e^{\iota(\mathbf{k}'-\mathbf{k})\mathbf{x}} d\mathbf{k}' d\mathbf{x} \tag{A.5b}$$

$$= \int \iota k_i' p(\mathbf{k}')\delta(\mathbf{k}'-\mathbf{k}) d\mathbf{k}' \tag{A.5c}$$

$$= \iota k_i p(\mathbf{k}), \tag{A.5d}$$

where a transform of $p(\mathbf{x})$ similar to eq. (A.2) has been used. $p(\mathbf{k})$ can be calculated from the Fourier transform of eq. (2.20b). From a similar calculation as in eq. (A.4) we first have

$$\mathcal{F}_-[\partial_{ii} p] = -k^2 p(\mathbf{k}) \tag{A.6}$$

while

$$\mathcal{F}_-[-\partial_i u_j \partial_j u_i] = -\frac{1}{(2\pi)^3}\int \partial_i u_j(\mathbf{x})\partial_j u_i(\mathbf{x}) e^{-\iota \mathbf{kx}} d\mathbf{x} \tag{A.7a}$$

$$= -\frac{1}{(2\pi)^3}\iint \iota k_i' u_j(\mathbf{k}') e^{\iota \mathbf{k}'\mathbf{x}} d\mathbf{k}' \int \iota k_j'' u_i(\mathbf{k}'') e^{\iota \mathbf{k}''\mathbf{x}} d\mathbf{k}'' e^{-\iota \mathbf{kx}} d\mathbf{x} \tag{A.7b}$$

$$= \frac{1}{(2\pi)^3}\iiint k_i' k_j'' u_j(\mathbf{k}') u_i(\mathbf{k}'') e^{\iota(\mathbf{k}'+\mathbf{k}''-\mathbf{k})\mathbf{x}} d\mathbf{k}' d\mathbf{k}'' d\mathbf{x} \tag{A.7c}$$

$$= \iint k_i' k_j'' u_j(\mathbf{k}') u_i(\mathbf{k}'')\delta(\mathbf{k}'+\mathbf{k}''-\mathbf{k}) d\mathbf{k}' d\mathbf{k}'' \tag{A.7d}$$

$$= \int k_i'(k_j - k_j') u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}' \tag{A.7e}$$

$$= \int k_i' k_j u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}', \tag{A.7f}$$

Combining eqs. (A.6) and (A.7f) gives

$$p(\mathbf{k}) = -\int \frac{k_i' k_j}{k^2} u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}'. \tag{A.8}$$

The spectral transform of the first and last term of eq. (2.18) are simply $\partial_t u_i(\mathbf{k})$ and $f_i(\mathbf{k})$. We can now combine all terms through a couple of index redefinitions to get the spectral form of the NSE (eq. (A.9b)):

$$\partial_t u_i(\mathbf{k}) = -\int \iota k_j u_j(\mathbf{k}') u_i(\mathbf{k}-\mathbf{k}') d\mathbf{k}'$$

$$
\begin{aligned}
&+\iota k_i \int \frac{k'_l k_j}{k^2} u_j(\mathbf{k}') u_l(\mathbf{k}-\mathbf{k}') d\mathbf{k}' \\
&-\nu k^2 u_i(\mathbf{k}) + f_i(\mathbf{k}) \qquad \text{(A.9a)} \\
&= -\iota k_j \int \Big(\delta_{il} - \frac{k'_l k_i}{k^2}\Big) u_j(\mathbf{k}') u_l(\mathbf{k}-\mathbf{k}') d\mathbf{k}' \\
&-\nu k^2 u_i(\mathbf{k}) + f_i(\mathbf{k}) \qquad \text{(A.9b)}
\end{aligned}
$$

A.2 Non-dimensional form

We define non-dimensional variables from the length scale, L, of the largest variations of the flow described by the NSE equation; L is normally defined from the size of the container holding the fluid or an obstacle embedded in the flow. The typical velocity of the flow at this scale we define as U, and the typical timescale at that scale is then $T = L/U$; that is, the typical time it takes for the fluid to travel the distance L (Ditlevsen, 2011). The non-dimensional variables (written with tildes) are then

$$x = L\tilde{x};\ \ u = U\tilde{u};\ \ t = (L/U)\tilde{t}. \qquad \text{(A.10)}$$

Inserting into eq. (2.18) we get

$$\frac{U^2}{L}\tilde{\partial}_t \tilde{u}_i + \frac{U^2}{L}\tilde{u}_j \tilde{\partial}_j \tilde{u}_i = -\frac{U^2}{L}\tilde{\partial}_i \tilde{p} + \frac{\nu U}{L^2}\tilde{\partial}_{jj}\tilde{u}_i + \tilde{f}_i, \qquad \text{(A.11)}$$

where we have used eq. (2.20c) to make the pressure gradient non-dimensional. Dividing through by U^2/L and dropping the tildes for convenience we get the non-dimensional NSE:

$$\partial_t u_i + u_j \partial_j u_i = -\partial_i p + \mathsf{R}e^{-1}\partial_{jj} u_i + f_i, \qquad \text{(A.12)}$$

where a factor U^2/L has been absorbed into the forcing term. $\mathsf{R}e$ is the Reynolds number defined in eq. (2.21).

B Timescale analyses

B.1 Lorentz-63 model

In this section, the two characteristic timescales of the Lorentz-63 model are estimated.

The short timescale, that corresponds to the average rotation period around the centre of a wing of the attractor, is estimated from a power spectrum analysis of the z-component of the full reference trajectory (see sec. 5.4). The spectrum is shown in fig. B.1. Using the Python package **scipy.signal** to find the period that corresponds to the peak with the largest amplitude, we get 0.76tu. This value corresponds well with L. Magnusson et al., 2008 that reported 0.75tu.

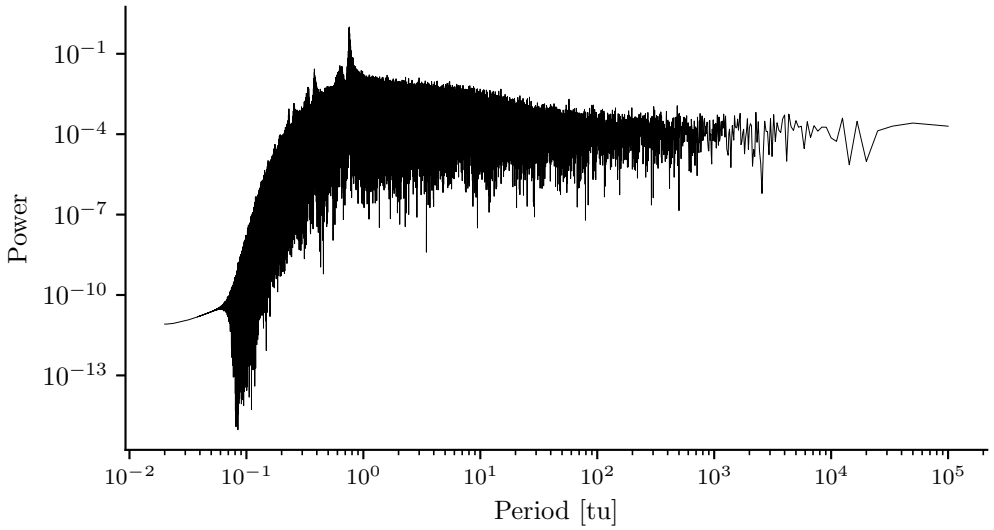


Figure B.1: Power spectrum of the z component of the Lorentz-63 model.

The long timescale, that corresponds to the average residence time in a wing of the attractor, is estimated from the distribution of residence times shown in fig. B.2. The residence times are calculated as the time between crossings of $x = 0$ of the full reference trajectory. As seen in the figure, the distribution follows an exponential distribution, which has a characteristic timescale known as the e-folding time. By fitting a linear function ($f(x) = ax$) to the logged distribution, we estimate the long timescale as $-1/a = 1.35$tu. L. Magnusson et al., 2008 finds a value of 1.8tu by using a (to the author) unknown statistical method, which makes it difficult to compare and explain the discrepancy.

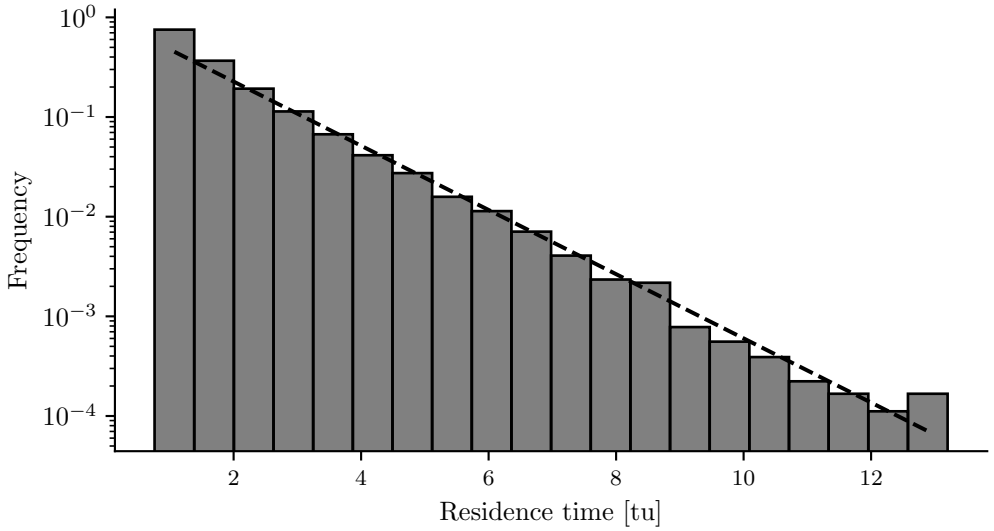


Figure B.2: The PDF of the residence time in a wing of the attractor. Distribution is made from the full reference run (see sec. 5.4). A linear fit is made ($f(x) = ax$ with $a = -0.74\text{tu}^{-1}$) to the logged histogram to get the residence time.

B.2 Sabra shell model

The characteristic time scales of the Sabra shell model, also called the *eddy turnover times*, can be estimated from the wave vectors and the shell velocities. Using dimensional arguments we have from eq. (3.9) that

$$\dot{u}_n \sim \frac{u_n}{t_n} \sim k_n u_n^2 \tag{B.1a}$$

$$t_n \sim \frac{2\pi}{k_n |u_n|}, \tag{B.1b}$$

where the factor 2π appears due to the conversion from wave vectors defined by eq. (3.8) and the constants defined in sec. 5.4 to wave vectors on the form $k_n = 2\pi n/L$. The absolute value of u_n is needed to get a real valued time estimate. The eddy turnover times are then calculated from a 1000tu average of the reference u_n (see sec. 5.4.2) and eq. (3.8):

Table B.1: The average eddy turnover times calculated from 1000tu of the reference run.

n	1	2	3	4	5	6	7	8	9
t_n [tu]	2.07	2.27	1.09	0.84	0.50	0.35	0.22	0.15	0.10
n	11	12	13	14	15	16	17	18	19
t_n [tu]	0.043	0.028	0.019	0.013	0.0082	0.0058	0.0040	0.0033	0.0042

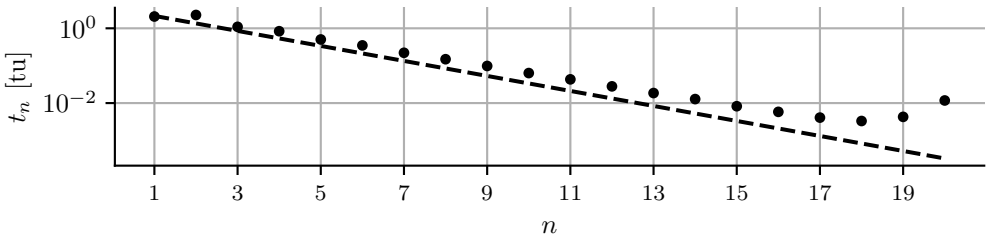


Figure B.3: The average eddy turnover times calculated from 1000tu of the reference run (dots). The scaling follows $k^{-2/3}$ (dashed line) given from eq. (2.24) and eq. (B.1b).

Notice how the eddy turnover times deviates from the $k^{-2/3}$ scaling at high shell numbers, which is due to the effect of viscosity on the shell velocities as seen in fig. 3.4a; the scaling relation is only valid in the inertial range $n \sim [3; 18]$ as assumed in the derivation of the K41 theory (sec. 2.3.2).

C Conservation of energy; Sabra shell model

In this appendix, we show the derivations of the conservation of energy for the Sabra shell model.

$$\dot{E} = \frac{d}{dt}\sum_{n=1}^{N}\frac{1}{2}u_n u_n^* \tag{C.1a}$$

$$= \frac{1}{2}\sum_{n=1}^{N}(\dot{u}_n u_n^* + u_n \dot{u_n^*}) \tag{C.1b}$$

We focus on the first term of the sum in the following, but a similar calculation can be done for the second term. In the second equal sign we interchange the last term in the sum with its complex conjugate, which appears in the before mentioned calculation.

$$\frac{1}{2}\sum_{n=1}^{N}\dot{u}_n u_n^* = \frac{1}{2}\sum_{n=1}^{N}\iota k_n(a u_n^* u_{n+1}^* u_{n+2} + b u_{n-1}^* u_n^* u_{n+1} + c u_{n-2} u_{n-1} u_n^*) \tag{C.2a}$$

$$= \frac{1}{2}\sum_{n=1}^{N}\iota k_n(a u_n^* u_{n+1}^* u_{n+2} + b u_{n-1}^* u_n^* u_{n+1} - c u_{n-2}^* u_{n-1}^* u_n) \tag{C.2b}$$

$$= \frac{1}{2}\Big[\sum_{n}^{N}\iota k_n a u_n^* u_{n+1}^* u_{n+2} + \sum_{n=-1}^{N+1}\iota k_{n+1} b u_n^* u_{n+1}^* u_{n+2}$$

$$- \sum_{n=-2}^{N+2}\iota k_{n+2} c u_n^* u_{n+1}^* u_{n+2}\Big] \tag{C.2c}$$

$$= \frac{1}{2}\sum_{n=1}^{N}\iota(a k_n + b k_{n+1} - c k_{n+2}) u_n^* u_{n+1}^* u_{n+2} + \text{boundary terms} \tag{C.2d}$$

Inserting $a = 1, b = -\frac{\epsilon}{\lambda}, c = -\frac{\epsilon - 1}{\lambda^2}$ and using that all boundary terms are zero, we get

$$= \frac{1}{2}\sum_{n=1}^{N}\iota(k_n - \frac{\epsilon}{\lambda}k_{n+1} + \frac{\epsilon - 1}{\lambda^2}k_{n+2}) u_n^* u_{n+1}^* u_{n+2} \tag{C.2e}$$

$$= \frac{1}{2}\sum_{n=1}^{N} \iota k_n (1 - \frac{\epsilon}{\lambda}\lambda + \frac{\epsilon - 1}{\lambda^2}\lambda^2) u_n^* u_{n+1}^* u_{n+2} = 0, \tag{C.2f}$$

where we used eq. (3.8) in the second to last equal sign.

D Time region analysis; Sabra shell model

In this appendix we present how we distinguish between time regions dominated by the small and large scales in the Sabra shell model. An algorithm has been developed for this purpose, such that the perturbation methods can be studied separately in those regions.

When the small scales dominate, i.e. a positive energy anomaly is present at large shell numbers (see fig. 3.4c), the dynamics happen mainly close to the Kolmogorov scale. This means that the viscosity has a significant impact on the dynamics, and energy is dissipated. Contrarily, when the large scales dominate, i.e. a positive energy anomaly is present at small shell numbers, the effect from viscosity is negligible and the energy content of the flow increases due to the forcing. For that reason, one way to distinguish between when the small or large scales dominate is to look at the slope of the total energy; negative slope means dissipation and the flow is dominated by the small scales, positive slope means increase in energy and the flow is dominated by the large scales.

After having calculated the derivative of the total energy and evaluated the sign of the slope as function of time (negative = False; positive = True), this boolean data is filtered through a so-called erosion and dilation algorithm in order to set a minimum duration of the time regions. While erosion removes entries in a data series by setting them to zero/False, dilation adds entries by setting them to one/True. The removal/adding of entries depends on the size of a kernel of the algorithm. The kernel determines the region around a data point that is used to evaluate if the data point should be removed/added or not. In fig. D.1 we see the effect of erosion and dilation on a one dimensional boolean array with a kernel of size 3.

To determine the distinct time regions in the Sabra shell model we use a kernel with size equal to 0.05tu and apply the sequence of operations: *erosion* $\rightarrow$ *dilation* $\rightarrow$ *dilation* $\rightarrow$ *erosion*. An example of the resulting regions is shown in fig. D.2c, while the effect of the erosion and dilation algorithm can be seen by comparing figs. D.2b and D.2c.

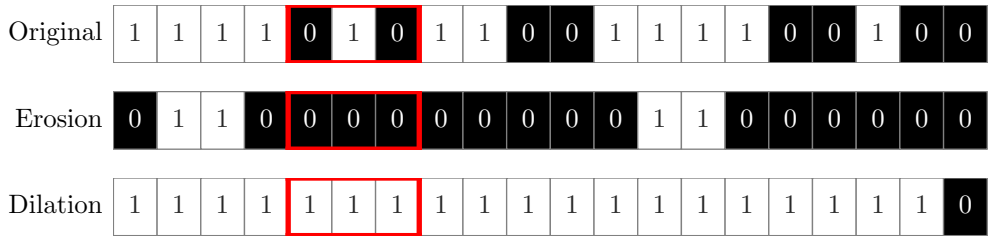


Figure D.1: Visualization of erosion and dilation with a kernel of size 3. To evaluate the result of these operations for a given entry in the original array, the kernel is imagined to be positioned to have its centre at this entry (visualized as the red rectangle). The algorithm evaluates if any value within the kernel is zero/False (erosion) or one/True (dilation). If yes, the value is changed; if not, it continues.

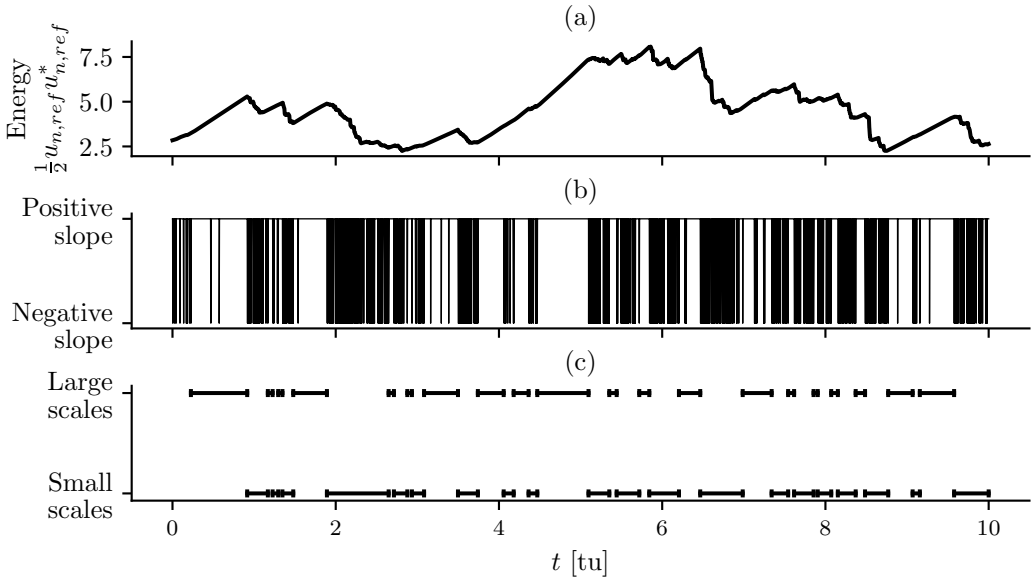


Figure D.2: An example of the effect of the erosion and dilation for determining the regions dominated by the small and large scales. (a) The total energy vs time. (b) The sign of the slope of the total energy curve. The vertical lines are drawn to show that the sign changes (c) The result of applying the erosion and dilation algorithm on the data from (b), after being converted to a boolean array. Regions dominated by the small/large scales are marked by the horizontal segments.

E Lanczos algorithm

In this appendix the detailed procedure of the Lanczos algorithm is outlined. The algorithm is used to find the eigenvectors and -values of a hermitian propagator or matrix (Demmel, 1997; Meurant et al., 2006).

Input: A hermitian propagator or matrix A of dimension $M \times M$ that maps a vector $\mathbf{v} \mapsto A\mathbf{v}$.

Output: The matrices $\mathbf{P}$ $(M \times N)$, $\mathbf{T}$ $(M \times M)$, where $\mathbf{T}$ is tridiagonal and the columns of $\mathbf{P}$ are given by the orthonormal basis $\{\mathbf{p}_i\}$.

Steps:

1. Choose a random vector $\mathbf{p}_1 \in \mathbb{C}^m$ with norm 1.
2. First iteration
   (a) $\mathbf{w}'_1 = A\mathbf{p}_1$
   (b) $\alpha_1 = \langle \mathbf{w}'^*_1; \mathbf{p}_1 \rangle$
   (c) $\mathbf{w}_1 = \mathbf{w}'_1 - \alpha_1 \mathbf{p}_1$
3. For $M \geq j \geq 2$
   (a) $\beta_j = ||\mathbf{w}_{j-1}||$
   (b) **If** $\beta_j \neq 0$ **then** $\mathbf{p}_j = \frac{1}{\beta_j}\mathbf{w}_{j-1}$,
   else Pick a random normalized vector orthogonal to all $\{\mathbf{p}_1, ..., \mathbf{p}_{j-1}\}$
   (c) $\mathbf{w}'_j = A\mathbf{p}_j$
   (d) $\alpha_j = \langle \mathbf{w}'^*_j; \mathbf{p}_j \rangle$
   (e) $\mathbf{w}_j = \mathbf{w}'_j - \alpha_j \mathbf{p}_j - \beta_j \mathbf{p}_{j-1}$ $\star$
   Re-orthogonalize
   (f) **for** $i < j$ **do**:
   $\mathbf{w}_j = \mathbf{w}_j - \langle \mathbf{w}_j; \mathbf{p}_i \rangle \mathbf{p}_i$
   end
4. Construct the matrices $\mathbf{P}$ and $\mathbf{T}$, where $T_{ij} = \alpha_i \delta_{i,j} + \beta_i \delta_{i,j+1} + \beta_j \delta_{i+1,j}$ or

$$\mathbf{T} = \mathbf{P}^* \mathbf{A} \mathbf{P} \tag{E.1}$$

5. Diagonalise $\mathbf{T}$ to get the eigenvalues, λ_i, and -vectors, $\tilde{\mathbf{v}}_i$, i.e. solve the eigenvalue problem

$$\mathbf{T}\tilde{\mathbf{v}}_i = \lambda_i \tilde{\mathbf{v}}_i \ \star \tag{E.2}$$

6. Project the $\tilde{\mathbf{v}}_i$ onto $\mathbf{P}$ to get the eigenvectors, $\mathbf{v}_i$, of A with corresponding eigenvalue λ_i, i.e.

$$\mathbf{v}_i = \mathbf{P}\tilde{\mathbf{v}}_i \tag{E.3}$$

7. If $A = L^*L$, the singular vectors of L is given by $\mathbf{v}_i$, and the singular values by $\sigma_i = \sqrt{\lambda_i}$

One can show the validity of item 6 by the following:

$$A\mathbf{v}_i = A\mathbf{P}\tilde{\mathbf{v}}_i = \mathbf{P}\mathbf{T}\mathbf{P}^*\mathbf{P}\tilde{\mathbf{v}}_i = \mathbf{P}\mathbf{T}\tilde{\mathbf{v}}_i = \mathbf{P}\lambda_i\tilde{\mathbf{v}}_i = \lambda_i\mathbf{P}\tilde{\mathbf{v}}_i = \lambda_i\mathbf{v}_i, \ \star \tag{E.4}$$

where we in the second equal sign from left used eq. (E.1); in the third equal sign from left used that the columns of $\mathbf{P}$ form an orthonormal basis such that $\mathbf{P}$ is unitary, i.e. $\mathbf{P}^*\mathbf{P} = \mathbf{I}$; in the third equal sign from the right used eq. (E.2).

F Supplementary plots; Lorentz-63 model

F.1 Perturbations along a trajectory

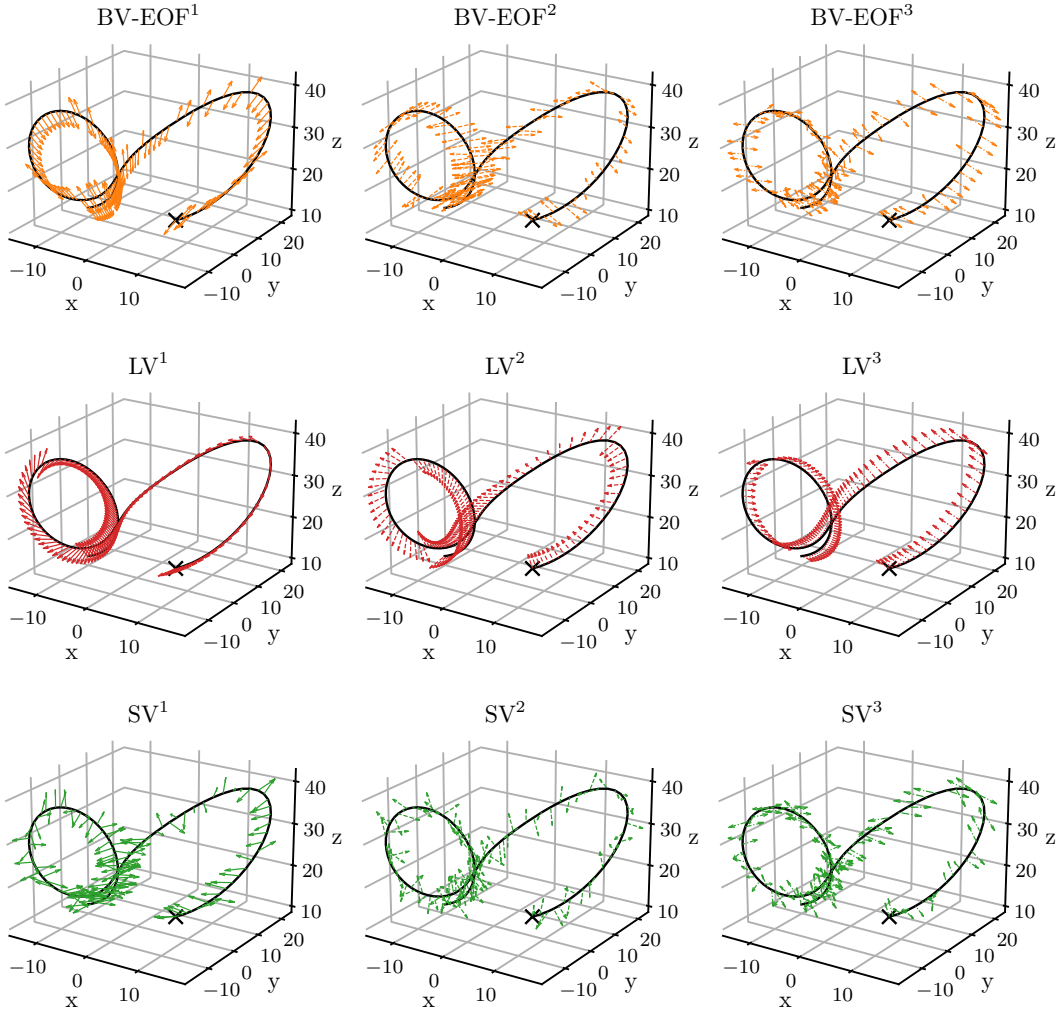


Figure F.1: The perturbations plotted as vectors for the BV-EOF$_{\text{ens}}$, LLV$_{\text{ens}}$ and SV$_{\text{ens}}$ along a trajectory (black line) starting at the black cross. The trajectory starts at $t = 10007.8$tu and ends at $t = 10009.2$tu (approximately the same time span as that shown in fig. 6.2). 140 ensembles of initial perturbations are calculated sequentially in time separated by 0.01tu. Note the very regular perturbations produced by the LLV method in contrary to those of the SV method.

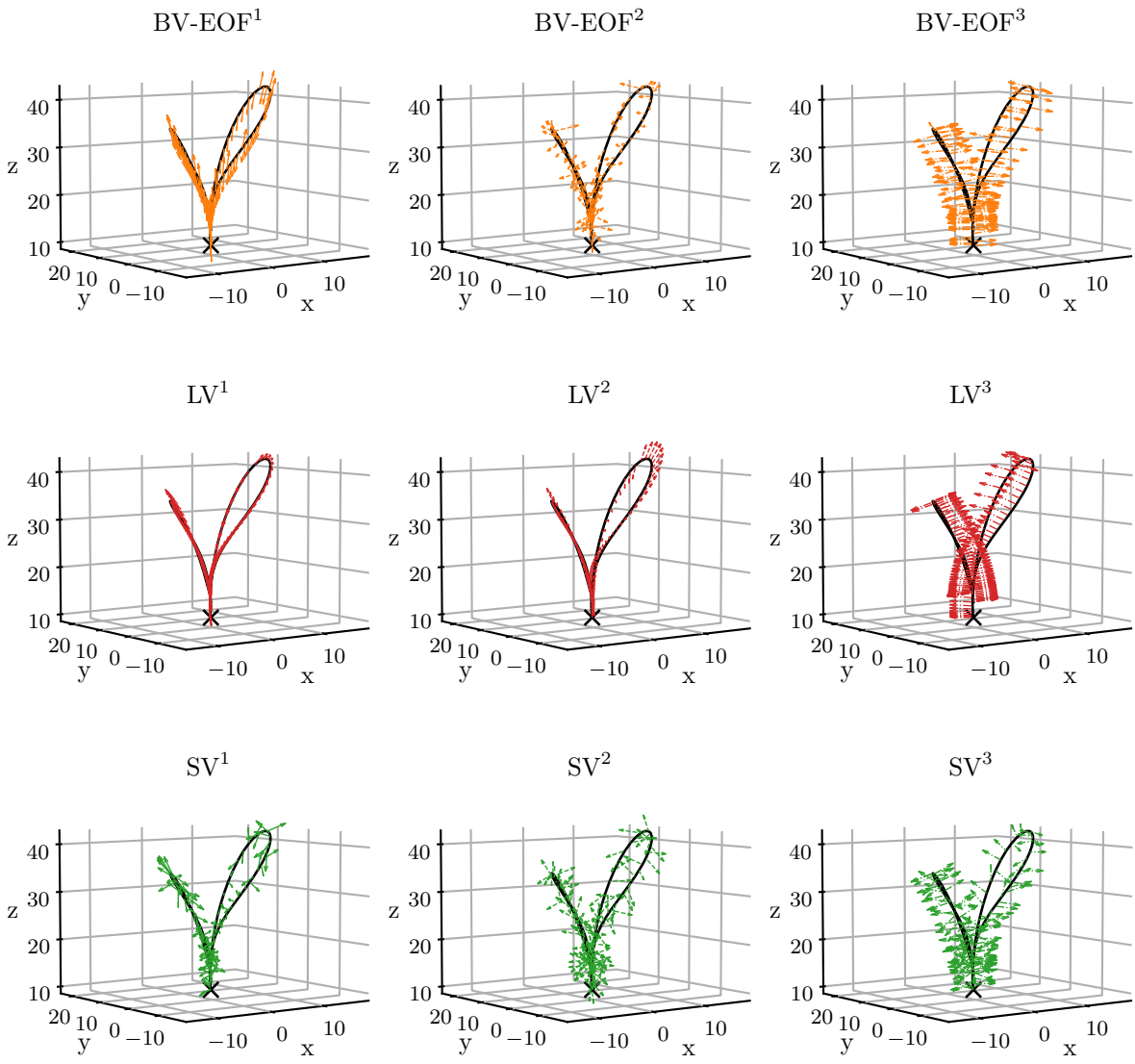


Figure F.2: As in fig. F.1, but with a different view angle (parallel to the attractor plane). Note how the third member perturbations clearly lie orthogonal to the attractor plane.

G Supplementary plots; Sabra shell model

G.1 Exponential growth rates

In this section we show the same curves as in fig. 6.6 but split up into three figures to be able to better distinguish them.

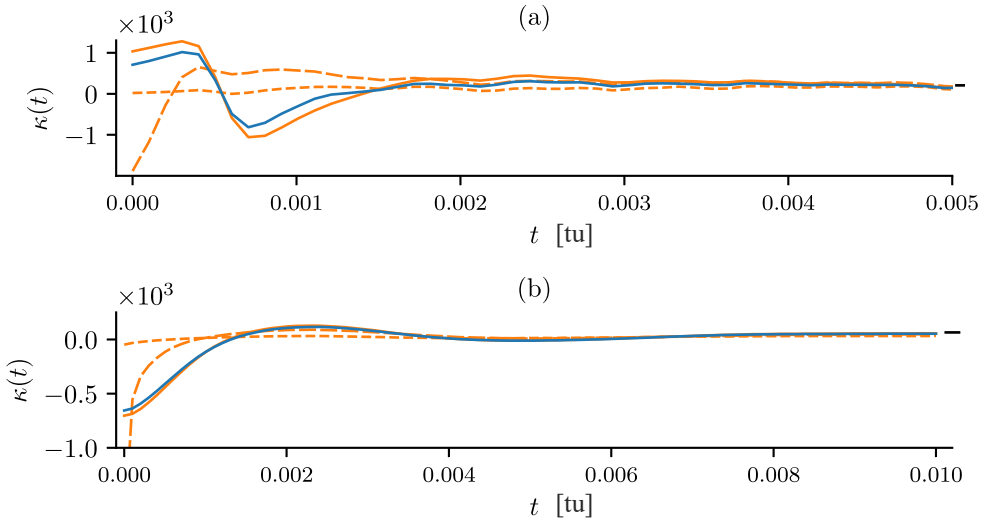


Figure G.1: Same as in fig. 6.6 but only for BV and BV-EOF methods.

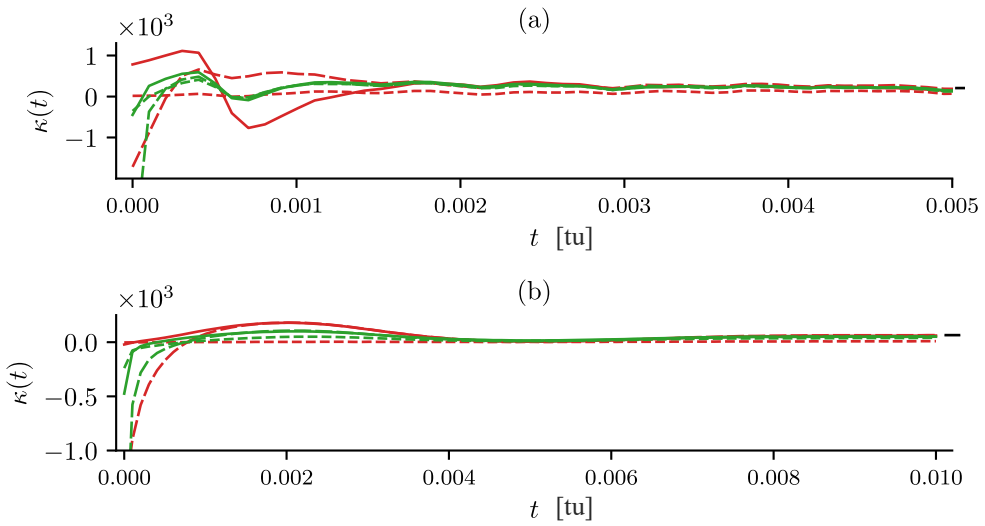


Figure G.2: Same as in fig. 6.6 but only for SV and LLV methods.

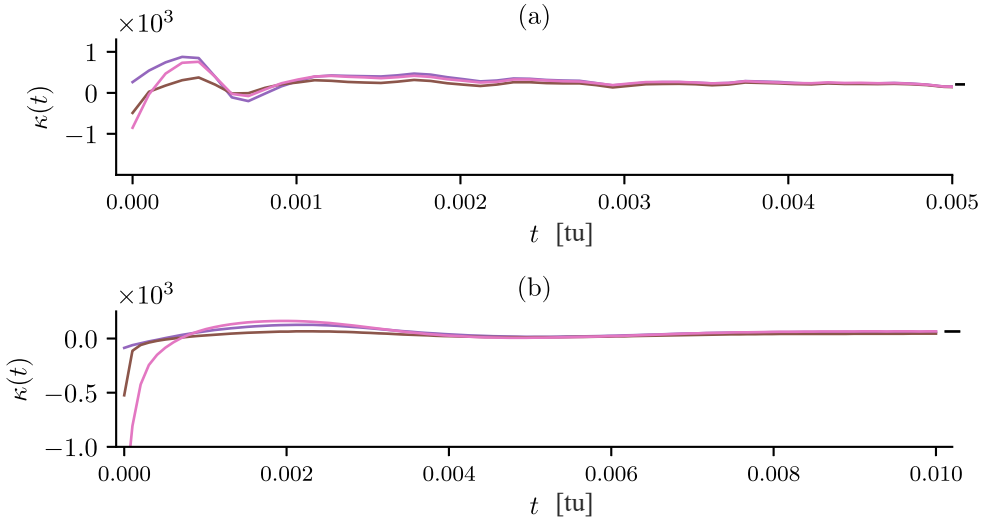


Figure G.3: Same as in fig. 6.6 but only for RD, NM and RF methods.

G.2 Exponential growth rates; $\mathrm{SV}_{\mathrm{ens}}$

In this section we present the exponential growth rates for all SVs in R_{small}.

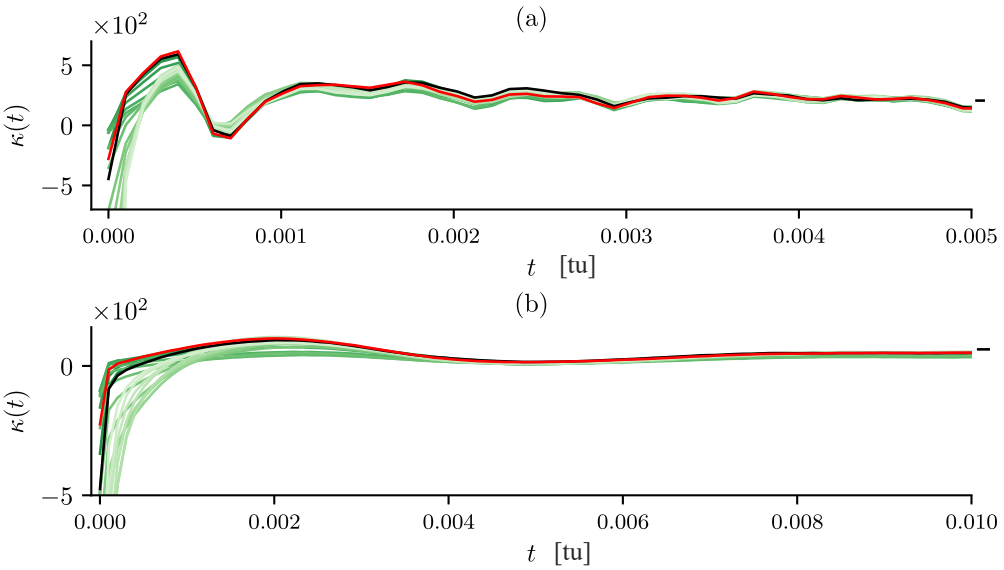


Figure G.4: Same as in fig. 6.6 but for all SVs. The curve that corresponds to the $\mathrm{SV}^1_{\mathrm{mbr}}$ is highlighted in black, while the curve that corresponds to the member experiencing the maximal growth rate is highlighted in red (in (a) $\mathrm{SV}^2_{\mathrm{mbr}}$ and (b) $\mathrm{SV}^4_{\mathrm{mbr}}$). For all other curves, the lighter the colour the higher the member index.

G.3 Distribution of adjoint LLVs

In this section we present the distribution of adjoint LLVs in R_{small}.

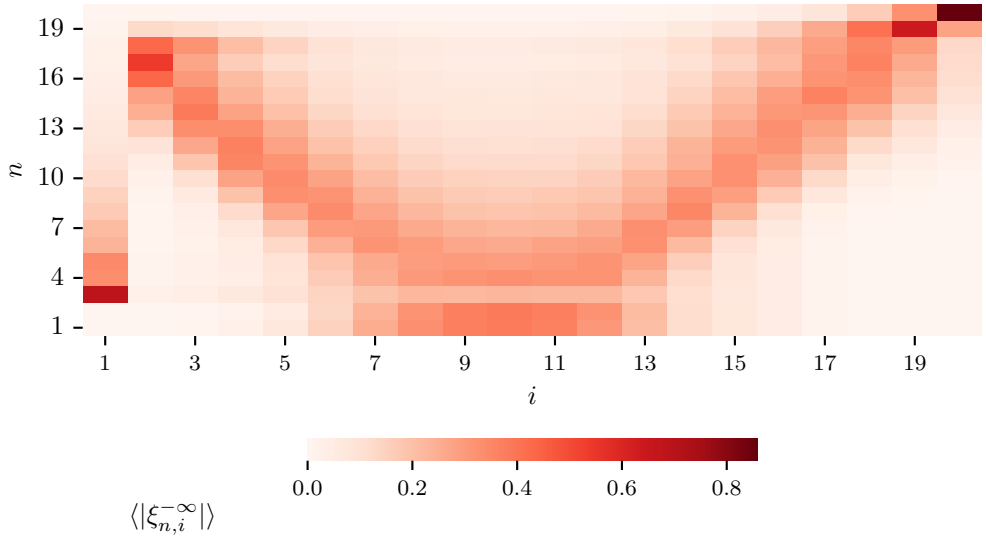


Figure G.5: The distribution of adjoint LLVs. Otherwise, as in fig. 6.7.

G.4 Distributions of perturbation vectors, R_{large}

In this section we present the average spectral distributions of perturbation vectors for the BV, BV-EOF, SV, LV and RF methods calculated in R_{large}.

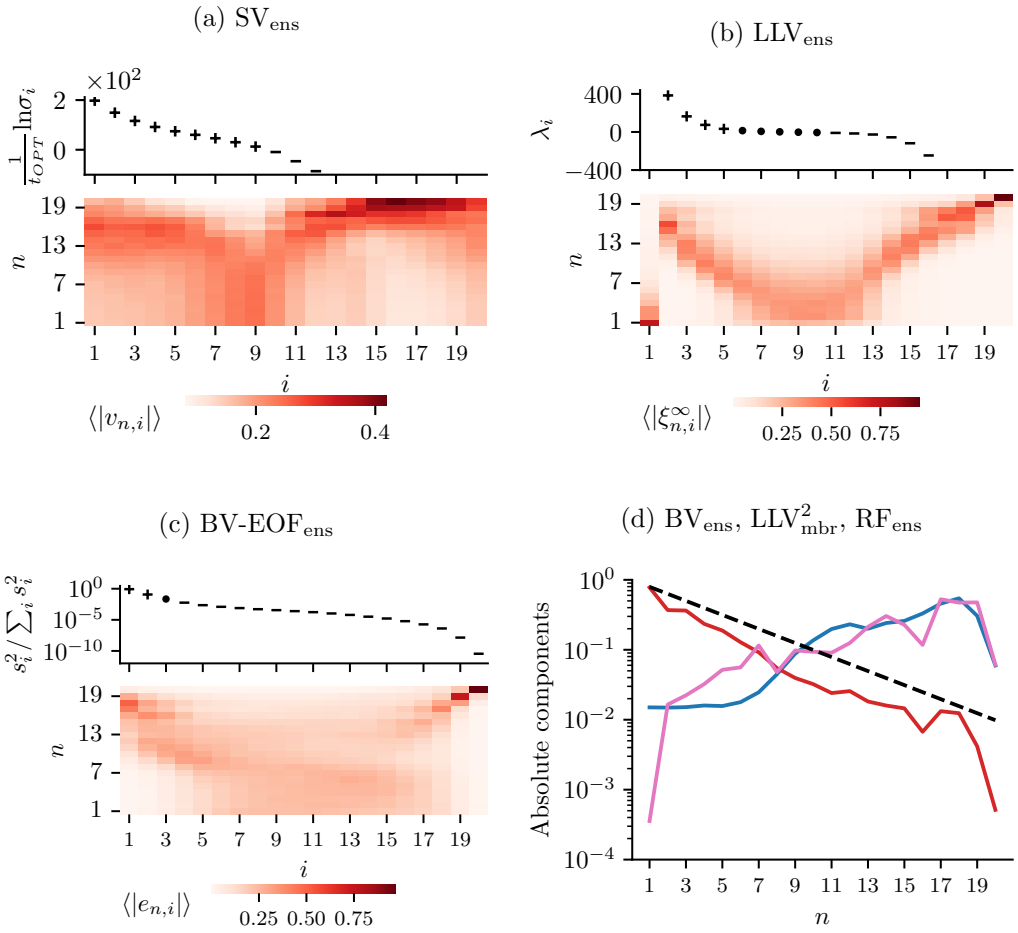


Figure G.6: As in fig. 6.7 but in R_{large}.

G.5 Results for $t_{OPT} = 0.001$tu in R_{small}

To investigate the influence

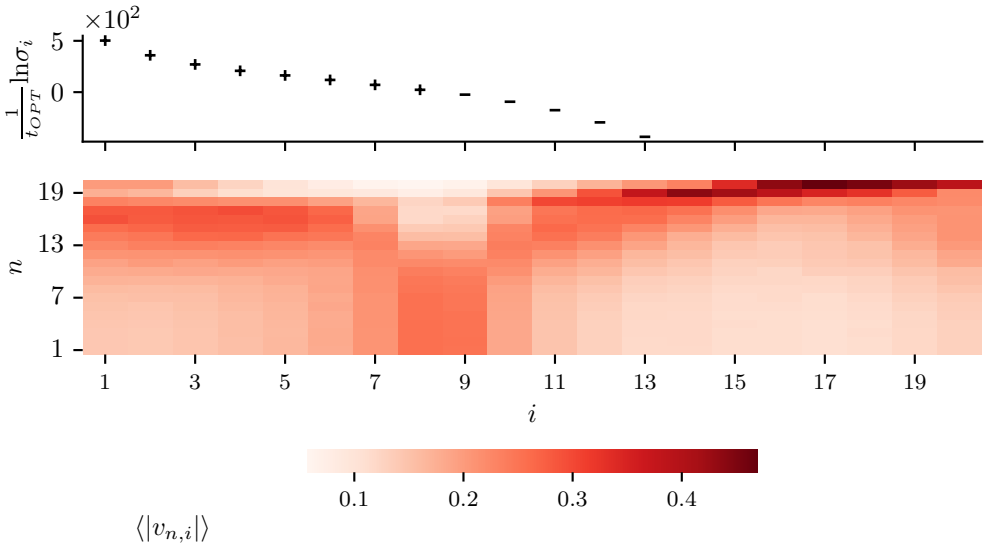


Figure G.7: The mean spectrum of SVs and corresponding exponential growth rates calculated from the singular values. As in fig. 6.7a

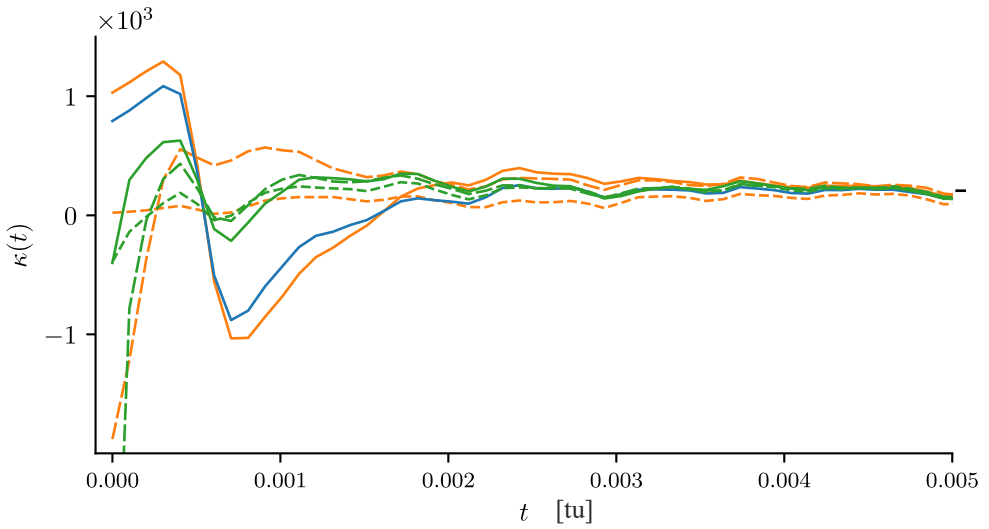


Figure G.8: The instantaneous exponential growth rate, $\kappa(t)$, for the BV, BV-EOF and SV method as in fig. 6.6a.

G.6 Results for $t_{OPT} = 0.005$tu in R_{small}

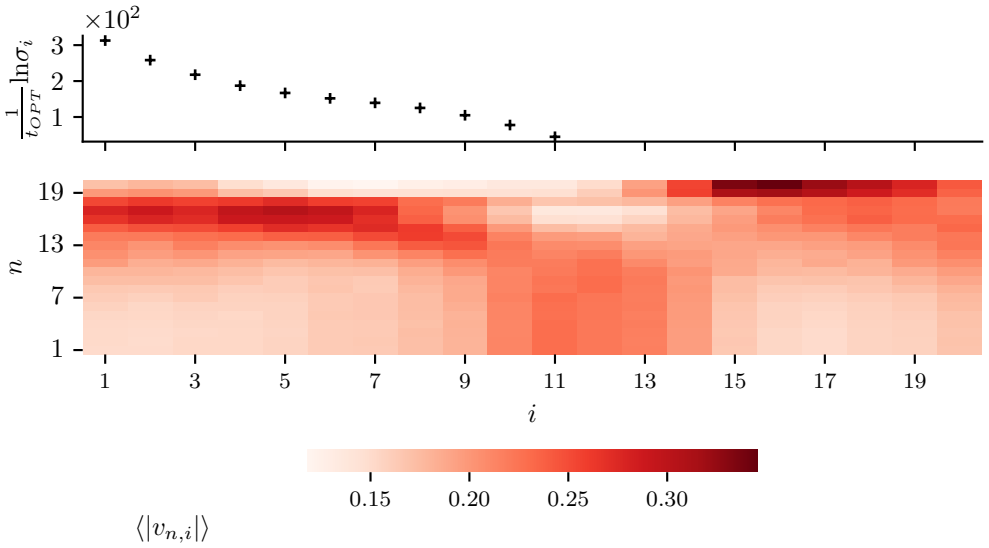


Figure G.9: The mean spectrum of SVs and corresponding exponential growth rates calculated from the singular values. As in fig. 6.7a

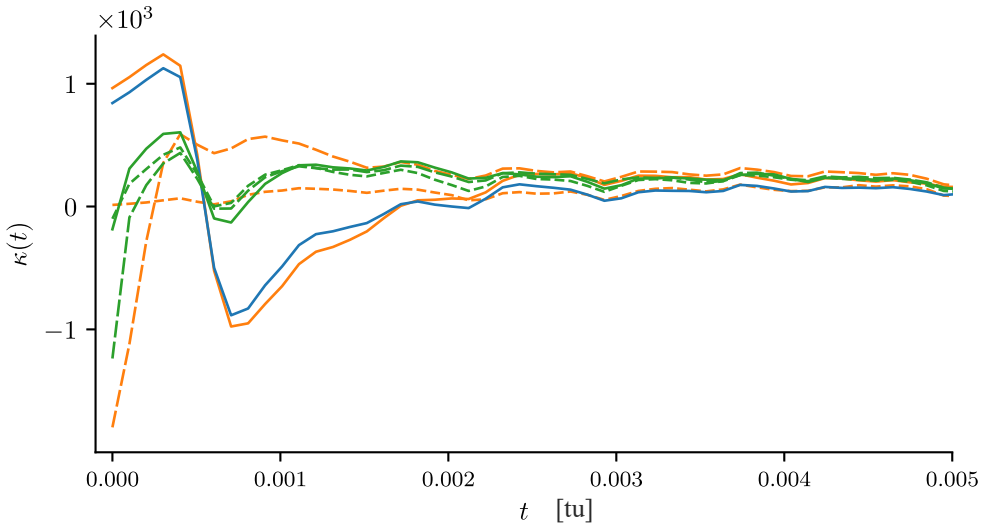


Figure G.10: The instantaneous exponential growth rate, $\kappa(t)$, for the BV, BV-EOF and SV method as in fig. 6.6a.

H Figures from (Magnusson et al., 2009, 2008)

In this chapter, we show figures from (Magnusson et al., 2009, 2008) for comparison of the results for the Lorentz-63 model.

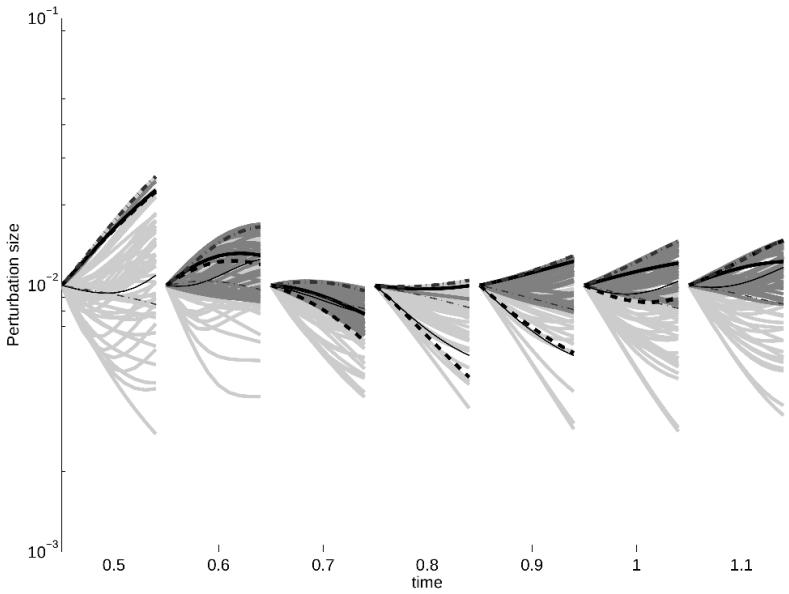


Fig. 1. Evolution of the difference between the central member and perturbed members, for 7 equally spread initial points taken along a trajectory. Rescaled every 0.1 tu. BV (black, dashed), RP (light-grey, solid),NM (grey, solid), first SV (dark-grey, thick dash-dotted), second SV (dark-grey, thin dash-dotted), first BV-EOF (black, thick solid), second BV-EOF(black, thin solid).

Figure H.1: Figure 1 from (L. Magnusson et al., 2008)

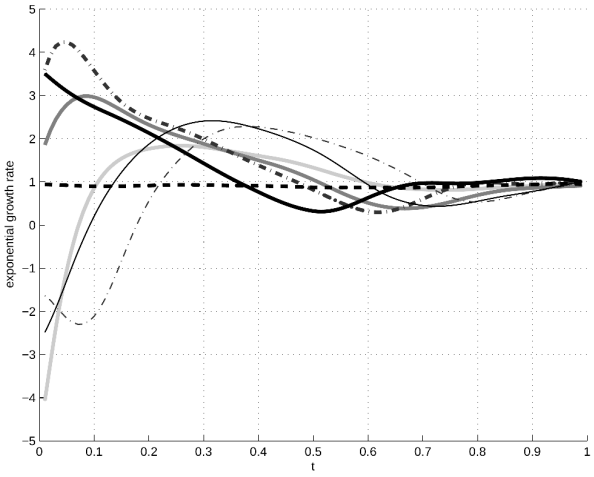


Fig. 3. Perturbation growth rate in terms of the instantaneous growth rate as an average over 5000 cases. Breeding cycle length and singular vector optimisation time 0.1 tu but the ensembles are run until 1 tu. BV (black, dashed), RP (grey, solid),NM (dark-grey, solid), first SV (black, thick dash-dotted), second SV (black, thin dash-dotted), first BV-EOF (black, thick solid), second BV-EOF (black, thin solid).

Figure H.2: Figure 3 from (L. Magnusson et al., 2008)

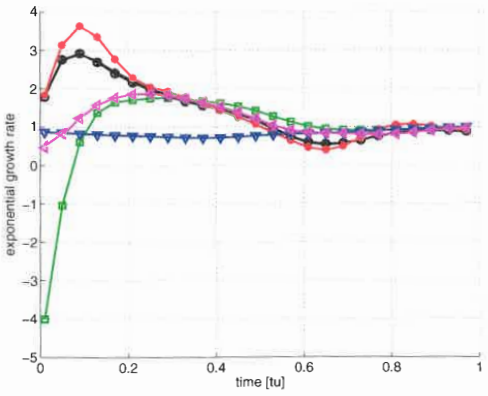


Figure 5.2: Exponential perturbation growth in the Lorenz-63 system. RP (□, green), NM (○, black), RF (+, black), BV (▽, blue), BV-EOF (◁, purple) and SV (∗, red). An average of 900 ensemble integrations, with 2 members in each (except SV there only the 1st SV is used).

Figure H.3: Figure 5.2 from (L. Magnusson, 2009). The black + symbols (RF) lie on top of the black circles (NM).

I Permission request for fig. 4.4

Subject: Re: Permission Request Form [#1077]
From: "Gumbel, Erin" <egumbel@ametsoc.org>
Date: 3/8/22, 03:28
To: rhk101@alumni.ku.dk
CC: permissions@ametsoc.org

Dear Mx. Frølund,

Thank you for your email. This signed message constitutes permission to use the material requested below.

You may include Figure 1 from Buizza and Palmer's 1995 JAS article in your Copenhagen University thesis, *Perturbation Methods for Ensemble Weather Forecasts*, with the following conditions:

1. Include the complete bibliographic citation of the original source.
2. Include the following statement with that citation: **© American Meteorological Society. Used with permission.**

Congratulations on completing your thesis! If you have any questions or need additional information, please feel free to contact me.

Please note: If the material in an AMS journal is credited to another source, the requester must obtain permission or license from that source directly. That material may not be used without permission or license from the copyright holder.

Best,

Ms. Erin Gumbel, she/her/hers
Senior Peer Review Support Associate
Senior Permissions Specialist
egumbel@ametsoc.org
617-226-3926

On Thu, Feb 3, 2022 at 3:10 AM Wufoo <no-reply@wufoo.com> wrote:

Use requested (examples include: reuse in a thesis/dissertation, journal, book, or coursepack materials; translation; inclusion in an institutional repository, etc) *	Reuse in Master Thesis at Niels Bohr Institue, Copenhagen University, Denmark. Supervisor: Peter D. Ditlevsen

Re: Permission Request Form [#1077]

Journal *	Journal of the Atmospheric Sciences
Title of AMS journal article *	The Singular-Vector Structure of the Atmospheric Global Circulation
First author of AMS journal article *	R. Buizza
Year of Publication *	1995
If you are requesting permission to reuse figures/images: List the figure number, caption, and the page(s) on which the material appears.	Fig. 1, page 1437
Requester name *	Martin Frølund
Requester Email *	rhk101@alumni.ku.dk
Requester type *	Academic/Non-Commercial
Title of new work *	Perturbation Methods for Ensemble Weather Forecasts
Format *	Electronic
Are you requesting permission for language(s) other than English? *	No