ABSTRACT

Title of Dissertation: Bubbling Transition and Data Assimilation

Aleksey V. Zimin, Doctor of Philosophy, 2003

Dissertation directed by: Professor Edward Ott
                             Department of Physics

The work presented in this thesis studies the phonomena of synchronization and bursting in dynamical systems.

Chapter I introduces the problems that are studied.

In the second chapter the bubbling transition is studied. We build theoretical models to derive the properties of a dynamical system containing an invariant manifold, when the system is perturbed in the direction transverse to the invariant manifold. We present a complete study that includes consideration of all generic bifurcations that may lead to bubbling.

In the third chapter a useful method of looking at weather model data is described. This method allows to accurately extract wave packet envelopes from atmospheric wind velocity data. We also describe possible practical applications of the method.

The last chapter describes a new data assimilation technique for numerical weather prediction. In the context of numerical weather prediction, data assimilation is the method of determining the best guess of the current state of the atmosphere, which is then used to make the subsequent forecast. From the point of view of dynamical systems, data assimilation acts as coupling between the numerical weather model and the atmosphere. In broad sense, we desire to achieve perfect syncronization between the atmosphere and the model, thus leading to good quality forecasts. Due to model imperfections, measurement errors, and sparceness of the atmospheric data, the synchronization is not perfect. We observed bursting in the RMS error of the best guess state that has a similar nature to the bursting in the synchronization studies of Chapter I. Although study of the bursts of RMS error in determining the best guess state from the bubbling transition point of view is beyond the scope of this thesis, some preliminary results hint that these bursts may be explained as bubbling.

Bubbling Transition and Data Assimilation

by

Aleksey V. Zimin

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2003

Advisory Committee:

Professor Edward Ott, Chairman/Advisor
Professor James A. Yorke, Co-Advisor
Doctor Istvan Szunyogh
Professor Eric Kostelich
Professor Brian Hunt (Dean's Representative)

DEDICATION

Dedicated to my Mother, Tatiana Zimina

and my Father, Valery Zimin,

for the support and inspiration they provide.

ACKNOWLEDGEMENTS

While working on my research on the way to my Doctoral degree, I had the great fortune to meet and collaborate with many great people. First and foremost, I would like to acknowledge and thank my advisor, Dr. Edward Ott, for his advice, encouragement, and patience, and Dr. James Yorke for his personal attention and guidance. These two people helped towards sharpening my skills of clearly perceiving and solving scientific problems by careful and complete examination and evaluation of the nature of the phenomena in question. I would like to thank Dr. Brian Hunt for his invaluable contributions to my research. I have learned many concepts in meteorology from Dr. Istvan Szunyogh who has become not only a colleague, but also a good friend. I had the good fortune to meet and collaborate with Dr. Eugenia Kalnay who has alway been very kind and supportive. I would like to specially mention Dr. D.J. Patil who originally recommended my involvement in the meteorology project and always offered support.

In the late stage of my research I have collaborated with Dr. Eric Kostelich, and I would like to thank him for his advice and patience.

I also have benefited a lot from valuable discussions with my fellow students and postdocs in the UMD Chaos group. In this regard I would like to mention Mike Ocszkovsky, Vasily Dronov, Romulus Breban, Doug Armstead, Michael Roberts, Susanne Sindi, Wayne Hayes and others.

My research progress would not have been as smooth without the technical support of Alfredo Nava-Tudela and Gyorgyi Gyarmati. I would like to also acknowledge the IPST staff for their help and trust.

Finally, I would like to thank my parents, Tatiana Zimina and Valery Zimin for their unconditional support and contributions to my education from the early stages of my life. The knowledge and genuine interest in science that I have acquired in the first decades of my life was entirely due to them, and this interest and desire to obtain more knowledge is still a major driving force for me. I would also like to thank my fiancee, Polina Vlasenko, who has been beside me for all the years in graduate school and whose advice and participation was valuable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The work presented in this thesis studies the phonomena of synchronization and bursting in dynamical systems.

Chapter I introduces the problems that are studied.

In the second chapter the bubbling transition is studied. We build theoretical models to derive the properties of a dynamical system containing an invariant manifold, when the system is perturbed in the direction transverse to the invariant manifold. We present a complete study that includes consideration of all generic bifurcations that may lead to bubbling.

In the third chapter a useful method of looking at weather model data is described. This method allows to accurately extract wave packet envelopes from atmospheric wind velocity data. We also describe possible practical applications of the method.

The last chapter describes a new data assimilation technique for numerical weather prediction. In the context of numerical weather prediction, data assimilation is the method of determining the best guess of the current state of the atmosphere, which is called analysis. Analysis is then used to make the subsequent forecast. From the point of view of dynamical systems, data assimilation acts as
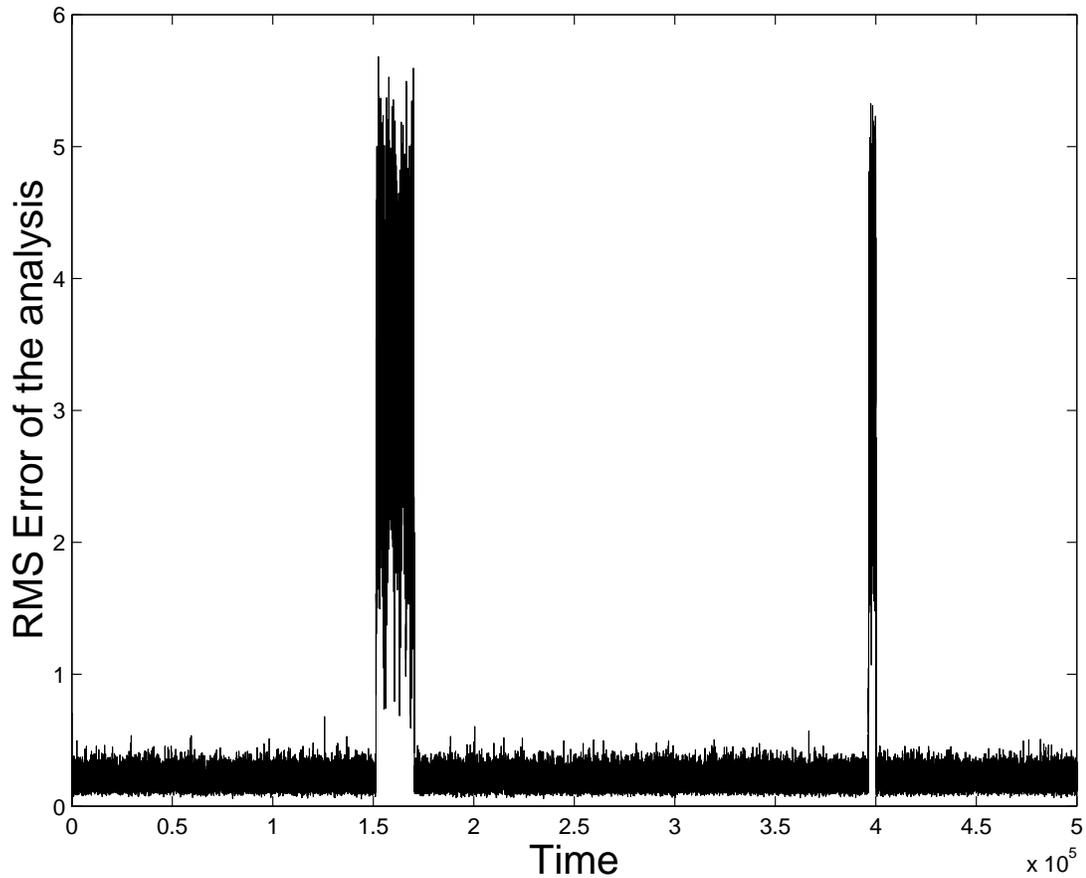
Figure 1.1: This plot shows the RMS error of the analysis verified by the "truth". Results from 500,000 data assimilation cycles are shown.

coupling between the numerical weather model and the atmosphere. In broad sense, we desire to achieve perfect syncronization between the atmosphere and the model, thus leading to good quality forecasts. Due to model imperfections, measurement errors, and sparceness of the atmospheric data, the synchronization is not perfect. We observed bursting in the RMS analysis error that has a similar nature to the bursting in the synchronization studies of Chapter I. A sample of the bursing behavior is presented in Figure 1.1. Although study of the bursts of RMS analysis error from the bubbling transition point of view is beyond the

scope of this thesis, some preliminary results hint that these bursts may be explained as bubbling. In the following paragraphs we provide short abstracts for the three projects described in this thesis.

Bifurcation scenarios for bubbling transition. Dynamical systems with chaos on an invariant submanifold can exhibit a type of behavior called bubbling, whereby a small random or fixed perturbation to the system induces intermittent bursting. The bifurcation to bubbling occurs when a periodic orbit embedded in the chaotic attractor in the invariant manifold becomes unstable to perturbations transverse to the invariant manifold. Generically the periodic orbit can become transversely unstable through a pitchfork, transcritical, period doubling, or Hopf bifurcation. In this work a unified treatment of the four types of bubbling bifurcations is presented. Conditions are obtained determining whether the transition to bubbling is soft or hard. For soft bubbling transitions, the scaling of the maximum burst amplitude with the parameter is derived. For both hard and soft transitions the scaling of the average interburst time with the bifurcation parameter is deduced. Both random (noise) and fixed (mismatch) perturbations are considered.

Extraction of the envelopes of Rossby wave packets. Packets of Rossby waves play an important role in the transfer of kinetic energy in the extra-tropics. The ability to locate, track, and detect changes in the envelope of these wave packets is vital to detecting baroclinic downstream development, tracking the impact of the analysis errors in numerical weather forecasts, and analyzing the forecast effects of targeted weather observations. In this work, it is argued that a well known technique of digital signal processing, which is based on the Hilbert transform, should be used for extracting the envelope of atmospheric wave packets.

3

This technique is robust, simple, and computationally inexpensive. The superiority of the proposed algorithm over the complex demodulation technique (the only technique previously used for this purpose in atmospheric studies) is demonstrated by examples. The skill of the proposed algorithm is also demonstrated by tracking wave packets in operational weather analyses from the National Centers for Environmental Prediction (NCEP).

A Local Ensemble Kalman Filter for atmospheric data assimilation. Recent studies have shown that, when the Earth's surface is divided up into local regions of moderate size, vectors of the forecast uncertainties in such regions tend to lie in a subspace of much lower dimension than that of the full atmospheric state vector of such a region. In this work we show how this finding can be exploited to formulate a potentially accurate and efficient data assimilation technique. The basic idea is that, since the expected forecast errors lie in a locally low dimensional subspace, the analysis resulting from the data assimilation should also lie in this subspace. This implies that operations only on relatively low dimensional matrices are required. The data assimilation analysis is done locally in a manner allowing massively parallel computation to be exploited. The local analyses are then used to construct global states for advancement to the next forecast time. The method, its potential advantages, properties, and implementation requirements are illustrated by numerical experiments on a 40-variable Lorenz model. It is found that near-optimal performance can be achieved with very modest computational cost.

Chapter 2

Bifurcation Scenarios for Bubbling Transition

[1]

2.1 Introduction

In this paper we will be concerned with dynamical systems that contain an invariant manifold embedded in their phase space and for which there exists a chaotic attractor in the invariant manifold. Such systems are common in a variety of physical situations, and they display interesting dynamical behaviors. Types of such dynamical behavior include on-off intermittancy (Pikovsky 1984; Fujisaka and Yamada 1986; Yu et al. 1990 Platt et al. 1993; Venkataramani et al. 1995), riddled basins of attraction (Alexander et al. 1992), and bubbling (Ashwin et al. 1994; Heagy et al. 1995; Gauthier and Bienfang 1996; Venkataramani et al. 1996). Examples of systems having invariant manifolds include systems with spatial symmetry (Ott and Sommerer 1994), predator-prey models (Ott et al. 1993; Sommerer and Ott 1993; Lai et al. 1996), magnetic dynamos (Sweet 2002),

---

[1]This chapter is a verbatim representation of the paper A. V. Zimin, B. R Hunt and E. Ott, 2003: Bifurcation Scenarios for Bubbling Transition. *Phys. Rev. E*, **67**, 016204.

and synchronized chaotic oscillators (Pecora and Carroll 1990). The subject of this paper is the transition to bubbling. Following (Ashwin et al. 1994), we introduce a *normal* parameter – a parameter whose variation effects the system dynamics off of the invariant manifold but leaves the dynamics within the invariant manifold unchanged. For example, in the case of synchronization of two coupled chaotic oscillators [as in Eqs. (2.1)], the coupling strength is the normal parameter. The bubbling transition occurs when, as a normal parameter is varied, a periodic orbit embedded within the chaotic attractor first becomes unstable to perturbations transverse to the invariant surface (Hunt and Ott 1996). Before the transition, all periodic orbits in the chaotic attractor are transversely stable. Beyond the bubbling transition, if the system is perturbed in a direction transverse to the invariant manifold, orbits that come close to the transversely unstable periodic orbits are repelled and move away from the invariant manifold. If there is no other attractor off the invariant manifold, the orbit returns, and, assuming that the perturbations continue, this process repeats, producing intermittent bursts away from the invariant manifold. If, on the other hand, there is another attractor not on the invariant manifold, orbits initially repelled from the periodic orbit on the invariant manifold may go to that attractor and never return. For definiteness, the following discussion will not consider the case where there is another attractor not on the invariant manifold, although later on in the paper (Sec. 6) we indicate how our results carry over to this case.

We consider the dependence of bursting on two parameters: the normal parameter and the size of perturbations transverse to the invariant manifold. These perturbations, which we assume to be small, may be random (noise) or fixed (as in the case of coupled oscillators when there may be a slight mismatch between

the two oscillators). For a given fixed value of the normal parameter beyond the bubbling transition, as the size of the perturbations goes to zero, but the typical size of bursts remains finite. Thus we can define a maximum burst amplitude as a function of the normal parameter as the maximum size of bursts in the limit as the perturbation size goes to zero.

We distinguish between two types of bubbling transitions: soft and hard. When a soft bubbling transition occurs, the maximum burst amplitude increases continuously from zero as the value of the system's normal parameter goes through its critical value. When a hard transition occurs, the maximum burst amplitude increases discontinuously from zero to a finite value at the bifurcation.

In this paper we investigate the transition to bubbling as the normal parameter of the system goes through its critical value and find the conditions on the other parameters of the system leading to soft or hard transitions. We present results for the dependence of the maximum burst amplitude on the normal parameter for a soft transition and the dependence of the average interburst time on the normal parameter and the size of the perturbations for both hard and soft transitions. As previously mentioned, the bubbling transition is marked by the loss of transverse stability by a periodic orbit of the chaotic attractor in the invariant manifold. In the presence of the invariant manifold, there are three ways by which such a loss of transverse stability can occur: the eigenvalue of the transverse map can increase through (i) $+1$, (ii) -1 or (iii) be complex and increase through the unit circle. These three ways generically correspond to: (i) a pitchfork or transcritical bifurcation, (ii) a period doubling bifurcation, and (iii) a Hopf bifurcation [2]. In (Venkataramani et al. 1996) the behavior of a system

[2]In case (i), the pitchfork bifurcation is generic for systems that are symmetric about the

7

undergoing a bubbling transition associated with a pitchfork and transcritical bifurcations (case(i)) was studied, and results for the average interburst time and maximum burst amplitude were derived. In (Yahchuk 2001) the stability of low period orbits and the transition to bubbling due to the three generic types of bifurcations was observed, and conditions determining whether the transition is hard or soft were derived for a system of coupled Rössler attractors. The bifurcation scenarios in case (i) are explored further in (Kim and Lim 2001;Kim et al. 2001; im et al. to be published)), while period-doubling (case(ii)) induced bubbling is observed in (Jalnine and Kim 2002; Kim et al. 2002)). In (Kim et al. to be published; Kim et al. 2002), the effect of both fixed and random perturbations on the average interburst time and maximum burst amplitude are considered. Our results for case (ii) in Section 2.4 are consistent with the scaling results in (Kim et al. to be published; Kim et al. 2002) for average interburst time as a function of perturbation size; however, we focus instead on the dependence on the normal parameter and on those scaling regions for which bursting is dominated by either the normal parameter or the random term.

In our paper we present a unified treatment covering all generic types of bubbling bifurcations. The contributions of the present paper are as follows: (i) we derive theoretically the conditions for hard and soft bubbling transitions for three generic types of bubbling bifurcations in terms of the coefficients of the canonical forms; (ii) we derive theoretically the scaling of the maximum burst amplitude and average interburst time with the normal parameter, and, in

invariant manifold, while the transcritical bifurcation is generic for systems without symmetry. A saddle-node bifurcation is not possible, because it would destroy the periodic orbit, whereas the normal parameter does not affect the dynamics in the invariant manifold.

the case of interburst time, the size of perturbations transverse to the invariant manifold; (iii) we verify our predicted scalings with the results of the numerical experiments. We present analyses of both mismatch induced and noise induced bubbling, but we pay particular attention to the case of noise-induced bubbling, where we use the Fokker-Planck diffusion approximation to obtain the interburst time scaling results. Our derivations are based on the model systems, where we use canonical forms of the bifurcations to represent the transverse dynamics.

As a specific example of a system that our analysis might be applied to, consider the general case of synchronization of two coupled oscillators, as described by the following system of equations,

$$\frac{d\boldsymbol{z}_1}{dt} = \boldsymbol{F}_1(\boldsymbol{z}_1) + k\boldsymbol{f_1}(\boldsymbol{z}_1 - \boldsymbol{z}_2), \tag{2.1a}$$

$$\frac{d\boldsymbol{z}_2}{dt} = \boldsymbol{F}_2(\boldsymbol{z}_1) + k\boldsymbol{f_2}(\boldsymbol{z}_2 - \boldsymbol{z}_1), \tag{2.1b}$$

where $\boldsymbol{f}_{1,2}$ and $\boldsymbol{F}_{1,2}$ are smooth functions, $\boldsymbol{f}_1(\boldsymbol{0}) = \boldsymbol{f}_2(\boldsymbol{0}) = \boldsymbol{0}$, and k is a coupling constant. For this situation k is the normal parameter in the system. First consider the case where the oscillators are identical, $\boldsymbol{F}_1(\boldsymbol{z}) = \boldsymbol{F}_2(\boldsymbol{z}) = \boldsymbol{F}(\boldsymbol{z})$. The synchronized state $\boldsymbol{z}_1 = \boldsymbol{z}_2$ represents an invariant surface embedded in the full $(\boldsymbol{z}_1, \boldsymbol{z}_2)$ phase space. Let $\boldsymbol{x} = (\boldsymbol{z}_1 + \boldsymbol{z}_2)/2$ and $\boldsymbol{y} = (\boldsymbol{z}_1 - \boldsymbol{z}_2)/2$. In these coordinates $(\boldsymbol{x}, \boldsymbol{y})$ the dynamics in $\boldsymbol{x}$ with $\boldsymbol{y} = 0$ represents the dynamics along the invariant manifold. We ask, what is the effect of small perturbations to this system caused by noise or mismatch on the dynamics in the $\boldsymbol{y}$ direction (i.e., transverse to the invariant manifold)? In particular, what is the effect of a small noise of order r added to the right hand sides of Eqs. (2.1) or of a small deviation such that the two oscillators are not identical (mismatch), $||\boldsymbol{F}_1 - \boldsymbol{F}_2|| \sim q << 1$? These perturbations will typically destroy the invariance of the invariant manifold $\boldsymbol{z}_1 = \boldsymbol{z}_2$. In what follows we consider discrete time systems, possibly obtained

via a surface of section from a continuous time system [e.g., Eqs. (2.1)]. If the synchronized state $z_1 = z_2$ without noise or mismatch is chaotic, we can model the dynamics using the following model systems.

2.2   Model systems

To simplify our analysis we assume that the chaotic dynamics on the invariant manifold is given by $x_{n+1} = 2x_n \mod 1$ and that the periodic orbit that becomes transversely unstable at the bubbling bifurcation is the fixed point $x = 0$. More generally, the bifurcation may occur at a higher period orbit, but this orbit will typically have low period (Pecora and Carroll 1990). Our results below depend on the dynamics within and transverse to the invariant manifold in the following ways. Within the invariant manifold, the results depend only on the largest Lyapunov exponent of the bifurcating orbit, which we denote by $h_{||}$. Transverse to the invariant manifold, our results depend on the local dynamics in the "center manifold" corresponding to the eigenvalue(s) on the unit circle at the bifurcation, which we represent by a complex variable z in the case of a Hopf bifurcation and by a real variable y for the other bifurcations. We include only the quadratic and cubic terms in y and z that are necessary to determine the *normal form* for the bifurcation – that is, to describe the bifurcation to the lowest order. For details on center manifolds and normal forms, see (Hale and Kocak 1991).

Model for Pitchfork and Transcritical Bifurcations: For the case of a transverse pitchfork or transcritical bifurcation we consider a model system of the form

(Venkataramani et al. 1996):

$$x_{n+1} = 2x_n \mod 1, \tag{2.2a}$$

$$y_{n+1} = [\cos(2\pi x_n) + p]y_n + ay_n^\sigma + q + r\nu_n, \text{ (for } |y| < 1), \tag{2.2b}$$

where x_n and y_n are real, p is the bifurcation parameter with $p > 0$ ($p < 0$) above (below) the bubbling transition and q as the mismatch parameter. The $r\nu_n$ term represents noise in the system with magnitude $r > 0$. We assume that ν_n are random numbers uniformly distributed on $[-1, 1]$. The term ay^σ with $\sigma = 2$ or 3 in (2.2b) represents the lowest order y nonlinearity of the system at the fixed point $x = 0$. We assume that $|p| << 1$, $|q| << 1$ and $|r| << 1$ but $|a| = O(1)$. The dynamics in x models the chaotic dynamics in the invariant manifold and the dynamics in y models the dynamics transverse to the invariant manifold. In this and the following models the equation for the y dynamics models the local evolution of the system close to the invariant manifold. For $|y| > 1$, it is presumed that (2.2b) does not apply and that there is a confining nonlinearity that sends the orbit back to the region $|y| < 1$ (in particular, there is no attractor in $|y| > 1$). Considering $q = 0$ and $r = 0$, the linearized y dynamics at the fixed point $x = 0$, is governed by $y_{n+1} = (1 + p)y_n$; thus, as p increases through zero, dy_{n+1}/dy_n increases through $+1$, corresponding to a pitchfork or transcritical bifurcation in the transverse dynamics at $x = 0$ fixed point. Symmetric coupling ($\boldsymbol{f}_1 = \boldsymbol{f}_2$) in (2.1) is modeled by $\sigma = 3$, and asymmetric coupling ($\boldsymbol{f}_1 \neq \boldsymbol{f}_2$) is modeled by $\sigma = 2$. For the symmetric case, the symmetry $y \to -y$ rules out the possibility of a y^2 term in (2.2b). The term q represents a small mismatch $\boldsymbol{F}_1 - \boldsymbol{F}_2$. In the absence of noise and mismatch we have an invariant line, $y = 0$, on which there exists a chaotic invariant set generated by (2.1a). The stability of this line is governed by the coefficient of y_n in the first term on the right hand side of (2.3b).

Since $\cos(2\pi x_n)$ is maximum at $x = 0$, the period one orbit at $(x, y) = (0, 0)$ with $q = 0$ and $r = 0$ becomes transversely unstable as p increases through zero, and the corresponding bifurcation is a transcritical bifurcation if $\sigma = 2$ (asymmetric coupling) or a pitchfork bifurcation if $\sigma = 3$ (symmetric coupling). In terms of Eqs.(2.1), p is analogous to $(k_c - k)$, where k_c is the critical bubbling value of the coupling strength k. We refer to $p = 0$ as the critical parameter value. Because the local structure of a pitchfork or transcritical bifurcation (e.g. subcritical or supercritical) is determined by the lowest order nonzero nonlinear term, we neglect all terms of order higher than y^σ. Note that the chaos in the invariant manifold $(y = 0)$ is unaffected by p (Eq. (2.2a) is independent of p).

Model for Period Doubling Bifurcations: For the period-doubling case we consider a model system of the form:

$$x_{n+1} = 2x_n \quad \text{mod } 1, \tag{2.3a}$$

$$y_{n+1} = -[\cos(2\pi x_n) + p]y_n + ay_n^2 + by_n^3 + q + r\nu_n, \quad \text{(for } |y| < 1\text{)}, \tag{2.3b}$$

where x_n and y_n are real, p is the bifurcation parameter with $p > 0$ $(p < 0)$ above (below) the bubbling transition, and a, b, q and $r > 0$ are parameters of the system whose values define the type of transition occurring as p goes through 0. Again we assume that $|p|, |q|, |r| << 1$ and $\max(|a|, |b|) = O(1)$. The ν_n are random numbers uniformly distributed on $[-1, 1]$. Considering $q = 0$ and $r = 0$, the linearized y dynamics at the $x = 0$ fixed point of (2.3a) is $y_{n+1} = -(1 + p)y_n$ which becomes unstable as p increases through zero with dy_{n+1}/dy_n decreasing through -1, corresponding to a period doubling bifurcation. In the case of coupled oscillators (2.1) with symmetric coupling $(\boldsymbol{f}_1 = \boldsymbol{f}_2)$, $a = 0$, but with asymmetric coupling $(\boldsymbol{f}_1 \neq \boldsymbol{f}_2)$, both quadratic and cubic terms may be present. Because the local structure of a period-doubling bifurcation is

determined by the terms of order up to 3, we neglect all terms of order higher than y^3, and, since both the quadratic and cubic terms are important, we do not treat the cases of symmetric and asymmetric coupling separately.

Model for Hopf Bifurcations: In the case where the transverse bifurcation of the periodic orbit is a Hopf bifurcation, we employ the following model:

$$x_{n+1} = 2x_n \mod 1, \tag{2.4a}$$

$$z_{n+1} = [\cos(2\pi x_n) + p]e^{i\theta} z_n + az_n^2 + bz_n z_n^* + c(z_n^*)^2 + d|z_n|^2 z_n + \tag{2.4b}$$

$$+ q + r\nu_n, \text{ (for } |z| < 1),$$

where z_n is complex and z_n^* is the complex conjugate of z_n. As in the previous cases, p and r are small real parameters. The quantities a, b, c, d, q, and r are complex parameters of the system with $|q|, |r| << 1$ and $|\max(a, b, c, d)| = O(1)$. The ν_n are i.i.d. complex random numbers uniformly distributed within the unit circle. In this model complex z models the 2-dimensional dynamics transverse to the invariant manifold, while the variable x models the dynamics along the invariant manifold. For $q = 0$ and $r = 0$, the quantity dz_{n+1}/dz_n evaluated at the fixed point $x = 0$ has its magnitude increase through unity as p increases through zero. For $\theta/2\pi$ irrational this corresponds to a Hopf bifurcation. This model includes the above mentioned pitchfork or transcritical and period doubling bubbling bifurcations as special cases with $\theta = 0$ being the pitchfork or transcritical case and $\theta = \pi$ being the period-doubling case. Although it is possible to get all results for period-doubling and pitchfork or transcritical bifurcations from the model (2.4), we will use the models (2.2) and (2.3) for these special cases to simplify the analysis. The mismatch parameter q again breaks the invariance of the line $z = 0 + 0i$. We included all possible terms quadratic in z. In previous work on the Hopf bifurcation it has been shown that out of all

13

possible terms cubic in z only the term proportional to $|z_n|^2 z_n$ is relevant to the local structure of the bifurcation (Wan 1978).

The results we obtain for models (2.2), (2.3), (2.4) are given in Tables 2.1-2.3. We treat the cases of mismatch ($q \neq 0$) and noise ($r \neq 0$) separately; when both q and r are nonzero, the average interburst time will scale like the smaller of the two expressions given. We claim that these results can be applied to any generic situation exhibiting a bubbling transition. For the case of the pitchfork or transcritical bifurcation this claim has been confirmed experimentally (see (Venkataramani et al. 1996)).

2.3 Pitchfork and transcritical bifurcations

2.3.1 Maximum burst amplitude and stability

In this subsection we review the derivation of the theoretical result for the maximum burst amplitude for the case of a pitchfork or transcritical bubbling bifurcation in the map (2.2). We derive the results for the noiseless case $r = 0$ first. This derivation serves as a model for the treatments of the period doubling and Hopf cases, and closely follows (Venkataramani et al. 1996). In the case of symmetric coupling ($\sigma = 3$) we will show that we have a soft bubbling transition if $a < 0$ and a hard bubbling transition if $a > 0$. In the case of asymmetric coupling ($\sigma = 2$), we will show that $qa > 0$ corresponds to a hard transition, and $qa < 0$ corresponds to a soft transition.

For small positive p, the factor $[p + \cos(2\pi x_n)]$ is greater than one only in a small region near the fixed point $x = 0$ of (2.1a) (since x is taken modulo 1, we consider values of x near 1 to be near 0). Thus a burst can only be initiated

when the chaotic x orbit comes near enough to $x = 0$ that it remains there long enough for y to burst. The burst ends when x moves away from the fixed point. In order to compute the maximum possible burst amplitude, we first consider the dynamics when $x = 0$. Then y satisfies the relation

$$y_{n+1} - y_n = py_n + ay_n^\sigma + q. \tag{2.5}$$

Assume for definiteness that $q > 0$. Then, in the case of a soft transition $(a < 0)$, if $x = 0$ then y increases but is bounded from above by the positive solution of

$$p\Delta + a\Delta^\sigma + q = 0. \tag{2.6}$$

Since the maximum value that y can reach is Δ, this value represents the maximum burst amplitude for all trajectories that start near $y = 0$. Since $a\Delta^\sigma$ is the only negative term on the left-hand side of (2.6) , the solution for Δ can be estimated by $\Delta^\sigma \simeq \max(p\Delta/|a|, |q/a|)$ which yields:

$$\Delta \simeq \max((p/|a|)^{1/(\sigma-1)}, |q/a|^{1/\sigma}). \tag{2.7}$$

(In this and further equations we use $\simeq$ to denote 'approximately equal' and $\sim$ to denote 'equal up to a constant'.) In particular, when $p >> q^{(\sigma-1)/\sigma}$, we have $\Delta \simeq |p/a|^{1/(\sigma-1)}$.

When $q < 0$, soft transition will occur if and only if (2.6) has a negative root Δ. Thus if $\sigma = 2$ we need $a > 0$ for a soft transition, while if $\sigma = 3$ we need $a < 0$. In these cases, the magnitude of the negative root Δ is given by the right side of (2.7). In the case of noise, for asymmetric coupling $(\sigma = 2)$ the transition is always hard; because the noise can take either sign, there will be both large and small bursts. For symmetric coupling $(\sigma = 3)$, we still have a soft transition if

$a < 0$ and a hard transition if $a > 0$. In the case of a soft transition, by arguments similar to the one above, we have $\Delta \simeq (p/|a|)^{1/(\sigma-1)}$ when $r^{2/3} << p << 1$.

2.3.2 Average interburst time - mismatch

In what follows we derive results for the average interburst time for soft and hard bubbling transitions. We first consider the effect of mismatch in the noiseless case $r = 0$. The analysis in the previous section shows that the nonlinear term ay^σ is insignificant when $|y| << y_c \sim \max((p/|a|)^{1/(\sigma-1)}, (q/|a|)^{1/\sigma})$, while when y grows close to y_c, the nonlinear term will either confine the burst (soft transition) or rapidly accelerate the orbit to $y = O(1)$ (hard transition). Either way, we can estimate the interburst time as the time for $|y|$ to reach y_c in the absence of the nonlinear term.

Assuming that x stays at its fixed point $x = 0$, the n-th iterate y_n for $n > 0$ in the absence of the nonlinearity can be written as

$$y_n = \sum_{i=0}^{n-1} (1+p)^i q = \frac{((1+p)^n - 1)q}{p} \tag{2.8}$$

Hence we can compute the time $\bar{n}$ for an initial point $(x_0, y_0) = (0, 0)$ to reach y_c by setting $y_{\bar{n}} = y_c$ which yields

$$\bar{n} \sim \frac{1}{p} \ln\left(\frac{py_c}{|q|}\right), \tag{2.9}$$

This expression is valid for $py_c >> q$, which corresponds to the case $p >> q^{(\sigma-1)/\sigma}$ and $y_c \sim (p/|a|)^{1/(\sigma-1)}$. To estimate the average interburst time τ we note that, in order to initiate a burst, an orbit must come within ε of $x = 0$ where ε is sufficiently small that the orbit remains near $x = 0$ for at least $\bar{n}$ iterates. Since the invariant density generated by (2.3) is uniform in x, we have that the average time τ between bursts is given by $\tau^{-1} = \varepsilon$. We express ϵ in terms of $\bar{n}$ as follows.

16

For small initial x_0 near $x = 0$, the subsequent iterates grow exponentially as $x_0 \exp(h_{||} n)$ where $h_{||} = \ln 2$ is the Lyapunov exponent of (2.1a). Requiring that $\varepsilon \exp(h_{||} \bar{n}) \lesssim \delta$, where $\delta < 1$ is $O(1)$, we obtain the desired estimate of ε and hence of the average interburst time in terms of $\bar{n}$

$$\ln \tau \sim \ln \varepsilon^{-1} = h_{||} \bar{n}. \tag{2.10}$$

Substituting $\bar{n}$ from (2.9) into (2.10) we obtain the result for the scaling of $\ln \tau$:

$$\ln \tau \sim \frac{h_{||}}{p} \ln \left(\frac{p y_c}{|q|} \right). \tag{2.11}$$

This result is recorded in Table 1 with the appropriate value of y_c substituted.

2.3.3   Average interburst time - noise

We now derive the scaling for the average interburst time in the presence of small bounded noise in the map (2.2). We isolate the effect of noise by taking $q = 0$. We consider the y dynamics in this case to be a drift-diffusion problem with drift proportional to terms linear in y, i.e. py, and diffusion due to noise $r\nu_n$. We split the problem into two parts: first we consider *drift-dominated* bubbling corresponding to the case where the effect of the py drift is dominant in developing a burst, and then we consider *noise-dominated bursting* corresponding to the case where the effect of the noise term $r\nu_n$ is dominant. For both cases we will derive asymptotic upper bounds on the mean interburst time, and our final result for the interburst time will be the minimum of these two bounds. We also derive a relation between p and r that determines what kind of bursting is dominant and thus which scaling applies.

First we consider noise-dominated bursting. In this case Eq. (2.2b) with $x = 0$ can be approximated as a random walk process with small drift. We

characterize this process with two quantities: the drift per iterate py_n and the diffusion coefficient $D = (1/2)r^2 Var(\nu)$. Here $Var(\nu)$ is the variance of the random variable ν_n (mean value of $\nu_n{}^2$). For our numerical experiments ν_n is uniformly distributed in $[-1, 1]$, in which case $Var(\nu) = 1/3$.

As in the previous section, we assume that there is a critical value y_c such that we can estimate the time to produce a burst as the time for $|y|$ to reach y_c in the absence of the nonlinear term ay^σ. We set y_c such that the size of the nonlinear term is equal to the typical size of the noise, $|a|y_c^\sigma \sim \sqrt{D}$. Hence we have

$$y_c \sim \left| \frac{\sqrt{D}}{a} \right|^{1/\sigma}. \qquad (2.12)$$

By defining y_c in this manner, we ensure that the nonlinear term dominates the noise for $|y| >> y_c$. However, it is possible that the nonlinear term becomes significant for $|y| << y_c$ because it behaves coherently from one iterate to the next while the noise term may not. Thus by ignoring the nonlinear term for $|y| < y_c$ we may be overestimating the interburst time. However, our estimate remains an upper bound on the interburst time, and our simulations show that this upper bound correctly describes the actual interburst time scaling in the noise-dominated case. The reason for this is that a burst most often occurs when the noise does behave coherently; we discuss this point further in Section 2.6.

The average interburst time τ is the inverse of the probability per unit time of initiating a burst. By initiating a burst, we mean that x maps close to 0 (having not been close on the previous iteration), and that a burst will happen during the time x remains close to 0. in the previous section, we could say exactly how many iterations $(\bar{n})$ x needed to remain close to 0 in order for a burst to occur, but in the noise-dominated case we cannot. Instead, we proceed as follows. Let

18

$Q(n)$ be the probability that $|y|$ has remained in the range $|y| < y_c$ up to time n. The probability that $|y|$ exceeds y_c for some time at or before time n is $1 - Q(n)$. As in the previous section, probability that x will map close enough to 0 to stay there for at least n iterations is proportional to $e^{-h_{||}n}$. Thus the probability that x remains close to zero for exactly n iterations is proportional to $e^{-h_{||}n} - e^{-h_{||}(n+1)}$. Hence $1/\tau$, the probability per unit time of initiating a burst, satisfies

$$\frac{1}{\tau} \simeq \sum_{n=0}^{\infty} \left[(e^{-h_{||}n} - e^{-h_{||}(n+1)})[1 - Q(n)] \right] \tag{2.13}$$

To estimate (2.13) we utilize a continuous time approximation for the y dynamics, with the continuous variable t replacing n. Equation (2.13) becomes

$$\frac{1}{\tau} \simeq h_{||} \int_0^{\infty} e^{-h_{||}t}[1 - Q(t)]dt = 1 - h_{||} \int_0^{\infty} e^{-h_{||}t} Q(t)dt. \tag{2.14}$$

The continuous time approximation of (2.13) requires that $h_{||}$ be small. We note, however, that (2.14) is valid as an order of magnitude estimate even when $h_{||}$ is of order one. Since $\tau >> 1$, and since we will use (2.14) only to estimate the logarithm of τ, an order of magnitude estimate is sufficient. To estimate $Q(t)$, we consider the time evolution of the probability distribution function for y, $P(y, t)$ for the situation in which an orbit starts at $y = 0$ at time $t = 0$ and is considered to burst when $|y|$ exceeds y_c. Accordingly, we assume that $P(y, 0) = \delta(y)$ and that $P(y, t) = 0$ for $|y| \geq y_c$, so that $P(y, t)$ for $|y| < y_c$ represent the probability distribution function for trajectories that have not yet reached $|y| = y_c$ at time t. Thus

$$Q(t) = \int_{-y_c}^{y_c} P(y, t)dy, \tag{2.15}$$

and

$$\ln \tau \sim -\ln \left\{ 1 - h_{||} \int_{-y_c}^{y_c} \int_0^{\infty} e^{-h_{||}t} P(y, t)dtdy \right\}. \tag{2.16}$$

19

We obtain the probability distribution function $P(y, t)$ using the Fokker-Planck diffusion approximation. Ignoring the nonlinear terms ay^σ, the evolution of the probability distribution function $P(y, t)$ is given by the solution of the following drift-diffusion equation:

$$D\frac{\partial^2 P}{\partial y^2} = \frac{\partial P}{\partial t} + \frac{\partial}{\partial y}(pyP), \tag{2.17}$$

where py and D are the above mentioned drift velocity and average diffusion per iterate parameters. Recall that the initial condition is $P(y, 0) = \delta(y)$ and the boundary conditions are $P(\pm y_c, t) = 0$. We solve this equation by first performing a Laplace transform with respect to the time variable t

$$\bar{P}(y, s) = \int_0^\infty e^{-st} P(y, t) dt. \tag{2.18}$$

Note that this integral is the same as the integration over dt on the right hand side of (2.16) with s replaced by $h_\parallel$. Thus

$$\ln \tau \sim -\ln\left\{1 - h_\parallel \int_{-y_c}^{y_c} \bar{P}(y, h_\parallel) dy\right\}. \tag{2.19}$$

The differential equation for $\bar{P}(y, s)$ is:

$$D\frac{\partial^2 \bar{P}}{\partial y^2} - py\frac{\partial \bar{P}}{\partial y} - (p + s)\bar{P} = -\delta(y), \tag{2.20}$$

with boundary conditions $\bar{P}(\pm y_c, s) = 0$. The exact solution of this equation satisfying the boundary conditions can be expressed in terms of parabolic cylinder functions (Abramowitz and Stegun 1965), and is rather cumbersome. For small p we have developed a perturbation expansion approach that gives the lowest order in p behavior. We first find the solution to (2.20) for $p = 0$ and call it $\bar{P}_0(y, s)$. Then we represent $\bar{P}(y, s)$ in the form $\bar{P}(y, s) = \bar{P}_0(y, s) + p\bar{P}_1(y, s) + O(p^2)$ and substitute this form into Eq.(2.20) and solve for $\bar{P}_1(y, s)$ subject to $\bar{P}_1(\pm y_c, s) = 0$.

20

Thus we obtain the first order in p correction to the solution (see Appendix). Writing τ in terms of $\bar{P}(y, s)$ we obtain:

$$\ln \tau \sim -\ln \left\{ 1 - h_{||} \int_{-y_c}^{y_c} [\bar{P}_0(y, h_{||}) + p\bar{P}_1(y, h_{||})]dy \right\} + O(1). \tag{2.21}$$

Performing the integral above, and making the appropriate approximations (see Appendix), we obtain the scaling of $\ln \tau$ with p and D:

$$\ln \tau \sim \sqrt{\frac{h_{||}}{D}} y_c - \frac{p y_c^2}{4D}. \tag{2.22}$$

This scaling is valid as long as the second term is small compared to the first one; that is, when $p^{\sigma/(\sigma-1)}/|r| << 1$. Upon substitution of the expression (2.12) for y_c into (2.22), we obtain our final expression for the scaling of $\ln \tau$:

$$\ln \tau \sim \sqrt{\frac{h_{||}}{D}} \left| \frac{\sqrt{D}}{a} \right|^{1/\sigma} - \frac{p}{4D} \left| \frac{\sqrt{D}}{a} \right|^{2/\sigma}. \tag{2.23}$$

This equation is also given in Table 1.

To numerically test the scaling results, we iterated map (2.2) starting at $y = 0$ and a typical (irrational) value of x with ν_n distributed uniformly on $[-1, 1]$, and measured the average interburst time for different values of the parameters p and r. Figure 2.1 compares the derived scaling (solid line) with the results from the numerical experiment (diamonds), where $p = 0$ and r is varied. In Figure 2.2, we vary p keeping r fixed.

We now consider drift-dominated bursting. In this case, we again claim that a burst occurs when y becomes greater than a critical value y_c beyond which the nonlinear term dominates. As an upper bound for the critical value y_c we use the same value for the burst threshold as we used in the case of mismatch, i.e. $y_c = (p/|a|)^{(\sigma-1)^{-1}}$. If $y \sim y_c$ the nonlinear term either confines the orbit (for $\sigma = 3$ and $a < 0$), or else rapidly accelerates the orbit to $y \sim O(1)$. Similar to

21
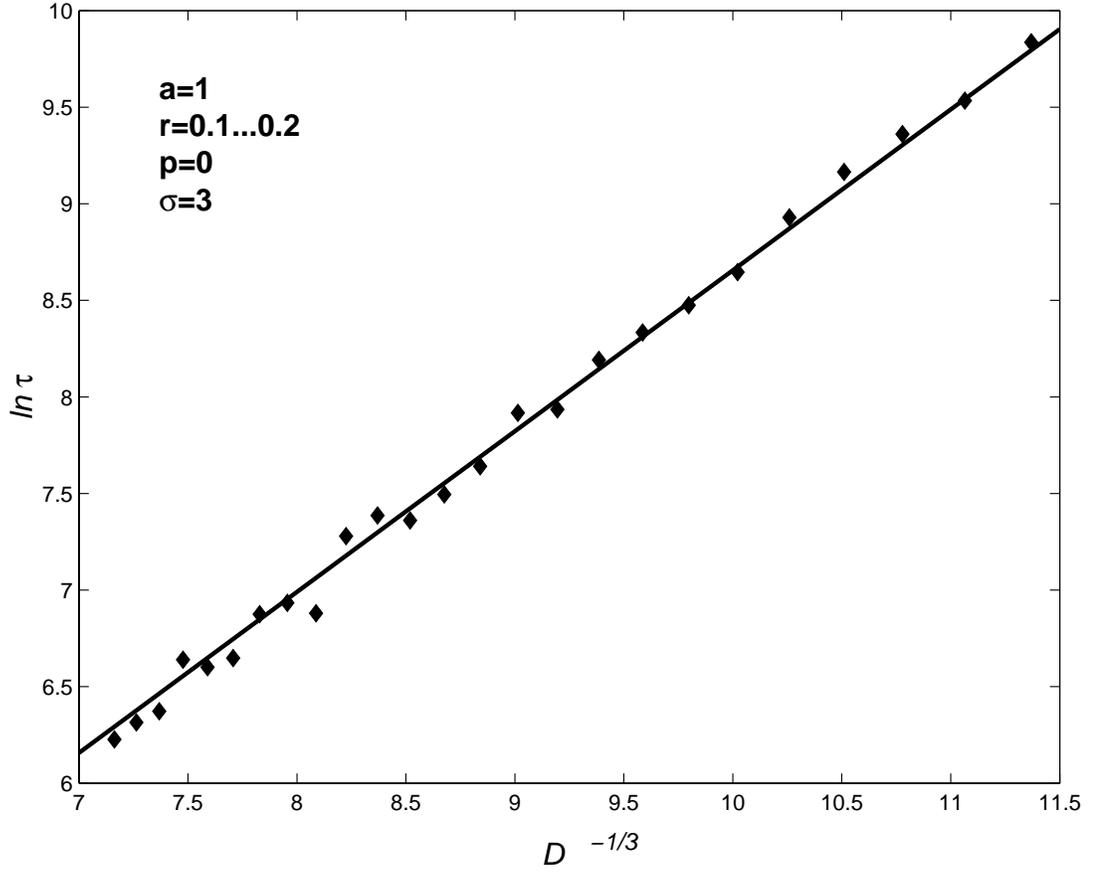
Figure 2.1: This plot shows $\ln\tau$ vs. $D^{-1/3}$ for bubbling induced by a pitchfork bifurcation with noise in noise-dominated case, $p << r^{2/3} << 1$. D is the diffusion coefficient, $D = (1/2)r^2 Var(\nu)$. Parameter values are $a = 1$, $p = 0$, $\sigma = 3$, and $r = 0.1...0.2$. The experimental data are plotted as diamonds. The solid curve has the slope $\sqrt{h_{||}} = \sqrt{\ln 2}$ predicted by scaling given in Table 1.
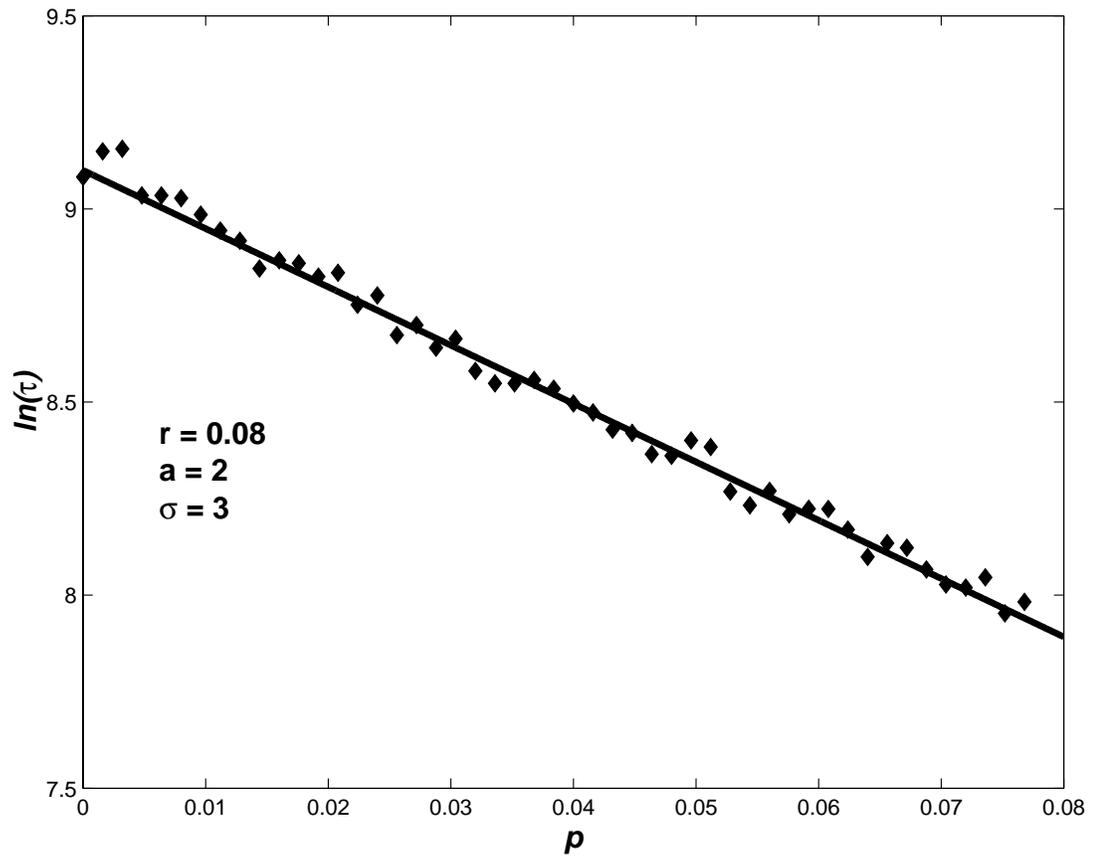
Figure 2.2: This plot shows $\ln \tau$ vs. p for bubbling induced by pitchfork bifurcation with noise in the noise-dominated case, $p << r^{2/3} << 1$. Parameter values are $a = 2$, $r = 0.08$ and $\sigma = 3$. The experimental data are plotted as diamonds. The solid curve has slope -15.1 predicted by scaling given in Table 1.

the mismatch case we first estimate the average number of iterates required for y to reach y_c starting at a value $y_0 = 0$, ignoring the nonlinear term and assuming that x stays at 0 (unlike for noise-dominated bursting, the number of iterates does not depend strongly on whether the noise behaves coherently). The n-th iterate then can be written as:

$$y_n = r \sum_{i=0}^{n-1} (1+p)^i \nu_i. \tag{2.24}$$

Since ν_i are random variables with mean zero, we deduce that

$$Var(y_n) = \frac{2[(1+p)^{2n} - 1]}{p(2+p)} D,$$

where $D = (1/2)r^2 Var(\nu)$ is the previously defined average diffusion per iterate. For $p \ll 1$, we can simplify the expression for the variance $Var(y_n) \simeq \frac{D}{p}[(1 + p)^{2n} - 1]$. Now we set the burst condition to $Var(y_{\bar{n}}) = y_c^2$, where $\bar{n}$ is the average number of iterates of the map required on average for y_n to become equal to or greater than y_c,

$$\frac{D}{p}[(1+p)^{2\bar{n}} - 1] \simeq y_c^2. \tag{2.25}$$

Solving the above equation for $\bar{n}$, and dropping higher order p and d terms, we obtain:

$$\bar{n} \simeq \frac{1}{2p} \ln\left(\frac{py_c^2}{D}\right). \tag{2.26}$$

Notice that if the noise behaved coherently (say $\nu_n = 1$ for all n), the result for $\bar{n}$ would differ only in that p would be replaced by p^2 inside the logarithm [one can see this by replacing $|q|$ with r in (2.9)]. Using $\ln \tau \sim h_{||}\bar{n}$, we obtain:

$$\ln \tau \sim \frac{h_{||}}{2p} \ln\left(\frac{py_c^2}{D}\right). \tag{2.27}$$

This is the final scaling result given in the Table 1.

We have numerically tested our theoretical result for drift-dominated bubbling by iterating the map (2.2) starting at $y = 0$ and a typical (irrational) value of x with ν_n distributed uniformly on $[-1, 1]$, and measuring the average interburst time. The parameter p was varied with the other parameter values set at $a = 2$, $\sigma = 3$ and $r = 0.005$. In Figure 2.3 we plot $\ln \tau$ from the numerical experiments vs. $\bar{n}(p)$ given by Eq. (2.26). The results of the numerical experiments are shown as diamonds. The solid curve has a slope of $h_{\parallel} = \ln 2$ predicted by the scaling given in Table 1.

Now we consider the condition on p and r that determines what kind of bursting prevails, and hence which scaling applies. To do that, we set the two relations, Eq. (2.23) and Eq. (2.27) equal in the lowest significant order:

$$\sqrt{\frac{h_{\parallel}}{D}} \left| \frac{\sqrt{D}}{a} \right|^{1/\sigma} \simeq \frac{h_{\parallel}}{p}. \tag{2.28}$$

We arrive at the conclusion that the two scalings agree when $p^{\sigma/(\sigma-1)}/|r| \simeq O(1)$. Thus noise-dominated bubbling prevails when $p^{\sigma/(\sigma-1)}/|r| << 1$ and the drift-dominated bubbling prevails if $p^{\sigma/(\sigma-1)}/|r| >> 1$. This result is consistent with the order-of-magnitude estimate presented in Section 2.6.

2.4 Period doubling bifurcation

2.4.1 Maximum burst amplitude and stability

In this section we present a derivation of theoretical results and results of numerical experiments for the period doubling bifurcation induced bubbling transition. We start with the case of no noise: $r = 0$ in our model system (2.3). Consider
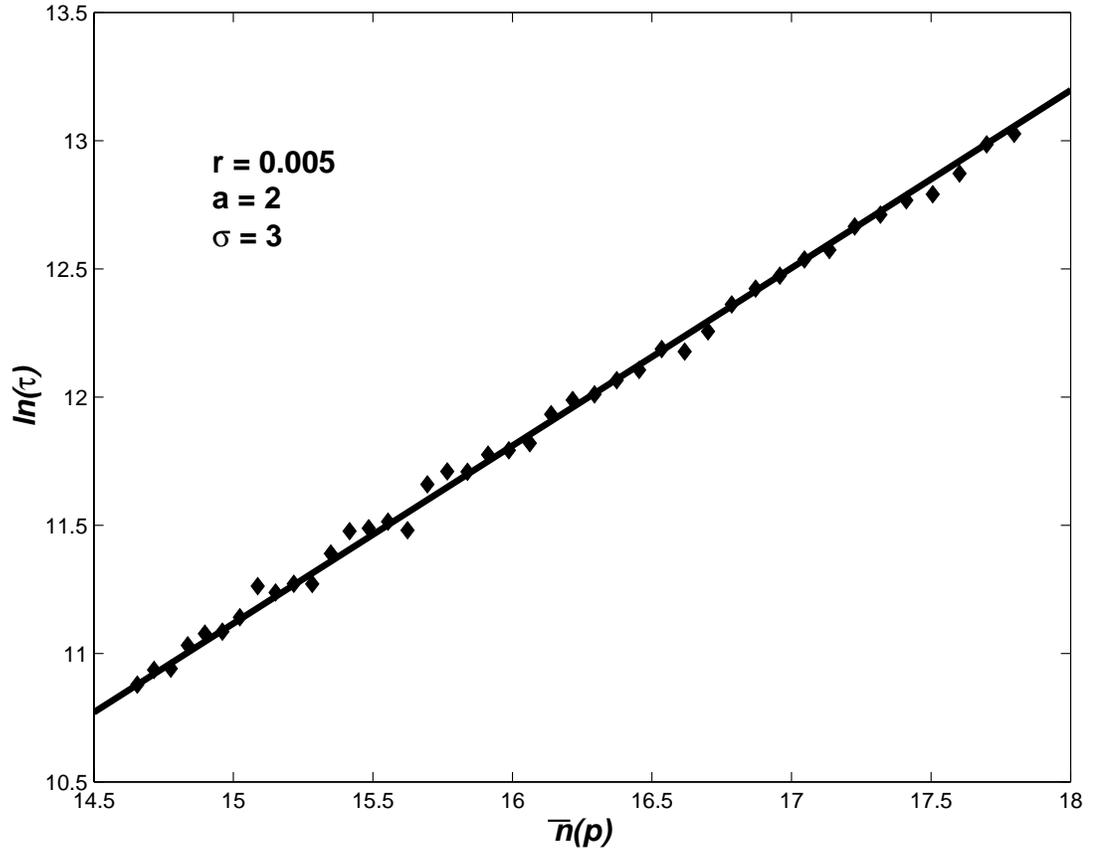
Figure 2.3: This plot shows $\ln \tau$ vs. $\bar{n}(p)$, where $\bar{n}(p)$ is given by (2.26), for bubbling induced by a pitchfork bifurcation with noise in the drift-dominated case, $r^{2/3} \ll p \ll 1$. Parameter values are $a = 2$, $\sigma = 3$, and $r = 0.005$. The experimental data are plotted as diamonds. The solid curve has slope $\ln 2$ predicted by scaling given in Table 1.

Table 2.1: Summary of results for the pitchfork and transcritical bifurcations using the model system (2.2), with $D = (1/2)r^2 Var(\nu)$.

	Symmetric coupling $(\sigma = 3)$	Asymmetric coupling $(\sigma = 2)$
Condition for hard transition	$a > 0$	$qa > 0$ or $r \neq 0$
Condition for soft transition	$a < 0$	$qa < 0$ and $r = 0$
Maximum burst amplitude (soft transition)	$\Delta \simeq \left(\frac{p}{|a|}\right)^{1/2}$ for $\max(|q|,|r|)^{2/3} << p << 1$	$\Delta \simeq \frac{p}{|a|}$ for $q^{1/2} << p << 1$
Average interburst time (mismatch)	$\ln \tau \sim \frac{h_{||}}{p} \ln(\frac{p^{3/2}}{|a|^{1/2}|q|})$ for $|q|^{2/3} << p << 1$	$\ln \tau \sim \frac{h_{||}}{p} \ln(\frac{p^2}{|aq|})$ for $|q|^{1/2} << p << 1$
Average interburst time (noise)	$\ln \tau \sim \frac{h_{||}}{2p} \ln\left(\frac{p^2}{|a|D}\right)$ for $r^{2/3} << p << 1$ $\ln \tau \sim \frac{\sqrt{\frac{h_{||}}{aD}}}{\sqrt[3]{aD}} - \frac{p}{4(aD)^{2/3}}$ for $p << r^{2/3} << 1$	$\ln \tau \sim \frac{h_{||}}{2p} \ln\left(\frac{p^3}{a^2 D}\right)$ for $r^{1/2} << p << 1$ $\ln \tau \sim \frac{\sqrt{\frac{h_{||}}{a\sqrt{D}}}}{\sqrt{a\sqrt{D}}} - \frac{p}{4a D^{1/2}}$ for $p << r^{1/2} << 1$

an orbit starting at $(x, y) = (0, 0)$ for the map (2.2). Note that $(0, 0)$ is a fixed point of the map for $q = 0$. The subsequent iterates obey the relation:

$$y_{n+2} - y_n = 2(p - aq)y_n - 2(a^2 + b)y_n^3 - (p - aq)q + O(py^2, p^2 y, q^2 y, qy^2, q^3, y^4, pqy).$$
(2.29)

Making the change of variables $\hat{p} = p - aq$, we have

$$y_{n+2} - y_n = 2\hat{p}y_n - 2(a^2 + b)y_n^3 - \hat{p}q + O(\hat{p}y^2, \hat{p}^2 y, q^2 y, q^3, qy^2, y^4, \hat{p}qy). \quad (2.30)$$

Of course, if $p >> |aq|$, then $\hat{p} \simeq p$, but for smaller values of p the distinction between $\hat{p}$ and p will be significant. We proceed by analyzing (2.29) in the same way we treated (2.5). Setting $y_{n+2} - y_n = 0$, we obtain the equation for the maximum burst amplitude Δ:

$$2\hat{p}\Delta - 2(a^2 + b)\Delta^3 - \hat{p}q = 0. \quad (2.31)$$

Similar to the derivation of (2.7) in the case $\sigma = 3$, we conclude that for $a^2 + b > 0$,

$$\Delta \simeq \max\left\{ \left(\frac{\hat{p}}{a^2 + b}\right)^{1/2}, \left(\frac{\hat{p}|q|}{a^2 + b}\right)^{1/3} \right\}. \quad (2.32)$$

In particular, if $|q|^2 << \hat{p} << 1$, then $\hat{p}^{1/2} >> (\hat{p}|q|)^{1/3}$ and we have

$$\Delta \simeq \sqrt{\frac{\hat{p}}{a^2 + b}}. \tag{2.33}$$

(A result that accounts for the effect of mismatch more accurately can be obtained by solving Eq. (2.31) for Δ and taking the appropriate root.) Thus for $x = 0$ and $a^2 + b > 0$, the linear exponential increase of y (namely $y_{n+2} - y_n \simeq 2\hat{p}y_n$) is eventually arrested by nonlinearity, and y reaches a maximum, $y = \Delta$, that is small, $O(\hat{p}^{1/2})$, for small $\hat{p}$, corresponding to a soft transition. For $a^2 + b < 0$, Δ does not exist, and, when $|y| \sim |\hat{p}/(a^2 + b)|^{1/2}$ the nonlinearity accelerates the growth of y, leading to a hard transition.

We now obtain the conditions on the parameters that will determine whether the transition is soft or hard. The type of transition is determined by the sign of the expression under the square root in the denominator of (2.33), positive corresponding to a soft transition, and negative corresponding to a hard transition. Thus we have that the transition is hard if $a^2 + b < 0$ and soft if $a^2 + b > 0$. We have tested the above results in numerical experiments on Eqs. (2.3). For $b = -4$ the transition is hard if $-2 < a < 2$. Figure 2.4a illustrates the soft transition if we iterate (2.3) starting at $(x_0, y_0) = (0, 0)$. The data from numerical experiments is plotted as dots. Figure 2.4b shows a bussting time series for a hard bubbling transition.

In the case of noise, we have a soft transition if $(a^2 + b) > 0$ and a hard transition if $(a^2 + b) < 0$. In the case of a soft transition, by arguments similar to the one above, we have $\Delta \simeq \sqrt{\frac{\hat{p}}{a^2+b}}$ when $r^{2/3} << p << 1$.
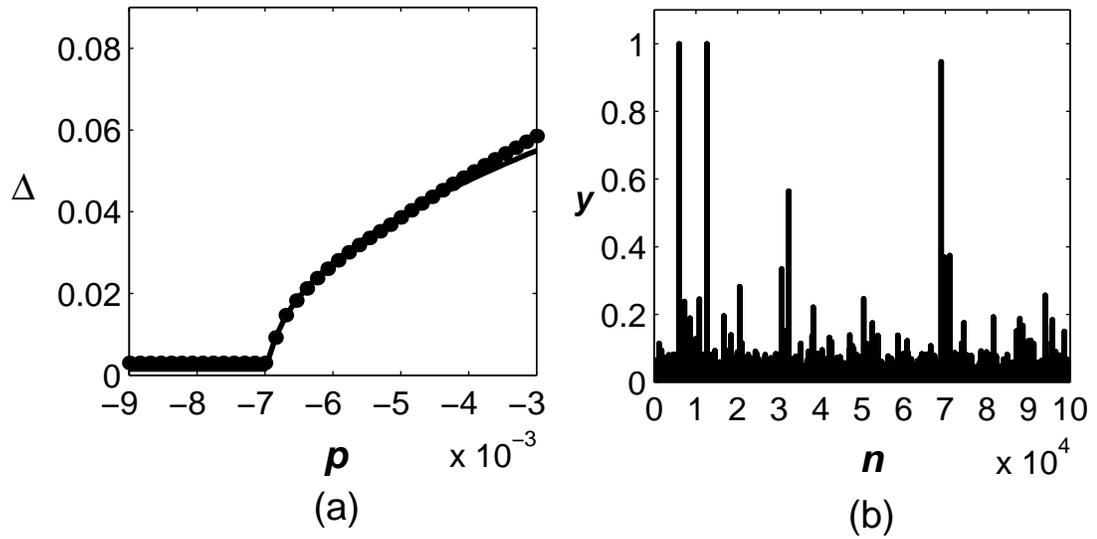
Figure 2.4: Plot (a) shows the maximum burst amplitude Δ vs. p for the soft transition in period-doubling induced bubbling, map (2.3). The experimental data are plotted as dots. The solid curve is the theoretical result from Eq. (2.33). Plot (b) shows the bursting time series for hard transition at $p = 0.27$. Parameter values are $q = 0.003$, $a = -1.9$ (a), $a = -2.3$ (b) and $b = -4$.

2.4.2 Average interburst time - mismatch

To begin the analysis, we first note from (2.3) that the average magnitude of y between bursts is of order q. We again consider the second iterate of the map (2.3b), for $x = 0$ with terms of higher order in p and q dropped, resulting in (2.29).

Eqn. (2.30) shows that in the case of a soft transition the nonlinearity limits the increase of $|y|$ at the value of $\sqrt{\frac{\hat{p}}{|a^2+b|}}$. Denote this value of y a critical value y_c. In the case of a hard transition, the nonlinear terms quickly push y to $|y| \sim O(1)$ as soon as $|y|$ grows to a value of the order y_c. As mentioned earlier, at the beginning of a burst y is of order q. The term $\hat{p}q$ only determines the direction of the burst and therefore it is rather insignificant, being at most the order of the $2\hat{p}y$ term. Thus for simplicity we ignore it and assume that, when x comes close to zero, $y \simeq q$. Then $y_n \cong (1 + 2\hat{p})^{n/2}q$ and setting $y_{\bar{n}} = y_c$ we obtain the following expression for the number of iterates $\bar{n}$ required for $|y|$ to reach y_c (i.e., to initiate a burst):

$$\bar{n} \simeq \frac{1}{\hat{p}} \ln\left(\frac{y_c}{|q|}\right), \tag{2.34}$$

for $|q|^2 << \hat{p} << 1$. From (2.9) and (2.34), we obtain the desired estimate for the average interburst time:

$$\ln \tau \sim h_{||}\bar{n} = \frac{h_{||}}{\hat{p}} \ln\left(\frac{y_c}{|q|}\right), \tag{2.35}$$

which is also given in Table 2.

We have obtained the scaling of τ with p numerically by starting the map (2.3) at a random initial x and $y = 0$ and measuring the average number of iterates that it took for $|y|$ to become greater than 1. For the parameters of Figure 2.5 the transition is hard since $a^2 + b < 0$. The experimental data is plotted as

diamonds. Figure 2.5 presents this data as $\ln \tau$ vs. $\bar{n}(p)$ where $\bar{n}(p)$ is obtained from Eq. (2.34). The solid line has the slope of $h_{||} = \ln 2$ predicted by (2.35) and is consistent with the data.

2.4.3 Average interburst time - noise

In this section we deduce the expression for average interburst time for period doubling induced bubbling. Similar to the previous section we consider every other iterate of (2.3b) in the presence of noise and the absence of mismatch $(q = 0)$ with x at its fixed point $x = 0$:

$$y_{n+2} - y_n \simeq 2py_n - 2(a^2 + b)y_n^3 + r(\nu_{n+1} - \nu_n) \qquad (2.36)$$

We redefine the noise variable $(\nu_{n+1} - \nu_n)$ as $\hat{\nu}_n$, where $Var(\hat{\nu}_n) = 2Var(\nu_n)$. With $\hat{\nu}_n$, Eqn.(2.36) is equivalent to the case of noise in pitchfork bifurcation $(\sigma = 3)$. The variance of $\hat{\nu}_n$ is double the variance of ν_n, but since we are considering every other iterate of y_n, these two effects cancel in the computation of the average interburst time. Thus all results derived in Section 2.3.3 apply, including the scaling ranges, if we use the derived expressions for average interburst time for $\sigma = 3$ with $a^2 + b$ as the coefficient of the cubic term. The results for period-doubling bifurcation induced bubbling are summarized in Table 2.

To test the scaling results we iterated map (2.3) starting at a typical irrational x and $y = 0$ with ν_n distributed uniformly on $[-1, 1]$ and measured the average interburst time for different values of p, keeping the noise magnitude r fixed. Figure 2.6 compares the derived scaling result with the experimental results. We plot $\ln \tau$ for different values of p keeping other parameters fixed. Numerical data is plotted as diamonds.
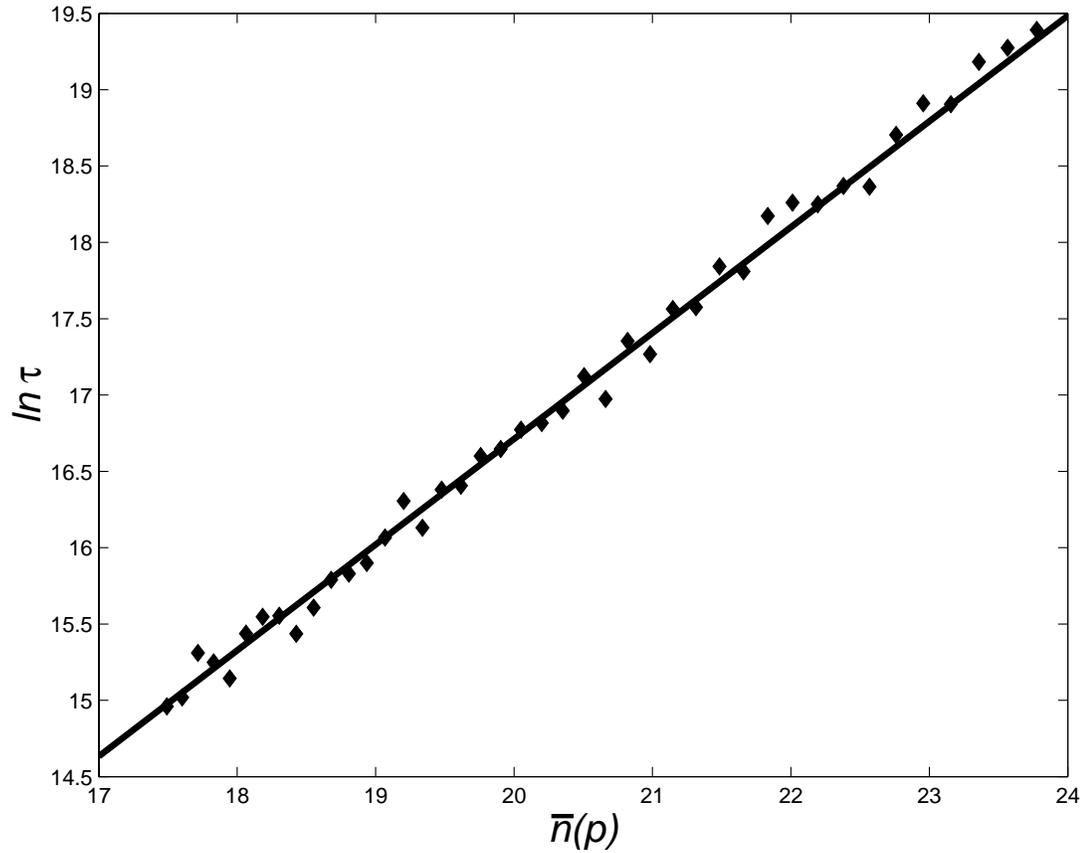
Figure 2.5: This plot shows $\ln \tau$ vs. $\bar{n}(p)$, where $\bar{n}(p)$ is given by (2.34), for bubbling induced by a period doubling bifurcation with asymmetry for $q^2 \ll p \ll 1$. Parameter values are $a = 1$, $b = -2$, $q = 0.008$ with p varying from 0.14 to 0.22. The experimental data are plotted as diamonds. The solid curve has a slope of $\ln 2$ predicted by scaling given in Table 2.
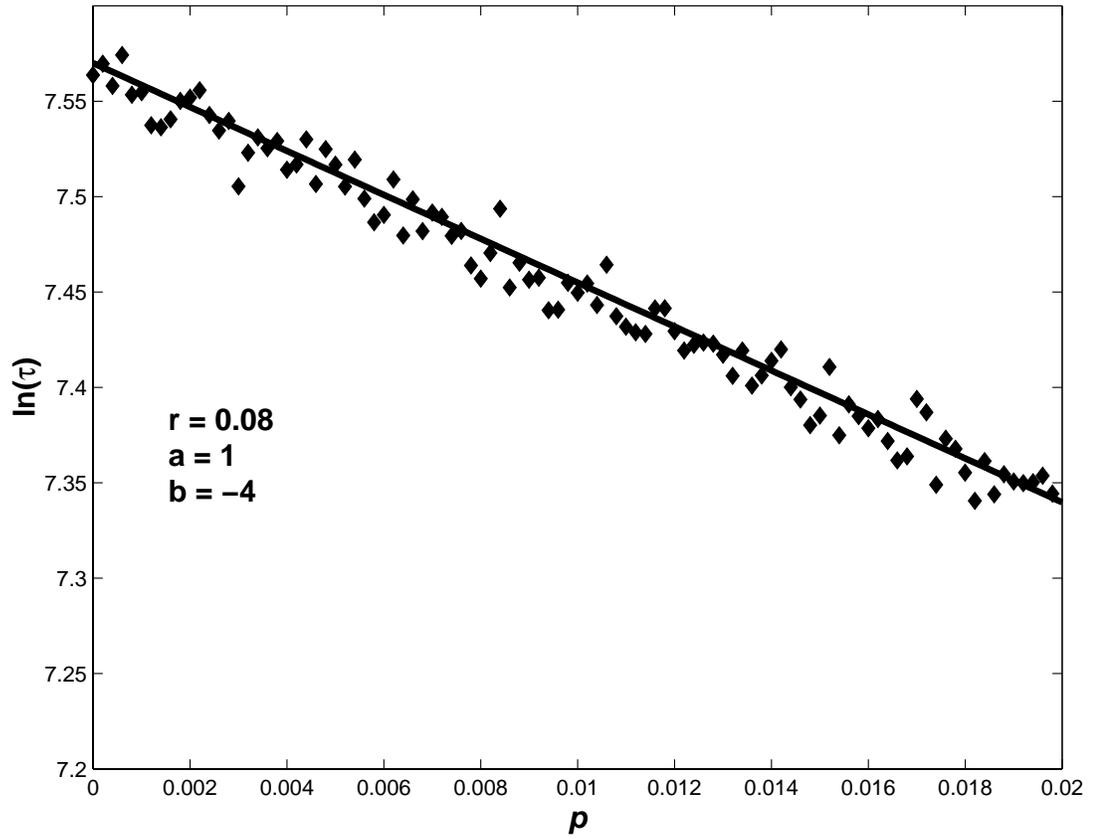
Figure 2.6: This plot shows $\ln \tau$ vs. p for bubbling induced by a period doubling bifurcation with noise in the noise-dominated case, $p << r^{2/3} << 1$. Parameter values are $a = 1$, $b = -4$ and $r = 0.08$. The experimental data are plotted as diamonds. The solid curve has the slope -11.5 predicted by scaling given in Table 2.

Table 2.2: Summary of results for the period doubling bifurcation using the model system (2.3). We use the notation $\hat{p} = p - aq$ and $D = (1/2)r^2 Var(\nu)$.

Condition for hard transition	$a^2 + b < 0$
Condition for soft transition	$a^2 + b > 0$
Maximum burst amplitude (soft transition)	$\Delta \simeq \sqrt{\frac{\hat{p}}{a^2+b}}$
	for $\max(|q|^2, r^{2/3}) << \hat{p} << 1$
Average interburst time (mismatch)	$\ln \tau \sim \frac{h_{\parallel}}{2\hat{p}} \ln \left(\frac{\hat{p}}{|a^2+b|q^2} \right)$
	for $|q|^2 << \hat{p} << 1$
Average interburst time (noise)	$\ln \tau \sim \frac{h_{\parallel}}{2p} \ln \left(\frac{p^2}{|a^2+b|D} \right)$
	for $r^{2/3} << p << 1$
	$\ln \tau \sim \frac{\sqrt{h_{\parallel}}}{\sqrt[3]{(a^2+b)D}} - \frac{p}{4((a^2+b)D)^{2/3}}$
	for $p << r^{2/3} << 1$

2.5 Hopf bifurcation

2.5.1 Maximum burst amplitude and stability

Next we examine the case of a Hopf bifurcation, Eqs. (2.4). We consider an orbit starting at $(x, z) = (0, 0 + 0i)$, and take $p = 0$ (i. e., we consider the map at the critical bifurcation point). For the validity of the analysis below we assume that the angle θ in the exponent in front of the linear coefficient is not equal to certain special values: $0, \pm\pi/2, \pm 2\pi/5, \pm 2\pi/3$ and $\pm\pi$. The cases $\theta = 0$ and $\theta = \pm\pi$ correspond to pitchfork or transcritical and period doubling bifurcations and have already been considered in Sections 2.3 and 2.4. The other non-allowed angles $(\pm\pi/2, \pm 2\pi/5, \pm 2\pi/3)$ correspond to non-generic cases which, unless special circumstances apply, are not expected to occur. When the above special θ values are excluded, it can be shown (Wan 1978; Lanford 1973) that by means of a coordinate transformation of the form

$$z'_n = z_n + \gamma_1 z_n^2 + \gamma_2 z_n z_n^* + \gamma_3 (z_n^*)^2, \tag{2.37}$$

where γ_1, γ_2 and γ_3 are complex numbers, all quadratic terms can be eliminated from (2.4c) with $x = 0$, yielding

$$z'_{n+1} = \lambda z'_n + d'|z'_n|^2 z'_n + q + r\nu_n + O(q^2, r^2, rz', qz', pz'^2, z'^4), \tag{2.38}$$

where $\lambda = (1 + p)e^{i\theta}$ and

$$d' = \frac{1 - 2\lambda}{\lambda(\lambda - 1)}ab + \frac{\lambda}{\lambda - 1}bb^* + \frac{2\lambda}{\lambda^3 - 1}cc^* + d. \tag{2.39}$$

Defining $\tilde{z} = z' + q/(1 - \lambda)$, and substituting into (2.38) cancels q in lowest significant order. Thus, in terms of $\tilde{z}$, Eq. (2.38) becomes:

$$\tilde{z}_{n+1} = \lambda\tilde{z}_n + d'|\tilde{z}_n|^2\tilde{z}_n + r\nu_n + O(q^2, r^2, r\tilde{z}, q\tilde{z}, p\tilde{z}^2, \tilde{z}^4). \tag{2.40}$$

Until we consider noise in Section 2.5.3, we assume that $r = 0$. Then (2.40) can be transformed into the following canonical form:

$$\begin{pmatrix} |\tilde{z}_{n+1}| \\ \arg(\tilde{z}_{n+1}) \end{pmatrix} = \begin{pmatrix} (1+p)|\tilde{z}_n| - f_1|\tilde{z}_n|^3 \\ \arg(\tilde{z}_n) + \Theta + f_2|\tilde{z}_n|^2 \end{pmatrix} + O(q^2, q\tilde{z}, \tilde{z}_n^4) \qquad (2.41)$$

with f_1, Θ, and f_2 being real coefficients. Equation (2.41) shows that the critical issue is the sign of f_1. A positive sign indicates a soft transition and a negative sign indicates a hard transition. As derived in (Wan 1978), $f_1 = -Re(\lambda^* d')$ at the bifurcation value $\lambda = e^{i\theta}$, or

$$f_1 = Re\left[\frac{(1 - 2e^{i\theta})e^{-2i\theta}}{1 - e^{i\theta}}ab\right] + \frac{1}{2}bb^* + cc^* - Re(de^{-i\theta}) \qquad (2.42)$$

The sign of the above expression determines whether the transition is hard ($f_1 < 0$) or soft ($f_1 > 0$).

Next we obtain an expression for the maximum burst amplitude for a soft transition in model (2.4). We assume that f_1 is of order 1, i.e., we are not close to the borderline between hard and soft transitions. Rewriting the radial part of (2.41) as

$$|\tilde{z}_{n+1}| - |\tilde{z}_n| = p|\tilde{z}_n| - f_1|\tilde{z}_n|^3 + O(q^2, q\tilde{z}, \tilde{z}_n^4), \qquad (2.43)$$

we see that, similar to the pitchfork case, the maximum burst amplitude $\tilde{\Delta}$ for $|\tilde{z}|$ is

$$\tilde{\Delta} \simeq \sqrt{\frac{p}{f_1}}. \qquad (2.44)$$

This relation is true if $p\Delta = f_1\Delta^3 >> \max(q^2, q\tilde{\Delta}, \tilde{\Delta}_n^4)$. Thus the scaling range for (2.44) is $|q| << p << 1$, and, since $z = z' + O(z')^2 = \tilde{z} + O(\tilde{z}^2, q)$, in this range we have $\Delta \simeq \tilde{\Delta} \simeq \sqrt{\frac{p}{f_1}}$. Since our final result does not depend on the magnitude of q in the scaling range specified, the result applies in the case of noise-induced bubbling as well. Figure 2.7 shows the scaling of the maximum burst amplitude

with p. Experimental data are plotted as dots. The solid curve is calculated from (2.44).

In the case of noise, we have a soft transition if $f_1 > 0$ and a hard transition if $f_1 < 0$. In the case of a soft transition, by arguments similar to the ones in previous sections, we have $\Delta \simeq \sqrt{\frac{p}{f_1}}$ when $r^{2/3} << p << 1$.

2.5.2   Average interburst time - mismatch

We now derive an expression for the average interburst time in the presence of mismatch ($q > 0$) and no noise ($r = 0$). Again we consider an orbit starting at $(x, z) = (0, 0+0i)$ and use the new coordinates z' to eliminate the quadratic terms. Note that the coefficient in front of the cubic term in the transformed coordinates is given by (2.39). As in the previous derivations (Sections 2.3 and 2.4), we make use of the fact that, in the case of a hard transition, when $|z_n|$ grows to the point where the nonlinear terms become significant, the nonlinearity pushes $|z_n|$ to $O(1)$ rapidly. Consider the linear terms in map (2.38), $z_{n+1} = \lambda z_n + q + O(z_n^3)$. We first find the number of iterates it takes to escape starting at 0. Starting at $z = 0$, the n-th iterate of the linearized map is $z_n = \frac{1-\lambda^{n+1}}{1-\lambda}q$. The nonlinear term becomes significant when after $\bar{n}$ iterates, $|z|$ reaches the critical value $z_c = \sqrt{\frac{p}{|f_1|}}$, thus we have the equation for $\bar{n}$:

$$\left| \frac{1 - \lambda^{\bar{n}+1}}{1 - \lambda} q \right|^2 \simeq \frac{p}{|f_1|}. \tag{2.45}$$

Solving the above equation, we find $\bar{n}$:

$$\bar{n} \simeq \frac{1}{2p} \ln \left| \frac{p(1-\lambda)^2}{|f_1||q|^2} \right|. \tag{2.46}$$

Knowing the number of iterates $\bar{n}$ it takes to escape assuming x_n stays close to the fixed point, we use (2.35), $\ln \tau \sim h_{||}n$, to derive the scaling with p of the

37
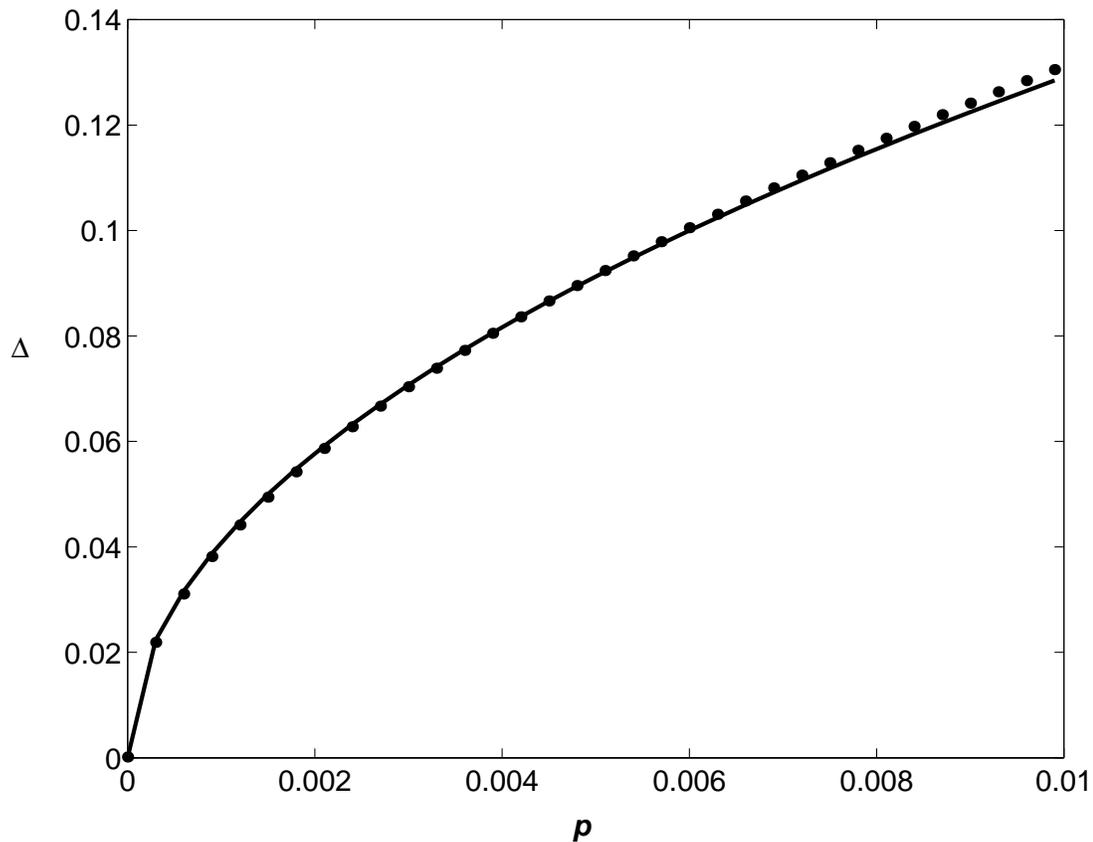
Figure 2.7: This plot shows the maximum burst amplitude Δ vs. p for a soft transition in Hopf bifurcation - induced bubbling. The experimental data are plotted as dots. The solid curve is the theoretical result from Eq. (2.44). Parameter values are $a = 0.1$, $b = 0$, $c = 0.1$, $d = 1.0$, $\theta = \pi\sqrt{5}$, and $q = 0.0001 - 0.0001i$.

average interburst time τ,

$$\ln \tau \sim h_{||} \bar{n} \simeq \frac{h_{||}}{2p} \ln \left| \frac{p(1-\lambda)^2}{|f_1||q|^2} \right|. \tag{2.47}$$

Finally, using the approximation $\lambda = (1+p)e^{i\theta} \simeq e^{i\theta}$, we obtain the final scaling given in Table 3:

$$\ln \tau \sim h_{||} \bar{n} \simeq \frac{h_{||}}{2p} \ln \left| \frac{p(1-e^{i\theta})^2}{|f_1||q|^2} \right|. \tag{2.48}$$

We numerically iterated the map (2.4) starting at a typical irrational x and $z = 0 + 0i$ and measured the interburst time. Figure 2.8 shows the result of numerical experiments (diamonds); $\ln \tau$ is plotted vs. $\bar{n}(p)$, where $\bar{n}(p)$ is given in terms of p by (2.48). The solid line has the predicted slope $h_{||} = \ln 2$ and is consistent with the data.

2.5.3 Average interburst time - noise

We now consider the case where bubbling is induced by noise, i.e. $r > 0$, but $q = 0$. According to (2.41), for small p and in the presence of noise, the evolution of the radial part of z_n with $x = 0$ is the same as the evolution of y in the pitchfork case with cubic nonlinearity with $-f_1$ as the cubic coefficient:

$$|\tilde{z}_{n+1}| \simeq (1+p)|\tilde{z}_n| - f_1|\tilde{z}_n|^3 + r\{\nu_n\}_z. \tag{2.49}$$

The noise term $r\{\nu_n\}_z$ is the projection of noise on the direction of z_n in the complex plane. Since the noise is distributed uniformly within the unit circle, $\{\nu_n\}_z$ has a one-dimensional probability distribution on $[-1, 1]$ with pdf given by $2\sqrt{1 - \{\nu\}_z^2}$. This distribution has variance of $1/4$. We can analyze (2.49) in the same way as in the Section 2.3.3, but with $D = (1/2)r^2 Var(\{\nu_n\}_z)$. The θ dependence disappears from the final scaling due to the fact that the distribution of noise is uniform within the unit circle.
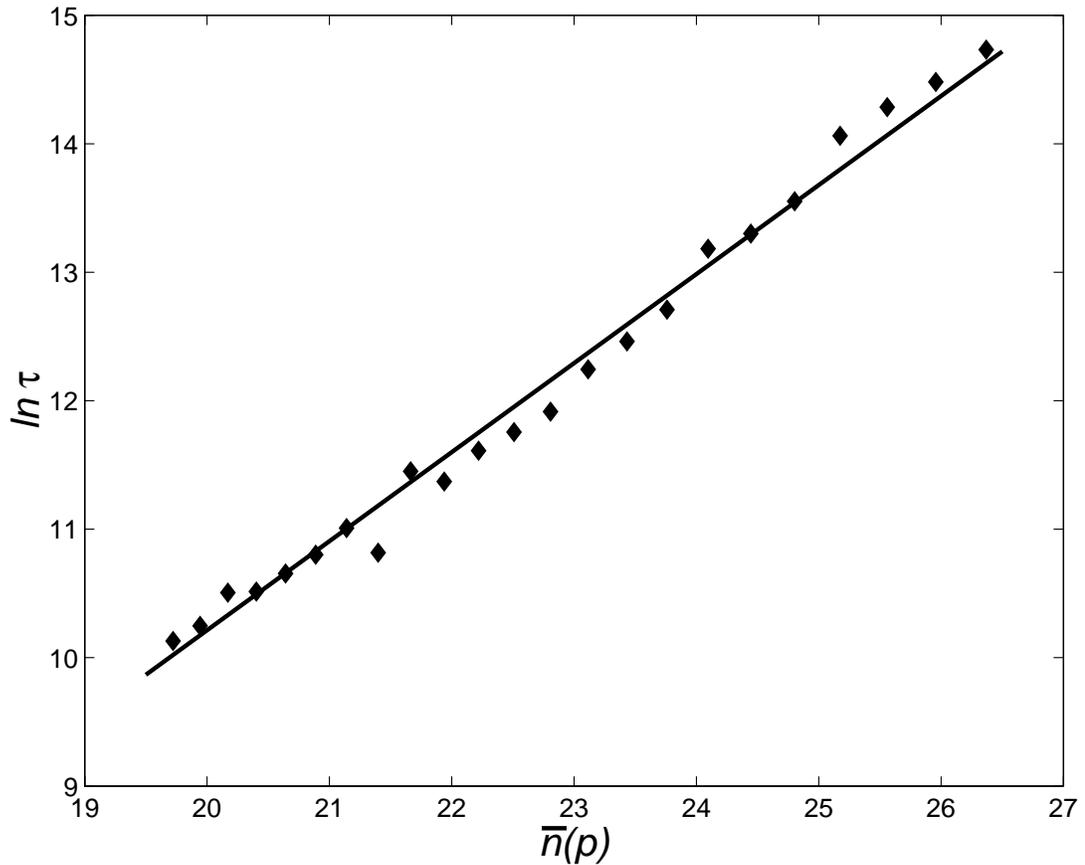
Figure 2.8: This plot shows $\ln \tau$ vs. $\bar{n}(p)$, where $\bar{n}(p)$ is given by (2.46), for bubbling induced by a Hopf bifurcation with asymmetry for $|q| \ll p \ll 1$. Parameter values are $a = 1$, $b = 1$, $c = 1$, $d = 2$, $p = 0.1, ..., 0.15$, $\theta = \pi\sqrt{5}$, and $q = 0.015$. The experimental data are plotted as diamonds. The solid curve has slope $\ln 2$ predicted by (2.48).

Table 2.3: Summary of results for the Hopf bifurcation (f_1 is defined in (2.42), and $D = (1/2)r^2 Var(\nu)$).

Condition for hard transition	$f_1 < 0$
Condition for soft transition	$f_1 > 0$
Maximum burst amplitude (soft transition)	$\Delta \simeq \sqrt{\frac{p}{f_1}}$
	for $\max(|q|, r^{2/3}) << p << 1$
Average interburst time (mismatch)	$\ln \tau \sim \frac{h_\parallel}{2p} \ln \left| \frac{p(1 - e^{i\theta})^2}{|f_1||q|^2} \right|$
	for $|q| << p << 1$
Average interburst time (noise)	$\ln \tau \sim \frac{h_\parallel}{2p} \ln \left(\frac{p^2}{|f_1|D} \right)$
	for $r^{2/3} << p << 1$
	$\ln \tau \sim \frac{\sqrt{h_\parallel}}{\sqrt[3]{f_1 D}} - \frac{p}{4(f_1 D)^{2/3}}$
	for $p << r^{2/3} << 1$

Our final results for Hopf bifurcation induced bubbling transitions are summarized in Table 3.

2.6 Further Discussion of the Noise-Induced Bubbling Mechanism

In this section we provide more insight on the nature of two types of noise-induced bubbling: noise-dominated and drift-dominated. Consider again the map (2.2). We define a critical value y_c such that for $y > y_c$ the nonlinear terms become

dominant and the burst quickly follows. A burst occurs when x comes close to the fixed point $x = 0$ and stays there for a large number of iterates $\bar{n}$. The probability per iterate of that event is of order $\exp(-h_{||}\bar{n})$. After x has entered the required vicinity of the fixed point, the linear coefficient $\cos(2\pi x) + p \simeq 1 + p$ immediately starts driving y away from the invariant manifold. At the time x enters the region near $x = 0$, y will typically be of the order of the size of the noise, $y \sim r$. Thus a burst will occur if $y_c \sim r \exp(\bar{n}p)$, or $\bar{n} \sim p^{-1} \ln(y_c/r) \sim p^{-1}$. We call this scenario drift-dominated bubbling. The probability per iterate of initiating a burst by this mechanism is of order $\exp(-h_{||}/p)$ and goes to zero exponentially as p comes close to the critical value $p = 0$. This would imply that bursts do not happen when $p = 0$, but the experimental results suggest otherwise. Thus we consider another possible route for a burst that becomes important when p sufficiently small. We call this second mechanism noise-dominated bubbling. In the noise-dominated case p is close enough to zero that it can be ignored. In that case a burst may occur if x comes close to the fixed point $x = 0$ and stays there for $\bar{n}$ iterates where $\bar{n}$ is in the range $y_c/r \lesssim \bar{n} \lesssim (y_c/r)^2$. With p neglected, the probability of reaching y_c in $\bar{n}$ iterates is of order $\exp\left[-\left(\frac{y_c^2}{\bar{n}r^2}\right)\right]$. This the probability of a burst in this case is of order $\exp\left[-\left(\frac{y_c^2}{\bar{n}r^2}\right)\right]\exp(-\bar{n}h_{||})$, which is maximized when $\bar{n} \sim y_c/r$. This suggests that when such a burst occurs, the noise behaves *coherently* over $\bar{n}$ iterates pushing y on average in the same direction away from the invariant manifold. Since we determined y_c for a coherent perturbation to be proportional to $r^{1/\sigma}$ (see Eq. (2.12)), the probability of the burst becomes of order $\exp(-h_{||}r^{(1-\sigma)/\sigma})$. Thus the average interburst time τ is of order

$$\tau \sim \min(\exp(h_{||}r^{(1-\sigma)/\sigma}), \exp(h_{||}/p)). \tag{2.50}$$

Eqn.(2.50) suggests that the noise-induced bursting mechanism prevails if $p^{\sigma/(\sigma-1)}/|r| << 1$, while the drift-induced bursting mechanism prevails if $p^{\sigma/(\sigma-1)}/|r| >> 1$.

2.7 Conclusion

The above discussions have assumed that there is no attractor away from the invariant manifold. In the situation where there is an attractor away from the invariant manifold, our analytical results derived in Sections 2.3-2.5 still apply, but the meaning of τ is different. Specifically, for the case that we previously referred to as a hard transition, $q, r = 0$ now yields a riddled basin attractor on the invariant manifold Aswin et al 1994;Ott et al. 1993; Sommerer and Ott 1993; Lai et al. 1996). For $q, r \neq 0$ this attractor is destroyed and converted to a chaotic transient whose mean lifetime is given by τ (Tables 2.1-2.3).

To summarize, in this paper we have presented an unified treatment of the bubbling transitions involving all generic types of bifurcations: pitchfork or transcritical, period doubling and Hopf. The novelty of this paper is in the theoretical derivation of results for scalings of the average interburst time and the maximum burst amplitude with the normal parameter as well as conditions for hard and soft bubbling transitions in the above three cases for both noise and mismatch induced bubbling.

Appendix

In this Appendix we provide details on the solution of Eq.(2.20) subject to the boundary conditions $\bar{P}(\pm y_c, s) = 0$. We look for an approximate solution in terms of a perturbation expansion $\bar{P}(y, s) = \bar{P}_0(y, s) + p\bar{P}_1(y, s) + O(p^2)$. First, we set $p = 0$ in Eq. (2.20) to obtain equation for $\bar{P}_0(y, s)$:

$$D\frac{\partial^2 \bar{P}_0}{\partial y^2} - s\bar{P}_0 = -\delta(y), \tag{2.51}$$

The solution of this equation satisfying the boundary conditions is:

$$\bar{P}_0(y, s) = \frac{\sinh\left(\frac{(y_c - |y|)\sqrt{s}}{\sqrt{D}}\right)}{2\sqrt{Ds}\cosh\left(\frac{y_c\sqrt{s}}{\sqrt{D}}\right)}. \tag{2.52}$$

Now since we know $\bar{P}_0(y, s)$, we can deduce the equation for $\bar{P}_1(y, s)$:

$$D\frac{\partial^2 \bar{P}_1}{\partial y^2} - y\frac{\partial \bar{P}_0}{\partial y} - \bar{P}_0 - s\bar{P}_1 = 0. \tag{2.53}$$

Solving the above equation subject to boundary conditions $\bar{P}_1(\pm y_c, s) = 0$ and $\left(\frac{\partial \bar{P}_1}{\partial y}\right)_{y=0} = 0$, we obtain the expression for $\bar{P}_1(y, s)$:

$$\bar{P}_1(y, s) = \frac{\left[1 - \frac{s}{D}|y|^2 + \sqrt{\frac{s}{D}}y_c \tanh\left(\sqrt{\frac{s}{D}}y_c\right)\right]\sinh\left(\sqrt{\frac{s}{D}}(|y| - y_c)\right) + \sqrt{\frac{s}{D}}(y_c - |y|)\cosh\left(\sqrt{\frac{s}{D}}(|y| - y_c)\right)}{8\cosh(\sqrt{s/D}y_c)}. \tag{2.54}$$

Upon setting $s = h_{||}$ and substitution of the expressions for $\bar{P}_0(y, s)$ and $\bar{P}_1(y, s)$ into the (2.21) and integration over y we obtain

$$\ln\tau \sim -\ln\left[\frac{1 + \frac{py_c^2}{4D} - \frac{py_c}{4\sqrt{h_{||}D}}\tanh\left(\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}\right)}{\cosh\left(\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}\right)}\right]. \tag{2.55}$$

The quantity $\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}$ in (2.55) is large which allows us to make the approximations $\cosh\left(\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}\right) \simeq \exp\left[\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}\right]/2$ and $\tanh\left(\frac{\sqrt{h_{||}}y_c}{\sqrt{D}}\right) \simeq 1$, and neglect the $\frac{y_c}{4\sqrt{h_{||}D}}$ term compared to $\frac{y_c^2}{4D}$ to obtain the final scaling given in Eq. (2.22).

Chapter 3

Extracting Envelopes of Rossby Wave packets

[1]

3.1 Introduction

In this note we consider a situation in which we are given a spatially dependent scalar atmospheric quantity $v(x)$, and that there is a range of wavenumbers $0 < k_{min} \leq k \leq k_{max}$ in which waves of physical interest are known to occur. Using this information, we seek to extract from $v(x)$ a suitable wave-like component of the form,

$$w(x) = A(x) \cos(\phi(x)) \tag{3.1}$$

where $A(x)$, the envelope, is slowly varying in x as compared to the phase $\phi(x)$ (see Figure 1 for an example). Often $A(x)$ will be spatially localized, in which case we refer to it as the *wave packet* envelope.

The desire to locate wave packets in observed atmospheric data is as old

---

[1]This chapter is a verbatim representation of the paper A. V. Zimin, I. Szunyogh, D.J. Patil B. R. Hunt and E. Ott 2003: Extracting Envelopes of Rossby Wave Packets, *Mon. Wea. Rev.*,**131**, 847-853.

as the knowledge that Rossby waves play a key role in shaping the weather in the mid-latitude extra-tropics (Rossby 1945,1949; Yeh 1949; Phillips 1990; Persson 2000). The first technique that was used to detect the propagation of synoptic scale Rossby waves (wavenumber 4-9) was the trough-ridge diagram proposed by Hovmöller (1949). In a Hovmöller diagram a selected atmospheric variable, usually the deviation from the mean of the meridional component of the wind or the geopotential height, is averaged over a latitude band and plotted as a time-longitude diagram. The signature of a propagating wave packet in the Hovmöller diagram is a series of alternating positive and negative regions aligned in a diagonal direction.

While the Hovmöller diagram is a powerful tool, it cannot detect the two-dimensional horizontal structure of the wave packets. This can lead to obscure results when wave packets coexist in close proximity at different latitudes. To remedy this problem, recent papers on downstream wave packet propagation (Lee and Held 1993; Chang and Yu 1999; Chang 2000) have utilized the method of *complex demodulation* (for an in depth review see Bloomfield 2000). This technique assumes that $v(x)$ can be written as $v(x) = \mathrm{Re}(A(x)\exp(ikx))$, where k is a supposed carrier wavenumber. Then the spectrum of $v(x)$ is shifted by multiplying $v(x)$ by $\exp(-ikx)$. Finally, the absolute value of the wave packet envelope $|A(x)|$, is extracted by low-pass filtering to remove high wave number components from the shifted signal. The general wisdom has been that the result is not sensitive to the choice of the carrier wavenumber when it is chosen from a plausible wave number range. (For instance, for carrier wavenumber 7 the position and amplitude of the maximum of $|A(x)|$ is usually identified with acceptable precision, even if the demodulation is done by wavenumber 6 or 8.)

While this assumption leads to reasonable results in most cases, demodulating by the wrong wave number results in incorrect computation of the wave packet envelope when wave packets of distinct carrier wavenumbers coexist at the same latitude.

The aforementioned problem can be easily demonstrated by an example using the following artificial signal that consists of two wave packets with carrier wave numbers 4 and 9 at the same latitude (see Figure 3.1)

$$v(x) = \exp(-(x - 4.5)^2)\cos(4x) + \exp(-(x - 7.5)^2)\cos(9x),\ \pi \le x \le 3\pi. \quad (3.2)$$

Figure 3.1 shows the resulting wave packet envelopes when the demodulation is performed using carrier wavenumbers from 4 through 11. It is evident that demodulating by a single wavenumber distorts the envelope of at least one of the wave packets.

The main aim of this note is to introduce a robust technique for extracting the envelope of atmospheric wave packets that is not affected by the aforementioned problem. The proposed technique does not require the specification of a carrier wave number, is easy to implement, and is computationally inexpensive. Although this technique is well known in digital signal processing (Gabor 1946, Oppenheim and Schafer 1975, Laine and Fan 1996), to the best of our knowledge, it has not previously been used for extracting the envelope of atmospheric wave packets. In what follows, we give a step-by-step description of the algorithm (Section 3.2) and demonstrate its skill using four examples (Section 3.3). These examples include two analytical cases for which the packet envelope is known; the tracking of an upper tropospheric wave packet in operational weather analyses from the National Centers for Environmental Prediction (NCEP), and the tracking of the impact of targeted dropsonde observations that were collected by
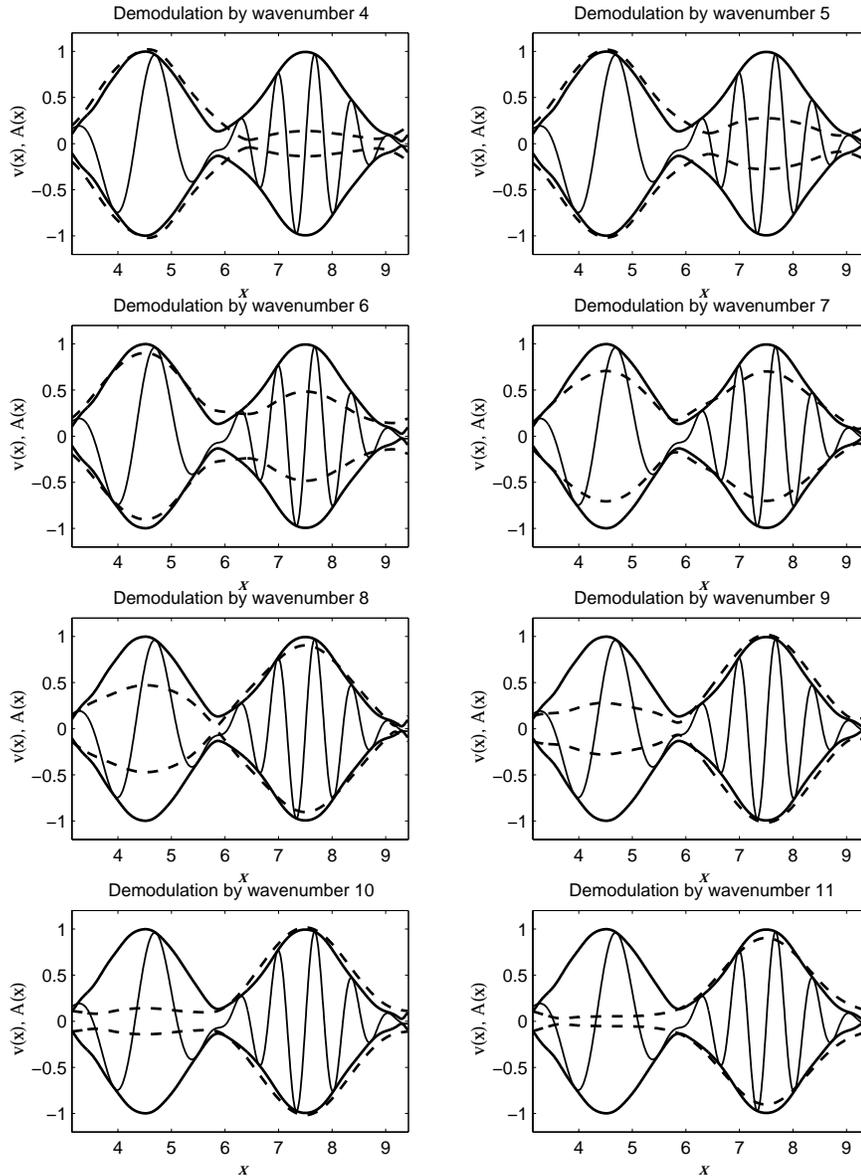
Figure 3.1: An analytical example of a function, $v(x)$, that consists of two wave packets with carrier wave numbers 4 and 9, respectively. Thin line represents $v(x)$, while dashed lines show the result of demodulation by wavenumbers from 4 through 11. Thick solid lines represent the envelope recovered by our method.

one of the flight missions during the 2000 Winter Storm Reconnaissance Program
(Szunyogh *et al.* 2002).

3.2   Description of the algorithm

In what follows, $v(x)$ is considered on an equidistant grid along a latitude circle,
which is parameterized by x, with $0 < x \leq 2\pi$. The grid points are located at
$x = 2\pi l/N$, where $l = 1, 2, ..., N$, and N is an even integer.

The following algorithm is proposed to isolate the wave packet envelopes:

- **Step 1**. The Fourier Transform of the real function $v(x)$ is computed:

$$\hat{v}_k = \frac{1}{N} \sum_{l=1}^{N} v(2\pi l/N)e^{-2\pi ikl/N}, \ (k = -N/2 + 1, ..., N/2). \qquad (3.3)$$

- **Step 2**. The inverse Fourier transform is applied to a selected band ($0 < k_{min} \leq k \leq k_{max}$) of the positive wavenumber half of the Fourier spectrum:

$$w(2\pi l/N) = 2 \sum_{k=k_{min}}^{k_{max}} \hat{v}_k e^{2\pi ikl/N}. \qquad (3.4)$$

- **Step 3**. The packet envelope is computed as follows:

$$A(2\pi l/N) = |w(2\pi l/N)|. \qquad (3.5)$$

This algorithm is a combination of the signal processing technique known as the
Hilbert Transform[2], and a simple filter that retains only the relevant wavenumber

[2]The Hilbert Transform is used in signal processing to recover a complex-valued signal from
the real-valued signal by removing the negative wavenumber components from the spectrum of
the real signal. If X_r is the real signal, the complex signal can be written as $X = X_r + iH(X_r)$,
where $H(\cdot)$ denotes Hilbert Transform.
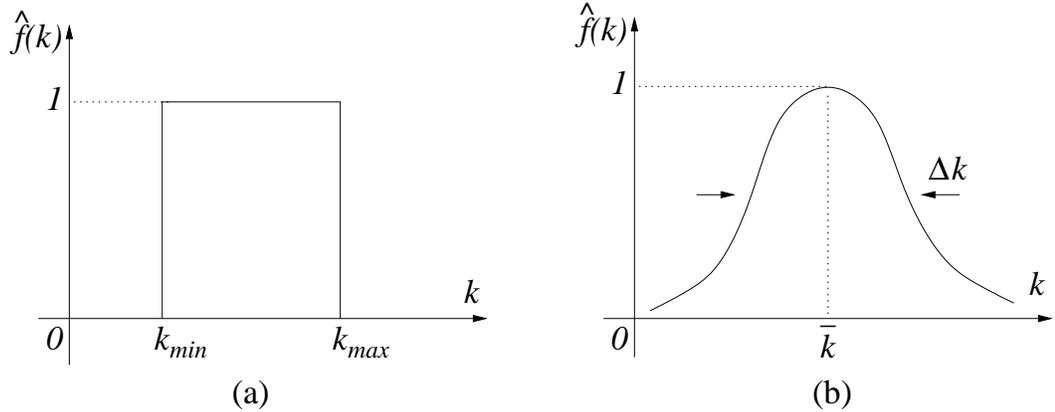
Figure 3.2: Two possible choices of the spectral filter function (see Section 3.2 for details).

components. Equation (3.4) also can be viewed as a wavelet transform in which the wavelet is the inverse Fourier transform $f(x)$ of the function $\hat{f}(k)$ shown in Figure 3.2a, which is $f(x) \sim \{\sin(\Delta kx)/x\} \exp\{-i\bar{k}x\}$, where $\Delta k = (k_{max} - k_{min})/2$ and $\bar{k} = (k_{max} + k_{min})/2$. A smoother filter (wavelet) might be thought to be more natural; e.g., $f(x) \sim \{\exp(\Delta kx)^2/2\}\Delta k \exp\{-i\bar{k}x\}$, corresponding to the k-space filter shown in Figure 3.2b. We also tested this smoother filter and found that it leads to negligible difference in the results while making the method more complicated.

3.3 Examples

Example 1: Let

$$v(x) = \cos[(k-1)x] + \cos[kx] + \cos[(k+1)x], \ x \in [\pi, 3\pi]. \tag{3.6}$$

Applying steps 1 and 2 of our algorithm, we get $w(x) = \exp[i(k-1)x] + \exp[ikx] + \exp[i(k+1)x]$, and the wave packet envelope for this example is $A(x) = |w(x)| = 1 + 2\cos(x)$. Indeed, $v(x)$ can be expressed in the following form:

$$v(x) = \mathrm{Re}(w(x)) = \mathrm{Re}[(1 + 2\cos(x))\exp(ikx)] = \mathrm{Re}[A(x)\exp(ikx)]. \qquad (3.7)$$

It can be easily verified that $w(x)$ can be obtained numerically by applying the Fourier transform to $v(x)$, removing the negative wavenumber part of the spectrum, and then applying the inverse Fourier transform to the resulting spectrum. The result for $k = 7$ is plotted in Figure 3.3.

Example 2: The proposed algorithm is applied to the analytical example, Equation (3.2), that was used in Section 3.1 to demonstrate the shortcomings of the complex demodulation technique. For this example we use $k_{min} = 1$ and $k_{max} = 12$. Figure 3.1 shows clearly that, in contrast to complex demodulation, the proposed algorithm accurately extracts the wave packet envelopes when wave packets with different carrier wave numbers coexist at the same latitude.

Example 3: The proposed algorithm is applied to analyze a case from the 2000 Winter Storm Reconnaissance targeted observations field program (Szunyogh *et al.* 2002). In particular, we consider the wave packet that significantly contributed to the deepening of a trough over the eastern United States on 25 and 26 January. The most significant failure of the numerical weather prediction models in the 1999-2000 winter season was associated with the prediction of the storm related to this trough (for details see Buizza *et al.* 2002; Langland *et al.* 2002; Zhang *et al.* 2002). Our goal is to track the propagation of this upper tropospheric wave packet. We restrict our analyses to the meridional wind component at the 300 hPa pressure level on a global 2.5° by 2.5° resolution grid ($N = 144$) with 24 hour temporal resolution. Since the goal is to track packets of
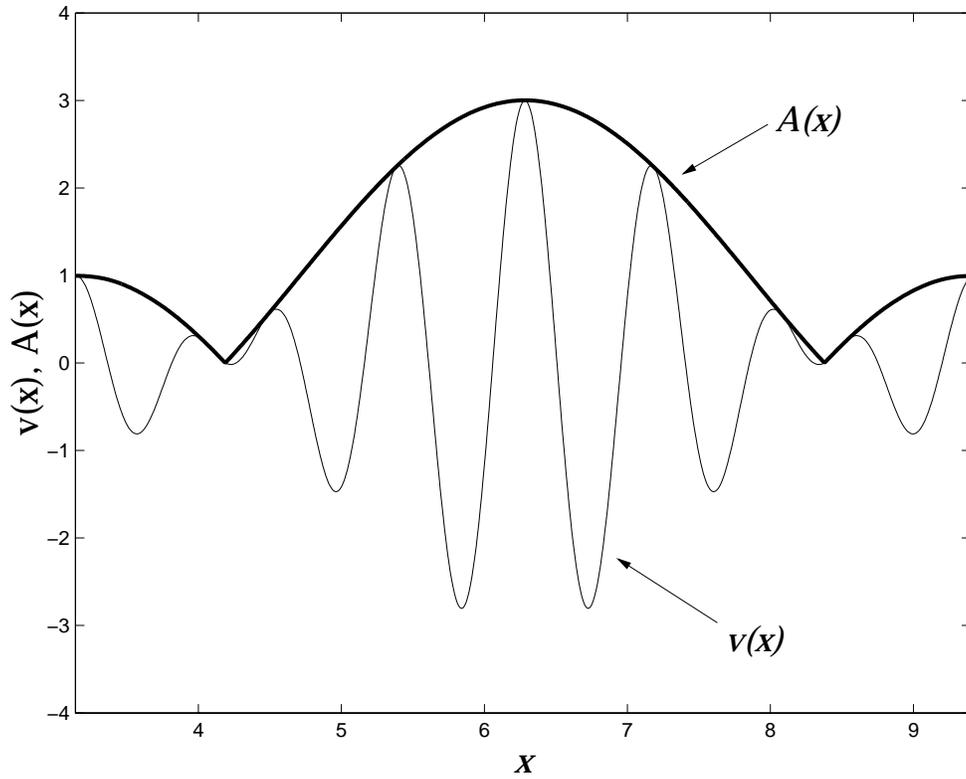
Figure 3.3: An analytical example of a wavepacket (for details see Example 1). Thin line represents the scalar function $v(x)$ and thick line represents the envelope $A(x)$.

short Rossby waves associated with baroclinic energy conversion, k_{min} and k_{max} are chosen to be 4 and 9, respectively. Our wave packet analysis shows clearly why this was a particularly difficult forecast situation. In Figure 3.4, the wave packet can be clearly identified even as early as 21 January 2000 0000 UTC (WP marks the location of the maximum of the packet envelope). At that time the wave packet is located over Japan. Subsequently the hydrodynamical influence in the upper troposphere (and the effect of analysis uncertainties) traveled with the wave packet at an approximate speed of 30 degrees/day from the Pacific regions. Once the wave packet reached the Atlantic, its envelope started to amplify rapidly due to local baroclinic energy conversion.[3] This result corroborates the conclusions of Langland *et al.* (2002), obtained by adjoint sensitivity calculations, that initial condition uncertainties over the Pacific had an influence on the quality of the 72-hour forecasts of the storm.

Example 4: Studies based on the analytical investigation of idealized atmospheric flows predicted long ago, that an initially localized disturbance in the initial condition of a synoptic scale numerical weather prediction would propagate as a packet of synoptic scale Rossby waves (Rossby 1949, Charney 1949). Targeted weather observations (e.g., Snyder 1996; Szunyogh *et al.* 1999; Palmer *et al.* 1998; Gelaro *et al.* 1999; Bergot *et al.* 1999; Pu and Kalnay 1999) change the initial conditions with the aim of removing localized initial condition errors that have potentially large negative forecast effects. As pointed out by Szunyogh et al. (2000 and 2002) based on analysis of data from the 1999 and 2000

---

[3]Our results on the local energetics of the wave packets, are reported in Szunyogh *et al.*, (2002). In that work, the envelope of the wave packet was extracted by the Multiple Wavenumber Packet Identification (MWPI) technique (manuscript is available from the corresponding author), which was an earlier version of the algorithm presented here.
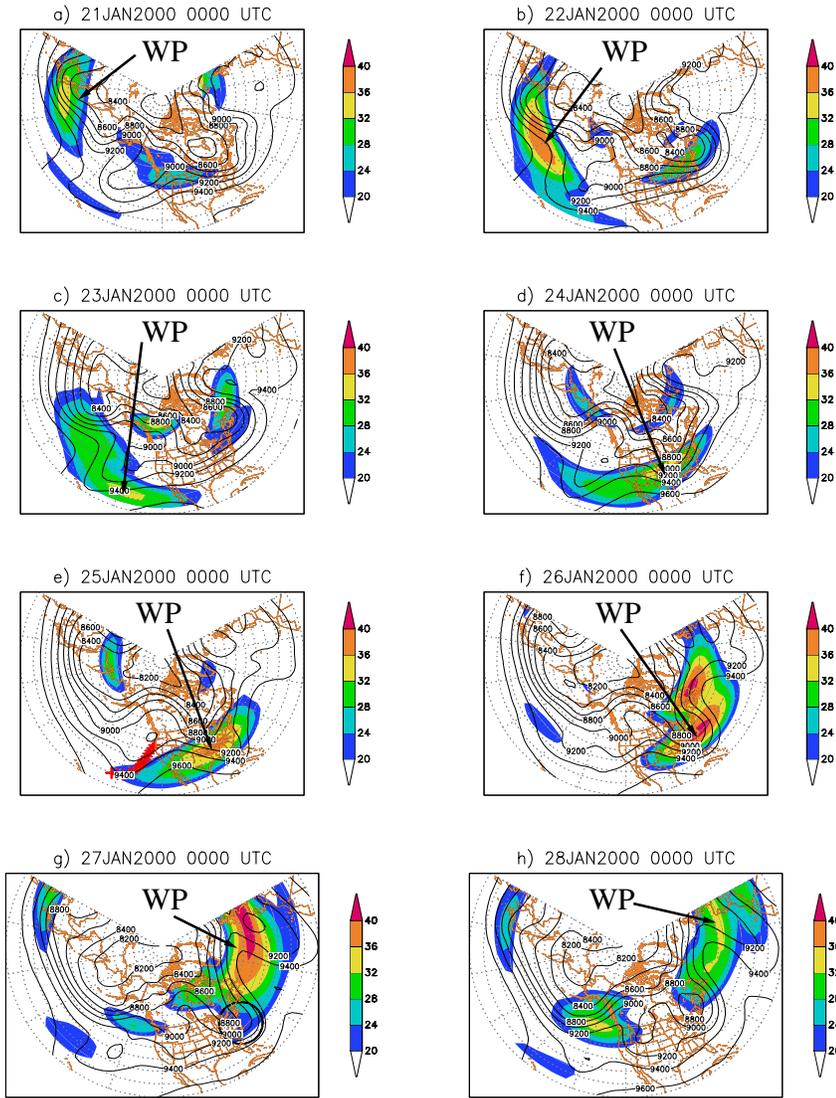
Figure 3.4: Time evolution of wave packet envelope $A(x)$ of the meridional wind velocity at 300 hPa for the eight-day period starting on 21 January 2000 0000 UTC. The values smaller than 20m/s are not shown. Contour lines show the NCEP analysis for the geopotential height at the 300 hPa surface. WP marks the center of the wave packet described in Example 3. The dropsonde locations are shown by red crosses at 25 January 2000 0000 UTC and the verification region is marked by a circle at 27 January 2000 0000 UTC (see Example 4 for details).

Winter Storm Reconnaissance programs, upper tropospheric wave packets play a twofold role in the targeted observation problem. Firstly, they build a dynamical relationship between the targeted region and the verification region, where the forecast is to be improved at a later time. This relationship can be detected by an objective targeting technique, such as the Ensemble Transform Kalman Filter (Bishop et al. 2001). Secondly, it can be expected that the local changes introduced by the targeted data would propagate in the form of Rossby wave packets in the upper troposphere, as it was predicted by the theory of Rossby (1949) and Charney (1949).

While the results of Szunyogh *et al.*(2000 and 2002) strongly indicated that the later statement is true even for complex atmospheric flow configurations, no strong evidence was offered to show that the impact of the dropsondes propagates in the form of synoptic scale Rossby wave packets. One way to demonstrate that this view is correct is to show that the wave packet can be detected in the dropsonde signal. (This signal is defined by the difference between the NCEP global model forecast that was initiated by assimilating all targeted and standard (non-targeted) observations and the NCEP global model forecast that was initiated by assimilating only the standard observations.)

In this example, the algorithm described in Section 3.2 is applied to the signal from a dropsonde mission that was associated with the atmospheric wave packet of Example 3. Targeted observations were collected on 25 January 2000 0000 UTC to improve the prediction of a secondary development on 27 January 2000 0000 UTC over the east coast, and not the major storm itself on 25-26 January 2000 (see Szunyogh et al. 2002 for further details). Figure 3.4 shows that the dropsonde observations targeted the tail of the observed atmospheric wave packet and that
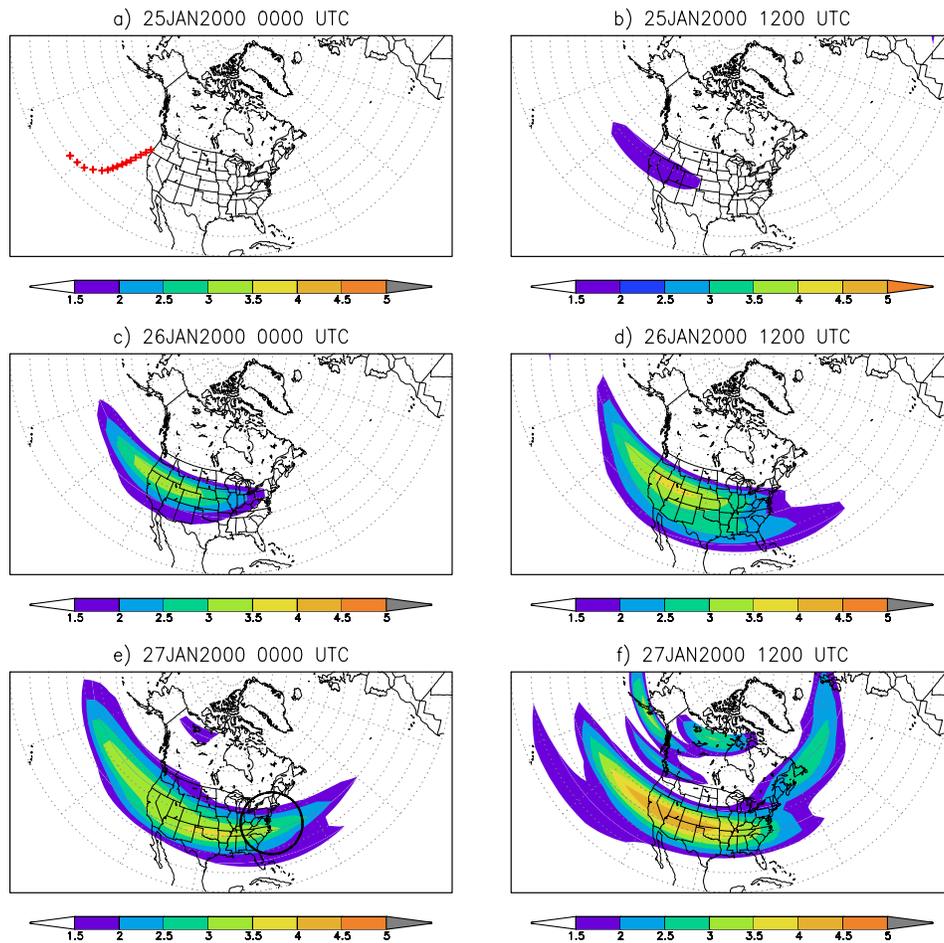
Figure 3.5: Time evolution of the envelope of the 300 hPa geopotential height signal for the four-day period starting 25 January 2000 0000 UTC. The signal is defined by the difference between a forecast that was initiated by assimilating all targeted and standard (non-targeted) observations and a forecast that was initiated by assimilating only the standard observations. The values smaller than 1.5 gpm are not shown. The dropsonde locations are shown by red crosses at 25 January 2000 0000 UTC and the verification region is marked by a circle at 27 January 2000 0000 UTC.

the verification region was also located at the tail of the eastward propagating wave packet at verification time. The results shown in Figure 3.5 demonstrate that after an initial transient, not longer than 12 hours, the data impact in the upper troposphere propagated in the form of a wave packet and the leading edge of the eastward expanding data impact propagated with the speed of the atmospheric wave packet.

3.4   Conclusion

In this note we have presented a robust objective method for extracting the envelope of packets of synoptic scale Rossby waves. We believe that our method provides a potentially useful tool in analyzing meteorological problems related to the propagation of wave packets, such as studying local baroclinic instability (Pierrehumbert and Swanson 1995; Orlanski and Sheldon 1993), tracking the origin of localized forecast errors (Persson 2000), and analyzing the propagation of the influence of targeted observations (Szunyogh *et.al.* 2002).

Acknowledgments

Chapter 4

Local Ensemble Data Assimilation

[1]

4.1   Introduction

The purpose of this paper is to develop and test a new atmospheric data assimilation scheme, which we call the Local Ensemble Kalman Filter method. Atmospheric *data assimilation* (*analysis*) is the process through which an estimate of the atmospheric state is obtained by using observed data and a dynamical model of the atmosphere (e.g., Daley 1991; Kalnay 2002). These estimates, called *analyses*, can then be used as initial conditions in operational numerical weather predictions. In addition, diagnostic studies of atmospheric dynamics and climate are also often based on analyses instead of raw observed data.

The analysis at a given time instant is a maximum likelihood estimate of the atmospheric state in which a short-term forecast, usually referred to as the *background* or *first guess field*, is used as a *prior* estimate of the atmospheric state

---

[1]This chapter is a verbatim representation of the paper *Local Ensemble Data Assimilation* by E. Ott et al. submitted to Monthly Weather Review.

(Lorenc 1986). Then the observations are assimilated into the background field by a statistical interpolation. This interpolation is performed based on the assumptions *(i)* that the uncertainties in the background field and the observations are normally distributed and *(ii)* that the covariance between different components of the background (formally the *the background covariance matrix*) and the covariances between uncertainties in the noisy observations (formally the *observational error covariance matrix*) are known. In reality, however, the background error covariance matrix cannot be directly computed. The implementation of a data assimilation system, therefore, requires the development of statistical models that can provide an estimate of the background error covariance matrix. The quality of a data assimilation system is primarily determined by the accuracy of this estimate.

The mathematically consistent technique to define an adaptive background covariance matrix is the Kalman Filter (Kalman 1960; Kalman and Bucy 1961) which utilizes the dynamical equations to evolve the most probable state and the error covariance matrix in time. Although the Kalman Filter approach has been successfully implemented for a wide range of applications and has been considered for atmospheric data assimilation for a long while (Jones 1965; Petersen 1973; Ghil et al. 1981, Dee et al., 1985), the computational cost involved does not allow for an operational implementation in the foreseeable future (see Daley 1991 for details).

The currently most popular approach to reduce the cost of the Kalman Filter is to use a relatively small (10-100 member) ensemble of background forecasts to estimate the background error covariances (e.g. Evensen 1994; Houtekamer and Mitchell 1998, 2001; Bishop et al. 2001; Hamill et al. 2001; Whitaker and Hamill

2002). In ensemble-based data assimilation schemes the ensemble of background forecasts is generated by using initial conditions distributed according to the result of the previous analysis.

The ensemble-based approach has the additional appeal of providing initial ensemble perturbations that are consistent with the analysis scheme. This is important because currently implemented operational techniques generate initial ensemble perturbations without use of direct information about the analysis errors (Toth and Kalnay 1997; Molteni et al. 1996). These techniques are obviously suboptimal considering the goal of ensemble forecasting, which is to simulate the effect of the analysis uncertainties on the ensuing forecasts.

The main difference between the existing ensemble-based schemes is in the generation of the analysis ensemble. One family of schemes is based on *perturbed observations* (Evensen and van Leeuwen 1996; Houtekamer and Mitchell 1998, 2001; Hamill and Snyder 2000, 2001). In this approach, the analysis ensemble is obtained by assimilating a different set of observations to each member of the background ensemble. The different sets of observations are created by adding random noise to the real observations, where the random noise component is generated according to the observational error covariance matrix. The main weakness of this approach is that the ensemble size must be large in order to accurately represent the probability distribution of the background errors. Thus a relatively large forecast ensemble has to be evolved in time, limiting the efficiency of the approach. The most recent papers demonstrated, however, that the required size of the ensemble can be reduced by evolving two smaller ensembles and filtering the long distance covariances from the background field (Houtekamer and Mitchell 1998 and 2001; Hamill et al. 2001).

The other family of schemes, the Kalman square-root filters, use a different approach to reduce the size of the ensemble. These techniques do the analysis only once, to obtain the mean analysis. Then the analysis ensemble perturbations (to the mean analysis) are generated by linearly transforming the background ensemble perturbations to a set of vectors that can be used to represent the *analysis error covariance matrix*. This strategy is feasible because the analysis error covariance matrix can be computed explicitly from the background and observational error covariance matrices. Since there is an infinite set of analysis perturbations that can be used to represent the analysis error covariance matrix, many different schemes can be derived following this approach (Tippett et al., 2002). Existing examples of the square root filter approach are the Ensemble Transform Kalman Filter (Bishop et al. 2001), the Ensemble Adjustment Filter (Anderson 2001), and the Ensemble Square-root Filter (Whitaker and Hamill 2001).

The scheme we propose is a Kalman square-root filter [2]. The most important difference between our scheme and the other Kalman square-root filters is that our analysis is done locally in model space. The assimilations in each local region are independent, thus facilitating a massively parallel approach. The development of this approach was motivated by the finding of Patil et al. (2001, 2002) that synoptic scale ensemble perturbations restricted to a small region frequently evolve in a low dimensional subspace of the full state space of that region. In this sense, our paper is related to previous work that attempted to construct a

---

[2]The basic algorithm was first described in the paper Ott, E., B. H. Hunt, I. Szunyogh, M. Corazza, E. Kalnay, D. J. Patil, J. A. Yorke, A. V. Zimin, and E. Kostelich, 2002: "Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering" (http://arxiv.org/abs/physics/0203058).

simplified Kalman Filter by explicitly taking into account the dominant unstable directions of the state space (Kalnay and Toth 1994; Fisher 1998).

Our scheme is based in the construction of local regions about each grid point. An outline of the scheme is as follows.

1. Advance the analysis ensemble of global atmospheric states to the next analysis time, thus obtaining a new background ensemble of global atmospheric states.

2. For each local region and each member of the background ensemble, form vectors of the atmospheric state information in that local region.

3. In each local region, project the 'local' vectors, obtained in step 2, onto a low dimensional subspace that best represents the ensemble in that region.

4. Do the data assimilation in each of the local low dimensional subspaces, obtaining analyses in each local region.

5. Use the local analyses, obtained in step 4, to form a new global analysis ensemble. (This is where the square root filter comes in.)

6. Go back to step 1.

This method is potentially advantageous in that the individual local analyses are done in low dimensional subspaces, so that matrix operations involve only relatively low dimensional matrices. Furthermore, since the individual analyses in different local regions do not interact, they can be done independently in parallel.

In the following sections we describe and test our new approach to data assimilation. Section 4.2 introduces the concept of local regions and explains how

the dimension of the local state vector can be further reduced. Section 4.3 explains the analysis scheme for the local regions. In section 4.4, the local analyses are pieced together to obtain the global analysis field and the ensemble of global analysis perturbations. Section 4.5 illustrates our data assimilation scheme by an application to the 40-variable Lorenz model (Lorenz 1996).

4.2   Local vectors and their covariance

A model state of the atmosphere is given by a vector field $\mathbf{x}(\mathbf{r}, t)$ where $\mathbf{r}$ is two dimensional and runs over discrete values $\mathbf{r}_{mn}$ (the grid in the physical space used in the numerical computations). Typically, the two components of $\mathbf{r}$ are the geographical longitude and latitude, and $\mathbf{x}$ at a fixed $\mathbf{r}$ is a vector of all relevant physical state variables of the model (e.g., wind velocity components, temperature, surface pressure, humidity, etc., at all height levels included in the model). Let u denote the dimensionality of $\mathbf{x}(\mathbf{r}, t)$ (at fixed $\mathbf{r}$); e.g., when five independent state variables are defined at 28 vertical levels, $u = 140$.

Data assimilation schemes generally treat $\mathbf{x}(\mathbf{r}, t)$ as a random variable with a time-dependent probability distribution. The distribution is updated over time in two ways: (i) it is evolved according to the model dynamics; and (ii) it is modified periodically to take into account recent atmospheric observations.

We do our analysis locally in model space (grid point by grid point). In this section we introduce our local coordinate system and the approximations we make to the local probability distribution of $\mathbf{x}(\mathbf{r}, t)$. Since all the analysis operations take place at a fixed time t, we will suppress the t dependence of all vectors and matrices introduced henceforth.

Motivated by the works of Patil et al. (2001, 2002) we introduce at each point

local vectors $\mathbf{x}_{mn}$ of the information $\mathbf{x}(\mathbf{r}_{m+m',n+n'}, t)$ for $-l \leq m', n' \leq l$. That is, $\mathbf{x}_{mn}$ specifies the model atmospheric state within a $(2l+1)$ by $(2l+1)$ patch of grid points centered at $\mathbf{r}_{mn}$. (This particular shape of the local region was chosen to keep the notations as simple as possible, but in general different (e.g., circular) shape regions can also be considered.) The dimensionality of $\mathbf{x}_{mn}$ is $(2l+1)^2 u$. We represent the construction of local vectors via a linear operator $\mathbf{M}_{mn}$,

$$\mathbf{x}_{mn} = \mathbf{M}_{mn}\mathbf{x}(\mathbf{r}, t). \tag{4.1}$$

We now consider local vectors obtained from the model as forecasts, using initial conditions distributed according to the result of the previous analysis, and we denote these by $\mathbf{x}_{mn}^b$ (where the superscript b stands for "background"). Let $F_{mn}(\mathbf{x}_{mn}^b)$ be our approximation to the probability density function for $\mathbf{x}_{mn}^b$ at the current analysis time t. A fundamental assumption is that this probability distribution can be usefully approximated as Gaussian,

$$F_{mn}(\mathbf{x}_{mn}^b) \sim \exp\left[-\frac{1}{2}(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b)^T (\mathbf{P}_{mn}^b)^{-1}(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b)\right], \tag{4.2}$$

where $\mathbf{P}_{mn}^b$ and $\bar{\mathbf{x}}_{mn}^b$ are the *local background error covariance matrix* and most probable state associated with $F_{mn}(\mathbf{x}_{mn}^b)$. Graphically, the level set

$$F_{mn}(\mathbf{x}_{mn}^b) = e^{-1/2} F_{mn}(\bar{\mathbf{x}}_{mn}^b) \tag{4.3}$$

is an ellipsoid as illustrated in Figure 4.1. The equation of this *probability ellipsoid* is

$$(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b)^T (\mathbf{P}_{mn}^b)^{-1}(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b) = 1. \tag{4.4}$$

As explained subsequently, the rank of the $(2l+1)^2 u$ by $(2l+1)^2 u$ covariance matrix $\mathbf{P}_{mn}^b$ for our approximate probability distribution function F_{mn} is much less than $(2l+1)^2 u$. Let

$$k = \text{rank}(\mathbf{P}_{mn}^b); \tag{4.5}$$
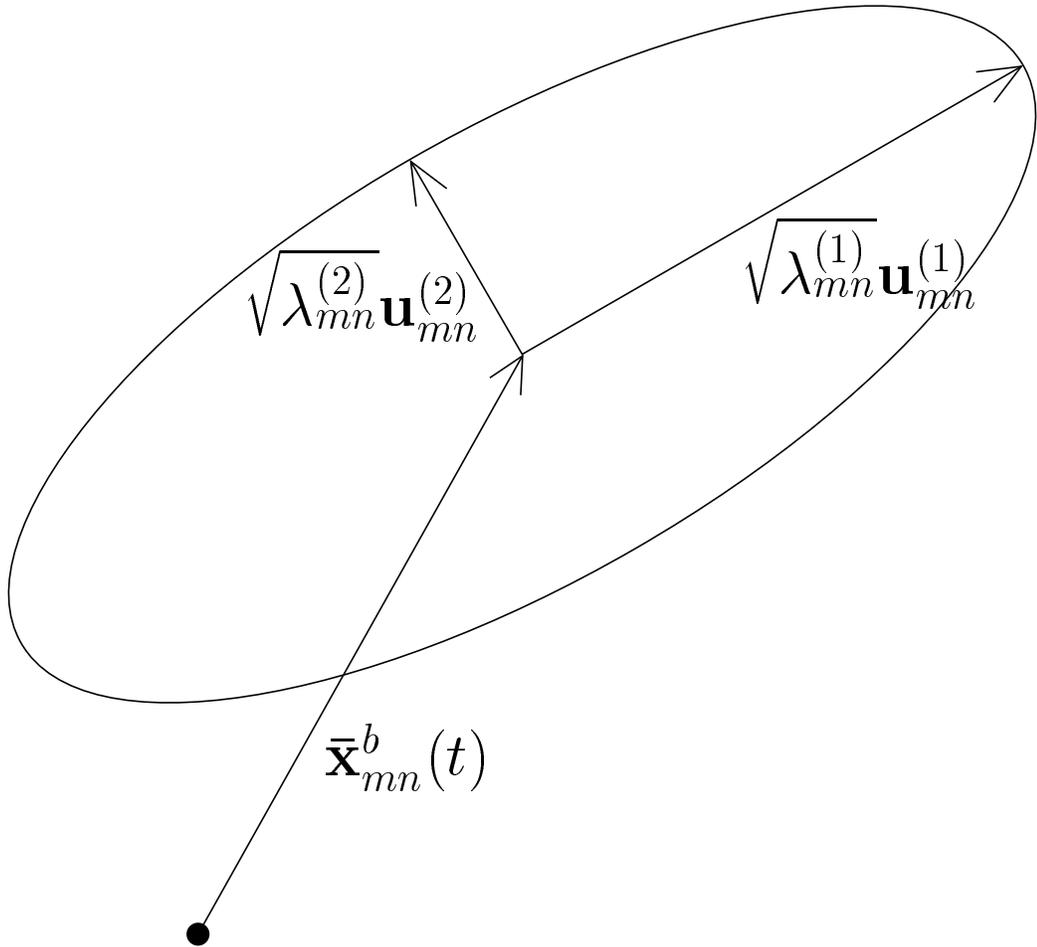
Figure 4.1: Probability ellipsoid for $\mathbf{x}_{mn}^b$.

($k = 2$ in Figure 1). Thus $\mathbf{P}_{mn}^b$ has a $(2l+1)^2 u - k$ dimensional null space $\bar{\mathbb{S}}_{mn}$ and the inverse $(\mathbf{P}_{mn}^b)^{-1}$ is defined for vectors $(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b)$ lying in the k dimensional subspace $\mathbb{S}_{mn}$ orthogonal to $\bar{\mathbb{S}}_{mn}$, with F_{mn} defined to be zero if $(\mathbf{x}_{mn}^b - \bar{\mathbf{x}}_{mn}^b)$ is not in $\mathbb{S}_{mn}$.

In the data assimilation procedure we describe in this work, the background error covariance matrix $\mathbf{P}_{mn}^b$ and the most probable background state $\bar{\mathbf{x}}_{mn}^b$ are derived from a $k' + 1$ member ensemble of global state field vectors $\{\mathbf{x}^{b(i)}(\mathbf{r}, t)\}$,

$i = 1, 2, \cdots, k' + 1$; $k' \geq k \geq 1$. The most probable state is given by

$$\bar{\mathbf{x}}^b_{mn} = \mathbf{M}_{mn}\big[(k'+1)^{-1}\sum_{i=1}^{k'+1}\mathbf{x}^{b(i)}(\mathbf{r}, t)\big]. \tag{4.6}$$

To obtain the local background error covariance matrix $\mathbf{P}^b_{mn}$ that we use in our analysis, we first consider a matrix $\mathbf{P}^{b'}_{mn}$ given by

$$\mathbf{P}^{b'}_{mn} = k'^{-1}\sum_{i=1}^{k'+1}\delta\mathbf{x}^{b(i)}_{mn}\big(\delta\mathbf{x}^{b(i)}_{mn}\big)^T, \tag{4.7}$$

where

$$\delta\mathbf{x}^{b(i)}_{mn} = \mathbf{M}_{mn}\mathbf{x}^{b(i)}(\mathbf{r}, t) - \bar{\mathbf{x}}^b_{mn}(\mathbf{r}, t). \tag{4.8}$$

It is also useful to introduce the notation

$$\mathbf{X}^b_{mn} = (k')^{-1/2}\big[\delta\mathbf{x}^{b(1)}_{mn} \mid \delta\mathbf{x}^{b(2)}_{mn} \mid \cdots \mid \delta\mathbf{x}^{b(k'+1)}_{mn}\big], \tag{4.9}$$

in terms of which (4.7) can be rewritten,

$$\mathbf{P}^{b'}_{mn} = \mathbf{X}^b_{mn}\mathbf{X}^{bT}_{mn}. \tag{4.10}$$

Patil et al. (2001, 2002), using 30-pair ensembles of bred vectors (Toth and Kalnay 1993, 1997), have found that forecast uncertainties in the mid-latitude extra-tropics tend to lie in a low dimensional subset of the $(2l+1)^2u$ dimensional local vector space. Thus we anticipate that we can approximate the background error covariance matrix by one of much lower rank than $(2l+1)^2u$, and this motivates our assumption that an ensemble of size of $k'+1$, where $k'+1$ is substantially less than $(2l+1)^2u$, will be sufficient to yield a good approximate representation of the background covariance matrix. Typically, $\mathbf{P}^{b'}_{mn}$ has rank k', i.e., it has k' positive eigenvalues. Let the eigenvalues of the matrix $\mathbf{P}^{b'}_{mn}$ be denoted by $\lambda^{(j)}_{mn}$, where the labeling convention for the index j is

$$\lambda^{(1)}_{mn} \geq \lambda^{(2)}_{mn} \geq \ldots \geq \lambda^{(k)}_{mn} \geq \cdots \geq \lambda^{(k')}_{mn}. \tag{4.11}$$

Since $\mathbf{P}_{mn}^{b'}$ is a symmetric matrix, it has k' orthonormal eigenvectors $\{\mathbf{u}_{mn}^{(j)}\}$ corresponding to the k' eigenvalues (4.11). Thus

$$\mathbf{P}_{mn}^{b'} = \sum_{j=1}^{k'} \lambda_{mn}^{(j)} \mathbf{u}_{mn}^{(j)} (\mathbf{u}_{mn}^{(j)})^T. \tag{4.12}$$

We approximate $\mathbf{P}_{mn}^{b'}$ by truncating the sum at $k \leq k'$

$$\mathbf{P}_{mn}^{b} = \sum_{j=1}^{k} \lambda_{mn}^{(j)} \mathbf{u}_{mn}^{(j)} \left(\mathbf{u}_{mn}^{(j)}\right)^T. \tag{4.13}$$

In terms of $\mathbf{u}_{mn}^{(j)}$ and $\lambda_{mn}^{(j)}$, the principal axes of the probability ellipsoid (Figure 1) are given by

$$\sqrt{\lambda_{mn}^{(j)}} \mathbf{u}_{mn}^{(j)}. \tag{4.14}$$

The basic justification for the approximation of the covariance by $\mathbf{P}_{mn}^{b}$ is our supposition, supported by Patil et al. (2002), that for reasonably small values of k, the error variance in all other directions is much less than the variance,

$$\sum_{j=1}^{k} \lambda_{mn}^{(j)}, \tag{4.15}$$

in the directions $\{\mathbf{u}_{mn}^{(j)}\}$, $j = 1, 2, \ldots, k$. The truncated covariance matrix $\mathbf{P}_{mn}^{b}$ is determined not only by the dynamics of the model but also by the choice of the components of $\delta\mathbf{x}_{mn}^{b(i)}$. In order to meaningfully compare eigenvalues, Equation (4.11), the different components of $\delta\mathbf{x}_{mn}^{b(i)}$ (e.g., wind and temperature) should be properly scaled to ensure that, if the variance (4.15) approximates the full variance, then the first k eigendirections, $\{\mathbf{u}_{mn}^{(j)}\}$, $j = 1, 2, \ldots, k$, explain the important uncertainties in the background, $\bar{\mathbf{x}}_{mn}^{b}$. In Patil et al. (2002), for instance, the weights for the different variables were chosen so that the Euclidean norm of the transformed vectors was equal to their *total energy norm* derived in Talagrand (1981). In what follows, we assume that the vector components are already

properly scaled. (We also note that if $k = k'$, the comparison of eigenvalues is not used and thus such a consistent scaling of the variables is not necessary.)

For the purpose of subsequent computation, we consider the coordinate system for the k dimensional space $\mathbb{S}_{mn}$ determined by the basis vectors $\{\mathbf{u}_{mn}^{(j)}\}$. We call this the *internal coordinate system* for $\mathbb{S}_{mn}$. To change between the internal coordinates and those of the local space, we introduce the $(2l+1)^2 u$ by k matrix,

$$\mathbf{Q}_{mn} = \left\{ \mathbf{u}_{mn}^{(1)} | \mathbf{u}_{mn}^{(2)} | \cdots | \mathbf{u}_{mn}^{(k)} \right\}. \tag{4.16}$$

We denote the projection of vectors into $\mathbb{S}_{mn}$ and the restriction of matrices to $\mathbb{S}_{mn}$ by a superscribed circumflex (hat). Thus for a $(2l+1)^2 u$ dimensional column vector $\mathbf{w}$, the vector $\hat{\mathbf{w}}$ is a k dimensional column vector given by

$$\hat{\mathbf{w}} = \mathbf{Q}_{mn}^T \mathbf{w}. \tag{4.17}$$

Note that this operation consists of both projecting $\mathbf{w}$ into $\mathbb{S}_{mn}$ and changing to the internal coordinate system. Similarly, for a $(2l+1)^2 u$ by $(2l+1)^2 u$ matrix $\mathbf{U}$, the matrix $\hat{\mathbf{U}}$ is k by k and given by

$$\hat{\mathbf{U}} = \mathbf{Q}_{mn}^T \mathbf{U} \mathbf{Q}_{mn}. \tag{4.18}$$

To go back to the original $(2l+1)^2 u$ dimensional local vector space, note that $\mathbf{Q}_{mn}^T \mathbf{Q}_{mn} = \mathbf{I}$ while $\mathbf{Q}_{mn} \mathbf{Q}_{mn}^T$ represents projection on $\mathbb{S}_{mn}$, i.e. it has null space $\bar{\mathbb{S}}_{mn}$ and acts as the identity on $\mathbb{S}_{mn}$. Then writing $\mathbf{w}$ as

$$\mathbf{w} = \mathbf{w}^{(\|)} + \mathbf{w}^{(\perp)}, \tag{4.19}$$

$$\mathbf{w}^{(\|)} = \mathbf{\Lambda}_{mn}^{(\|)} \mathbf{w} = \mathbf{Q}_{mn} \hat{\mathbf{w}}, \qquad \mathbf{w}^{(\perp)} = \mathbf{\Lambda}_{mn}^{(\perp)} \mathbf{w}, \tag{4.20}$$

where $\mathbf{w}^{(\|)}$ and $\mathbf{w}^{(\perp)}$ denote the components of $\mathbf{w}$ in $\mathbb{S}_{mn}$ and $\bar{\mathbb{S}}_{mn}$, respectively, and the projection operators $\mathbf{\Lambda}_{mn}^{(\|)}$ and $\mathbf{\Lambda}_{mn}^{(\perp)}$ are given by

$$\mathbf{\Lambda}_{mn}^{(\|)} = \mathbf{Q}_{mn} \mathbf{Q}_{mn}^T, \qquad \mathbf{\Lambda}_{mn}^{(\perp)} = \mathbf{I} - \mathbf{Q}_{mn} \mathbf{Q}_{mn}^T \tag{4.21}$$

In addition, if $\mathbf{U}$ is symmetric with null space $\bar{\mathbb{S}}_{mn}$,

$$\mathbf{U} = \mathbf{Q}_{mn} \hat{\mathbf{U}} \mathbf{Q}_{mn}^T. \tag{4.22}$$

Note that

$$\hat{\mathbf{P}}_{mn}^b = \text{diag}\big(\lambda_{mn}^{(1)}, \lambda_{mn}^{(2)}, ..., \lambda_{mn}^{(k)}\big), \tag{4.23}$$

and thus it is trivial to invert.

4.3 Data assimilation

With Section 4.2 as background, we now consider the assimilation of observational data to obtain a new specification of the probability distribution of the local vector. In what follows, the notational convention of Ide et al. (1997) is adopted whenever it is possible.

Let $\mathbf{x}_{mn}^a$ be the random variable at the current analysis time t representing the local vector after knowledge of the observations is taken into account. For simplicity, we assume that all observations collected for the current analysis were taken at the same time t. Let $\mathbf{y}_{mn}^o$ be the vector of current observations within the local region, and assume that the errors in these observations are normally distributed with covariance matrix $\mathbf{R}_{mn}$. We also assume that the expected observation vector $\bar{\mathbf{y}}_{mn}^o$ can be written as a linear operator $\mathbf{H}_{mn}$ times the true local state of the atmosphere. (If there are s scalar observations in the local $(2l + 1)$ by $(2l + 1)$ region at analysis time t, then $\bar{\mathbf{y}}_{mn}^o$ is s dimensional and the rectangular matrix $\mathbf{H}_{mn}$ is s by $(2l + 1)^2 u$). Then, since we have assumed the background (pre-analysis) state $\mathbf{x}_{mn}^b$ to be normally distributed, it will follow below that $\mathbf{x}_{mn}^a$ is also normally distributed. Its distribution is determined by the most probable state $\bar{\mathbf{x}}_{mn}^a$ and the associated covariance matrix $\mathbf{P}_{mn}^a$. The data

assimilation step determines $\bar{\mathbf{x}}_{mn}^a$ (the *local analysis*) and $\mathbf{P}_{mn}^a$ (the *local analysis covariance matrix*).

Since our approximate background covariance matrix $\mathbf{P}_{mn}^b$ has null space $\bar{\mathbb{S}}_{mn}$, we consider the perturbation component $\Delta\mathbf{x}_{mn}^{a(\|)} = \Lambda_{mn}^{(\|)}(\mathbf{x}_{mn}^a - \bar{\mathbf{x}}_{mn}^b)$ within the k-dimensional subspace $\mathbb{S}_{mn}$, and do the data assimilation in $\mathbb{S}_{mn}$. Thus the data assimilation is done by minimizing the quadratic form,

$$
\begin{aligned}
J\big(\Delta\hat{\mathbf{x}}_{mn}^a\big) &= \big(\Delta\hat{\mathbf{x}}_{mn}^a\big)^T\big(\hat{\mathbf{P}}_{mn}^b\big)^{-1}\Delta\hat{\mathbf{x}}_{mn}^a \\
&+ \big(\hat{\mathbf{H}}_{mn}\Delta\hat{\mathbf{x}}_{mn}^a + \mathbf{H}_{mn}\bar{\mathbf{x}}_{mn}^b - \mathbf{y}_{mn}^o\big)^T\mathbf{R}_{mn}^{-1}\times \\
&\quad \big(\hat{\mathbf{H}}_{mn}\Delta\hat{\mathbf{x}}_{mn}^a + \mathbf{H}_{mn}\bar{\mathbf{x}}_{mn}^b - \mathbf{y}_{mn}^o\big).
\end{aligned}
\tag{4.24}
$$

Here $\hat{\mathbf{H}}_{mn} = \mathbf{H}_{mn}\mathbf{Q}_{mn}$ maps $\mathbb{S}_{mn}$ to the observation space, using the internal coordinate system for $\mathbb{S}_{mn}$ introduced in the previous section, so that $\Delta\mathbf{x}_{mn}^{a(\|)} = \mathbf{Q}_{mn}\Delta\hat{\mathbf{x}}_{mn}^a$. The most probable value of $\Delta\hat{\mathbf{x}}_{mn}^a$,

$$
\Delta\hat{\bar{\mathbf{x}}}_{mn}^a = \hat{\mathbf{P}}_{mn}^a\hat{\mathbf{H}}_{mn}^T\mathbf{R}_{mn}^{-1}\big(\mathbf{y}_{mn}^o - \mathbf{H}_{mn}\bar{\mathbf{x}}_{mn}^b\big),
\tag{4.25}
$$

is the minimizer of $J\big(\Delta\hat{\mathbf{x}}_{mn}^a\big)$, where the analysis covariance matrix $\hat{\mathbf{P}}_{mn}^a$ is the inverse of the matrix of second derivatives (Hessian) of $J\big(\Delta\hat{\mathbf{x}}_{mn}^a\big)$ with respect to $\Delta\hat{\mathbf{x}}_{mn}^a$,

$$
\hat{\mathbf{P}}_{mn}^a = \big[\big(\hat{\mathbf{P}}_{mn}^b\big)^{-1} + \hat{\mathbf{H}}_{mn}^T\mathbf{R}_{mn}^{-1}\hat{\mathbf{H}}_{mn}\big]^{-1}.
\tag{4.26}
$$

For computational purposes, we prefer to use the alternate form,

$$
\hat{\mathbf{P}}_{mn}^a = \hat{\mathbf{P}}_{mn}^b\big[\mathbf{I} + \hat{\mathbf{H}}_{mn}^T\mathbf{R}_{mn}^{-1}\hat{\mathbf{H}}_{mn}\hat{\mathbf{P}}_{mn}^b\big]^{-1},
\tag{4.27}
$$

both in place of (4.26) and in computing (4.25). A potential numerical advantage of (4.27) over (4.26) is that (4.26) involves the inverse of $\hat{\mathbf{P}}_{mn}^b$, which may be problematic if $\hat{\mathbf{P}}_{mn}^b$ has a small eigenvalue.

Another alternative is to compute (4.25) and (4.26) in terms of the "Kalman gain" matrix

$$\hat{\mathbf{K}}_{mn} = \hat{\mathbf{P}}^b_{mn}\hat{\mathbf{H}}^T_{mn}\big(\hat{\mathbf{H}}_{mn}\hat{\mathbf{P}}^b_{mn}\hat{\mathbf{H}}^T_{mn} + \mathbf{R}_{mn}\big)^{-1}. \tag{4.28}$$

Then it can be shown (e.g., Kalnay 2002, p. 171) that (4.25) and (4.26)/(4.27) are equivalent to

$$\Delta\hat{\bar{\mathbf{x}}}^a_{mn} = \hat{\mathbf{K}}_{mn}\big(\mathbf{y}^o_{mn} - \mathbf{H}_{mn}\bar{\mathbf{x}}^b_{mn}\big), \tag{4.29}$$

and

$$\hat{\mathbf{P}}^a_{mn} = \big(\mathbf{I} - \hat{\mathbf{K}}_{mn}\hat{\mathbf{H}}_{mn}\big)\hat{\mathbf{P}}^b_{mn}. \tag{4.30}$$

Again, the inverse of $\hat{\mathbf{P}}^b_{mn}$ is not required.

Though (4.25) and (4.27) are mathematically equivalent to (4.28)–(4.30), the former approach may be significantly more efficient computationally for the following reasons. In both cases, one must invert an s by s matrix, where s is the number of local observations. While these matrices are considerably smaller than those involved in global data assimilation schemes, they may still be quite large. Generally the s by s matrix $\mathbf{R}_{mn}$ whose inverse is required in (4.27) will be diagonal or close to diagonal, and thus less expensive to invert than the matrix inverted in (4.28). (Furthermore, in some cases one may be able to treat $\mathbf{R}_{mn}$ as time-independent and avoid recomputing its inverse for each successive analysis.) The additional inverse required in (4.27) is of a k by k matrix, where $k \leq k'$ may be relatively small compared to s if the number of observations in the local region (m, n) is large.

Finally, going back to the local space representation, we have

$$\bar{\mathbf{x}}^a_{mn} = \mathbf{Q}_{mn}\Delta\hat{\bar{\mathbf{x}}}^a_{mn} + \bar{\mathbf{x}}^b_{mn}, \tag{4.31}$$

4.4 Updating the ensemble

We now wish to use the analysis information, $\hat{\mathbf{P}}^a_{mn}$ and $\bar{\mathbf{x}}^a_{mn}$, to obtain an ensemble of global analysis fields $\{\mathbf{x}^{a(i)}(\mathbf{r}, t)\}$; $i = 1, 2, \cdots, k' + 1$. Once these fields are determined, they can be used as initial conditions for the atmospheric model. Integrating these global fields forward in time to the next analysis time $t + \Delta t$, we obtain the background ensemble $\{\mathbf{x}^{b(i)}(\mathbf{r}, t + \Delta t)\}$. This completes the loop, and, if the procedure is stable, it can be repeated for as long as desired. Thus at each analysis time we are in possession of a global initial condition that can be used for making forecasts of the desired durations.

Our remaining task is to specify the ensemble of global analysis fields $\{\mathbf{x}^{a(i)}(\mathbf{r}, t)\}$ from our analysis information, $\hat{\mathbf{P}}^a_{mn}$ and $\bar{\mathbf{x}}^a_{mn}$. Denote $(k'+1)$ local analysis vectors by

$$\mathbf{x}^{a(i)}_{mn} = \bar{\mathbf{x}}^a_{mn} + \delta\mathbf{x}^{a(i)}_{mn} \tag{4.32}$$

$$\delta\mathbf{x}^{a(i)}_{mn} = \delta\mathbf{x}^{a(i)(\|)}_{mn} + \delta\mathbf{x}^{a(i)(\perp)}_{mn}, \tag{4.33}$$

$$\delta\mathbf{x}^{a(i)(\|)}_{mn} = \mathbf{\Lambda}^{(\|)}_{mn}\delta\mathbf{x}^{a(i)}_{mn}, \qquad \delta\mathbf{x}^{a(i)(\perp)}_{mn} = \mathbf{\Lambda}^{(\perp)}_{mn}\delta\mathbf{x}^{a(i)}_{mn}, \tag{4.34}$$

$$\delta\mathbf{x}^{a(i)(\perp)}_{mn} = \delta\mathbf{x}^{b(i)(\perp)}_{mn}. \tag{4.35}$$

Combining (4.20) and (4.32)-(4.35), we have

$$\mathbf{x}^{a(i)}_{mn} = \bar{\mathbf{x}}^a_{mn} + \mathbf{Q}_{mn}\delta\hat{\mathbf{x}}^{a(i)} + \mathbf{\Lambda}^{(\perp)}_{mn}\delta\mathbf{x}^{b(i)}_{mn}. \tag{4.36}$$

Equation (4.35) results because our analysis uses the observations only to reduce the variance in the space $\mathbb{S}_{mn}$, leaving the variance in $\bar{\mathbb{S}}_{mn}$ unchanged. (We note, however, that by our construction of $\bar{\mathbb{S}}_{mn}$ in section 4.2, the total variance in $\bar{\mathbb{S}}_{mn}$ is expected to be small compared to that in $\mathbb{S}_{mn}$.) We require that

$$\sum_{i=1}^{k'+1} \delta\mathbf{x}^{a(i)}_{mn} = \mathbf{0}, \tag{4.37}$$

which, by virtue of (4.34), (4.35), and (from (4.6) and (4.8))

$$\sum_{i=1}^{k'+1} \delta\mathbf{x}_{mn}^{b(i)} = \mathbf{0},$$ (4.38)

implies

$$\sum_{i=1}^{k'+1} \delta\mathbf{x}_{mn}^{a(i)(\|)} = \sum_{i=1}^{k'+1} \mathbf{Q}_{mn}\delta\hat{\mathbf{x}}_{mn}^{a(i)} = \mathbf{Q}_{mn}\sum_{i=1}^{k'+1} \delta\hat{\mathbf{x}}_{mn}^{a(i)} = \mathbf{0}.$$ (4.39)

Thus we require that

$$\sum_{i=1}^{k'+1} \delta\hat{\mathbf{x}}_{mn}^{a(i)} = \mathbf{0}.$$ (4.40)

In addition, $\hat{\mathbf{P}}_{mn}^a$ is given by

$$\hat{\mathbf{P}}_{mn}^a = k'^{-1}\sum_{i=1}^{k'+1} \delta\hat{\mathbf{x}}_{mn}^{a(i)}\big(\delta\hat{\mathbf{x}}_{mn}^{a(i)}\big)^T.$$ (4.41)

Hence the local analysis state $\bar{\mathbf{x}}_{mn}^a$ (determined in Section 4.3) is the mean over the local analysis ensemble $\{\mathbf{x}_{mn}^{a(i)}\}$, and, by (4.41), $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$ gives a representation of the local analysis error covariance matrix. We now turn to the task of determining the analysis perturbations $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$. Once these are known $\{\mathbf{x}_{mn}^{a(i)}\}$ is determined from (4.36).

4.4.1 Determining the ensemble of local analysis perturbations

There are many choices for $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$ that satisfy (4.40) and (4.41), and in this section we will describe possible methods for computing a set of solutions to these equations. See also Tippett et al. (2002) for different approaches to this problem in the global setting. In a given forecasting scenario, one could compare the accuracy and speed of these methods in order to choose among them. There are two main criteria we have in mind in formulating these methods.

First, the method for computing $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$ should be numerically stable and efficient. Second, since we wish to specify global fields that we think of as being similar to physical fields, we desire that these fields be slowly varying in m and n. That is, if $\hat{\mathbf{P}}_{mn}^a$ is slowly varying, we do not want to introduce any artificial rapid variations in the individual $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ through our method of constructing a solution of (4.40) and (4.41). For this purpose we regard the background vectors as physical states, and hence slowly varying in m and n. (This is reasonable since the background ensemble is obtained from evolution of the atmospheric model from time $t - \Delta t$ to time t.)

Thus we are motivated to express the analysis ensemble vectors $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ as formally linearly related to the background ensemble vectors. We consider two possible methods for doing this. In the first method, we relate $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ to the background vector with the same label i,

$$\delta\hat{\mathbf{x}}_{mn}^{a(i)} = \mathbf{Z}_{mn}\delta\hat{\mathbf{x}}_{mn}^{b(i)}, \tag{4.42}$$

where

$$\delta\hat{\mathbf{x}}_{mn}^{b(i)} = \mathbf{Q}_{mn}^T \delta\mathbf{x}_{mn}^{b(i)}. \tag{4.43}$$

(Note that the apparent linear relation between the background and analysis in (4.42) is only formal, since our solution for $\mathbf{Z}_{mn}$ will depend on the background.) In the second method, we formally express $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ as a linear combination of the vectors, $\delta\hat{\mathbf{x}}_{mn}^{b(1)}, \delta\hat{\mathbf{x}}_{mn}^{b(2)}, \cdots, \delta\hat{\mathbf{x}}_{mn}^{b(k'+1)}$,

$$\hat{\mathbf{X}}_{mn}^a = \hat{\mathbf{X}}_{mn}^b \mathbf{Y}_{mn}. \tag{4.44}$$

where

$$\hat{\mathbf{X}}_{mn}^{a,b} = \left(k'\right)^{-1/2}\left\{\delta\hat{\mathbf{x}}_{mn}^{a,b(1)}|\delta\hat{\mathbf{x}}_{mn}^{a,b(2)}|\cdots|\delta\hat{\mathbf{x}}_{mn}^{a,b(k'+1)}\right\} \tag{4.45}$$

Using (4.45) the analysis and the background covariance matrices can be expressed as

$$\hat{\mathbf{P}}^{a,b} = \hat{\mathbf{X}}_{mn}^{a,b} \hat{\mathbf{X}}_{mn}^{a,bT}. \tag{4.46}$$

The $k \times k$ matrix $\mathbf{Z}_{mn}$ or the $(k'+1) \times (k'+1)$ matrix $\mathbf{Y}_{mn}$ can be thought of as a generalized 'rescaling' of the original background fields. This 'rescaling' can be viewed as being similar to the techniques employed in the breeding method (Toth and Kalnay, 1993) and in the Ensemble Transform Kalman Filter approach (Bishop et al., 2001; Wang and Bishop, 2002). If $\mathbf{Z}_{mn}$ or $\mathbf{Y}_{mn}$, respectively, vary slowly with m and n, then by (4.42) and (4.44) so will $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$.

Considering (4.42), we see that (4.40) is automatically satisfied because by (4.38) and (4.43) the background perturbations $\delta\hat{\mathbf{x}}_{mn}^{b(i)}$ sum to zero,

$$\hat{\mathbf{X}}_{mn}^{b}\mathbf{v} = \mathbf{0}, \tag{4.47}$$

where $\mathbf{v}$ is a column vector of $(k'+1)$ ones. The analysis perturbations given by (4.42) will satisfy (4.41) [equivalently (4.46)] if, and only if,

$$\hat{\mathbf{P}}_{mn}^{a} = \mathbf{Z}_{mn}\hat{\mathbf{P}}_{mn}^{b}\mathbf{Z}_{mn}^{T}. \tag{4.48}$$

Considering (4.44), we see that (4.46) yields the following equation for $\mathbf{Y}_{mn}$

$$\hat{\mathbf{P}}_{mn}^{a} = \hat{\mathbf{X}}_{mn}^{b}\mathbf{Y}_{mn}\mathbf{Y}_{mn}^{T}\hat{\mathbf{X}}_{mn}^{bT} \tag{4.49}$$

Unlike (4.42) for $\mathbf{Z}_{mn}$, (4.44) does not imply automatic satisfaction of (4.40). We note that (4.40) can be written as

$$\hat{\mathbf{X}}_{mn}^{a}\mathbf{v} = \mathbf{0}. \tag{4.50}$$

Thus, in addition to (4.49), we demand that $\mathbf{Y}_{mn}$ must also satisfy

$$\hat{\mathbf{X}}_{mn}^{b}\mathbf{Y}_{mn}\mathbf{v} = \mathbf{0}. \tag{4.51}$$

Equation (4.48) has infinitely many solutions for $\mathbf{Z}_{mn}$. Similarly, equations (4.49) and (4.51) have infinitely many solutions for $\mathbf{Y}_{mn}$. In order for the results to vary slowly from one grid point to the next, it is important that we use an algorithm for computing a particular solution that depends continuously on $\hat{\mathbf{P}}_{mn}^a$ and $\hat{\mathbf{P}}_{mn}^b$.

4.4.2 Solutions of Equation (4.48)

Solution 1

One solution $\mathbf{Z}_{mn}$ is

$$\mathbf{Z}_{mn} = \left(\hat{\mathbf{P}}_{mn}^a\right)^{1/2}\left(\hat{\mathbf{P}}_{mn}^b\right)^{-1/2}, \tag{4.52}$$

where in (4.52), by the notation $\mathbf{M}^{1/2}$, we mean the *unique* positive symmetric square root of the positive symmetric matrix $\mathbf{M}$. In terms of the eigenvectors and eigenvalues of $\mathbf{M}$, the positive symmetric square root is

$$\mathbf{M}^{1/2} = \sum_{j=1}^{k} \sqrt{\nu^{(j)}}\mathbf{m}^{(j)}\left(\mathbf{m}^{(j)}\right)^{T}, \tag{4.53}$$

where

$$\mathbf{M}\mathbf{m}^{(j)} = \nu^{(j)}\mathbf{m}^{(j)}. \tag{4.54}$$

Recall that $\hat{\mathbf{P}}_{mn}^b$ is diagonal, so that its inverse square root in (4.52) is easily computed.

Solution 2

Pre- and post-multiplying (4.48) by $\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}$ and taking $\mathbf{Z}_{mn}$ to be symmetric,

$$\left[\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}\mathbf{Z}_{mn}\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}\right]^{2} = \left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}\hat{\mathbf{P}}_{mn}^a\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}. \tag{4.55}$$

Taking the positive symmetric square root of (4.55), we obtain a second possible solution of (4.48),

$$\mathbf{Z}_{mn} = \left(\hat{\mathbf{P}}_{mn}^b\right)^{-1/2}\left[\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}\hat{\mathbf{P}}_{mn}^a\left(\hat{\mathbf{P}}_{mn}^b\right)^{1/2}\right]^{1/2}\left(\hat{\mathbf{P}}_{mn}^b\right)^{-1/2}. \qquad (4.56)$$

In contrast to solution 1 (given by (4.52)) and solutions 3 (given below), this solution yields a $\mathbf{Z}_{mn}$ that is symmetric, $\mathbf{Z}_{mn} = \mathbf{Z}_{mn}^T$.

Family of solutions

We can create a family of solutions for $\mathbf{Z}_{mn}$ by introducing an arbitrary positive symmetric matrix $\mathbf{D}_{mn}$ and by pre- and post-multiplying (4.48) by $\mathbf{D}_{mn}^{-1/2}$. This yields

$$\tilde{\mathbf{P}}_{mn}^a = \tilde{\mathbf{Z}}_{mn}\tilde{\mathbf{P}}_{mn}^b\tilde{\mathbf{Z}}_{mn}^T, \qquad (4.57)$$

where

$$\tilde{\mathbf{Z}}_{mn} = \mathbf{D}_{mn}^{-1/2}\mathbf{Z}_{mn}\mathbf{D}_{mn}^{1/2}, \qquad (4.58)$$

$$\tilde{\mathbf{P}}_{mn}^{a,b} = \mathbf{D}_{mn}^{-1/2}\hat{\mathbf{P}}_{mn}^{a,b}\mathbf{D}_{mn}^{-1/2}. \qquad (4.59)$$

Applying solution 2 to (4.57) we obtain (4.56) with $\mathbf{Z}_{mn}$ and $\hat{\mathbf{P}}_{mn}^{a,b}$ replaced by $\tilde{\mathbf{Z}}_{mn}$ and $\tilde{\mathbf{P}}_{mn}^{a,b}$, Then, applying (4.58) and (4.59), we find that the unique solution to (4.48) such that $\mathbf{D}_{mn}^{-1/2}\mathbf{Z}_{mn}\mathbf{D}_{mn}^{1/2}$ is symmetric is

$$\mathbf{Z}_{mn} = \mathbf{D}_{mn}^{1/2}\left(\tilde{\mathbf{P}}_{mn}^b\right)^{-1/2}\left[\left(\tilde{\mathbf{P}}_{mn}^b\right)^{-1/2}\tilde{\mathbf{P}}_{mn}^a\left(\tilde{\mathbf{P}}_{mn}^b\right)^{-1/2}\right]^{1/2}\left(\tilde{\mathbf{P}}_{mn}^b\right)^{-1/2}\mathbf{D}_{mn}^{-1/2}. \quad (4.60)$$

Thus for any choice of $\mathbf{D}_{mn}$ we obtain a solution $\mathbf{Z}_{mn}$ of (4.48), and this is the unique solution for $\mathbf{Z}_{mn}$ subject to the added condition that $\mathbf{D}_{mn}^{-1/2}\mathbf{Z}_{mn}\mathbf{D}_{mn}^{1/2}$ is symmetric.

Another way to generate a family of solutions is to replace (4.52) by

$$\mathbf{Z}_{mn} = \sqrt{\hat{\mathbf{P}}_{mn}^a}\sqrt{\hat{\mathbf{P}}_{mn}^{b-1}}^T, \qquad (4.61)$$

where for a positive definite symmetric matrix $\mathbf{M}$, we mean by $\sqrt{\mathbf{M}}$ any matrix for which $\sqrt{\mathbf{M}}\sqrt{\mathbf{M}}^T = \mathbf{M}$. Note that this equation does not uniquely determine $\sqrt{\mathbf{M}}$, and that given any solution $\sqrt{\mathbf{M}} = \mathbf{W}$, the most general solution is $\sqrt{\mathbf{M}} = \mathbf{WO}$ where $\mathbf{O}$ is any orthogonal matrix. In particular, the positive symmetric square root (which we denote $\mathbf{M}^{1/2}$) is a specific choice for $\sqrt{\mathbf{M}}$, and, in general, $\sqrt{\mathbf{M}} = \mathbf{M}^{1/2}\mathbf{O}$. Furthermore, by considering all possible matrices $\sqrt{\hat{\mathbf{P}}^a_{mn}}$ we obtain all possible solutions $\mathbf{Z}_{mn}$ of (4.48). Thus we can write a general solution of (4.48) as

$$\mathbf{Z}_{mn} = \left(\hat{\mathbf{P}}^a_{mn}\right)^{1/2}\mathbf{O}_{mn}\left(\hat{\mathbf{P}}^{b-1}_{mn}\right)^{1/2}, \tag{4.62}$$

where $\mathbf{O}_{mn}$ is an arbitrary orthogonal matrix. (Note that $\mathbf{O}_{mn}$ can be a function of $\hat{\mathbf{P}}^a_{mn}$ and $\hat{\mathbf{P}}^b_{mn}$.) The family of solutions of (4.48) given by (4.60) with different $\mathbf{D}_{mn}$ is smaller than the family given by (4.62) with different $\mathbf{O}_{mn}$. In particular, the family (4.62), being the most general solution of (4.48), must contain the family corresponding to (4.60). To see that the later family is indeed smaller than the former family, consider the special case, $\hat{\mathbf{P}}^a_{mn} = \hat{\mathbf{P}}^b_{mn}$. For $\hat{\mathbf{P}}^a_{mn} = \hat{\mathbf{P}}^b_{mn}$, (4.60) always gives $\mathbf{Z}_{mn} = \mathbf{I}$, while (4.62) gives

$$\mathbf{Z}_{mn} = \left(\hat{\mathbf{P}}^a_{mn}\right)^{1/2}\mathbf{O}^{(o)}_{mn}\left(\hat{\mathbf{P}}^a_{mn}\right)^{1/2}, \tag{4.63}$$

which is never $\mathbf{I}$ unless the orthogonal matrix $\mathbf{O}^{(o)}_{mn}$ is $\mathbf{I}$. (Here $\mathbf{O}^{(o)}_{mn}$ denotes $\mathbf{O}_{mn}$ evaluated at $\hat{\mathbf{P}}^a_{mn} = \hat{\mathbf{P}}^b_{mn}$.) Based on our treatment in section 4.4.3, we believe that the smaller family, given by (4.60) with different $\mathbf{D}_{mn}$, gives results for $\hat{\mathbf{X}}^a_{mn}$ that are more likely to be useful for our purposes.

Solution 3

Subsequently, special interest will attach to the choices $\mathbf{D}_{mn} = \hat{\mathbf{P}}^b_{mn}$ and $\mathbf{D}_{mn} = \hat{\mathbf{P}}^a_{mn}$ in (4.60). Although these two choices yield results from (4.59) and (4.60)

that appear to be of quite different form, the two results for $\mathbf{Z}_{mn}$ are in fact the same. We call this solution for $\mathbf{Z}_{mn}$ solution 3. To see that these two $\mathbf{D}_{mn}$ choices yield the same $\mathbf{Z}_{mn}$, we note that (4.48) can be put in the form,

$$
\begin{aligned}
\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{1/2}&\left[\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{-1/2}\mathbf{Z}_{mn}\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{1/2}\right]^{-1}\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{1/2} = \\
&= \left(\hat{\mathbf{P}}_{mn}^{b}\right)^{1/2}\left[\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{-1/2}\mathbf{Z}_{mn}\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{1/2}\right]^{T}\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{1/2}.
\end{aligned} \tag{4.64a}
$$

Thus symmetry of $\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{-1/2}\mathbf{Z}_{mn}\left(\hat{\mathbf{P}}_{mn}^{a}\right)^{1/2}$ (i.e., $\mathbf{D}_{mn} = \mathbf{P}_{mn}^{a}$ in (4.60)) implies symmetry of $\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{-1/2}\mathbf{Z}_{mn}\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{1/2}$ (i.e., $\mathbf{D}_{mn} = \hat{\mathbf{P}}_{mn}^{b}$ in (4.60)) and vice versa. Hence, the two choices for $\mathbf{D}_{mn}$ necessarily yield the same $\mathbf{Z}_{mn}$. Explicitely, setting $\mathbf{D}_{mn} = \hat{\mathbf{P}}_{mn}^{b}$ in (4.60) we can write solution 3 as

$$
\mathbf{Z}_{mn} = \left(\hat{\mathbf{P}}_{mn}^{b}\right)^{1/2}\left[\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{-1/2}\hat{\mathbf{P}}_{mn}^{a}\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{-1/2}\right]^{1/2}\left(\hat{\mathbf{P}}_{mn}^{b}\right)^{-1/2}. \tag{4.65}
$$

4.4.3 'Optimal' choices for $\mathbf{Z}_{mn}$

Since we think of the background ensemble members as physical fields, it is reasonable to seek to choose the analysis ensemble members $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ in such a way as to minimize their difference,

$$
\mathcal{F}(\delta\hat{\mathbf{x}}_{mn}^{a(i)}) = \sum_{i=1}^{k'+1}\left\|\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right\|^{2} = \sum_{i=1}^{k'+1}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right]^{T}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right], \tag{4.66}
$$

with the analysis ensemble members $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$, subject to the requirement that (4.41) be satisfied. Thus, introducing a $k \times k$ matrix $\mathbf{B}_{mn}$ of Lagrange multipliers, we form the following quantity,

$$
\begin{aligned}
\mathcal{L} = &\sum_{i=1}^{k'+1}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right]^{T}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right] - \\
&- \sum_{p,q=1}^{k}(\mathbf{B}_{mn})_{p,q}\left[\left(\hat{\mathbf{P}}_{mn}^{a}\right)_{p,q} - \frac{1}{k'}\sum_{i=1}^{k'+1}\left(\delta\hat{\mathbf{x}}_{mn}^{a(i)}\right)_{p}\left(\delta\hat{\mathbf{x}}_{mn}^{a(i)}\right)_{q}\right]
\end{aligned} \tag{4.67a}
$$

79

which we minimize with respect to $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$ and $\mathbf{B}_{mn}$. Forming the first and second derivatives of $\mathcal{L}$ with respect to $\delta\hat{\mathbf{x}}_{mn}^{a(i)}$, we have

$$\frac{1}{2}\frac{\partial\mathcal{L}}{\partial\delta\hat{\mathbf{x}}_{mn}^{a(i)}} = \mathbf{Z}_{mn}^{-1}\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}, \tag{4.68}$$

$$\frac{1}{2}\frac{\partial^2\mathcal{L}}{\partial\delta\hat{\mathbf{x}}_{mn}^{a(i)}\partial\delta\hat{\mathbf{x}}_{mn}^{a(i)}} = \mathbf{Z}_{mn}^{-1}, \tag{4.69}$$

where we have defined $\mathbf{Z}_{mn}^{-1}$ as

$$\mathbf{Z}_{mn}^{-1} = \mathbf{I} + \frac{1}{2k'}\left(\mathbf{B}_{mn} + \mathbf{B}_{mn}^T\right). \tag{4.70}$$

Since $\mathcal{L}$ is stationary, (4.68) implies (4.42), and the derivative with respect to $\mathbf{B}_{mn}$ returns (4.41). Since $\mathcal{L}$ is minimum, (4.69) implies that $\mathbf{Z}_{mn}$ is positive, while (4.70) gives $\mathbf{Z}_{mn} = \mathbf{Z}_{mn}^T$. Thus the solution that minimizes $\mathcal{F}(\hat{\mathbf{x}}_{mn}^{a(i)})$ is obtained from the *unique* symmetric positive solution for $\mathbf{Z}_{mn}$. This is given by solution 2 (4.56).

It is also of interest to consider different metrics for the distance between the analysis ensemble $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$ and the background ensemble $\{\delta\hat{\mathbf{x}}_{mn}^{b(i)}\}$. Thus we minimize the quadratic form,

$$\mathcal{F}_D(\delta\hat{\mathbf{x}}_{mn}^{a(i)}) = \sum_{i=1}^{k'+1}\left\|\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right\|_D^2 = \sum_{i=1}^{k'+1}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right]^T\mathbf{D}_{mn}^{-1}\left[\delta\hat{\mathbf{x}}_{mn}^{a(i)} - \delta\hat{\mathbf{x}}_{mn}^{b(i)}\right], \tag{4.71}$$

where the positive symmetric matrix $\mathbf{D}_{mn}$ specifies the metric. (The quadratic form $\mathcal{F}(\delta\hat{\mathbf{x}}_{mn}^{a(i)})$ is the special case of $\mathcal{F}_D(\delta\hat{\mathbf{x}}_{mn}^{a(i)})$ when the metric is defined by the identity matrix, $\mathbf{D}_{mn} = \mathbf{I}$). The introduction of the metric matrix $\mathbf{D}_{mn}$ is equivalent to making the change of variables, $\tilde{\mathbf{X}}_{mn}^{a,b} = (\mathbf{D}_{mn})^{-1/2}\hat{\mathbf{X}}_{mn}^{a,b}$. Inserting this change of variables in (4.56), we obtain (4.60).

Solution 3, namely $\mathbf{D}_{mn}$ equal to $\hat{\mathbf{P}}_{mn}^b$ or $\hat{\mathbf{P}}_{mn}^a$, appears to be favorable in that it provides a natural intuitive normalizations for the distance. We thus conjecture that solutions 3 may yield better performance than solutions 1 and 2.

4.4.4 Solution of (4.49) and (4.51)

Another way of solving for the analysis fields is to use the 'Potter method' (e.g., Biermann 1977). To see how this solution is obtained, let

$$\mathbf{A}_{mn} = \mathbf{Y}_{mn}\mathbf{Y}_{mn}^T \tag{4.72}$$

so that (4.49) becomes

$$\hat{\mathbf{P}}_{mn}^a = \hat{\mathbf{X}}_{mn}^b \mathbf{A}_{mn} \hat{\mathbf{X}}_{mn}^{bT} \tag{4.73}$$

Because $\hat{\mathbf{P}}_{mn}^a$ is $k \times k$ and $\mathbf{A}_{mn}$ is $(k'+1) \times (k'+1)$, there is a lot of freedom in choosing $\mathbf{A}_{mn}$. It seems reasonable that, if the analysis covariance and the background covariance are the same (i.e., $\hat{\mathbf{P}}_{mn}^a = \hat{\mathbf{P}}_{mn}^b$), then the ensemble analysis perturbations should be set equal to the ensemble background perturbations:

$$\mathbf{Y}_{mn} = \mathbf{I} \qquad \text{if} \qquad \hat{\mathbf{P}}_{mn}^a = \hat{\mathbf{P}}_{mn}^b. \tag{4.74}$$

A solution for $\mathbf{A}_{mn}$ consistent with (4.72)-(4.74) is

$$\mathbf{A}_{mn} = \mathbf{I} + \hat{\mathbf{X}}_{mn}^{bT} \hat{\mathbf{P}}_{mn}^{b-1} \big[\hat{\mathbf{P}}_{mn}^a - \hat{\mathbf{P}}_{mn}^b\big] \hat{\mathbf{P}}_{mn}^{b-1} \hat{\mathbf{X}}_{mn}^b. \tag{4.75}$$

This solution for $\mathbf{A}_{mn}$ is symmetric and can also be shown to be positive definite. Equation (4.75) yields $\mathbf{A}_{mn} = \mathbf{I}$ if $\hat{\mathbf{P}}_{mn}^a = \hat{\mathbf{P}}_{mn}^b$, as required by (4.72) and (4.74), and satisfaction of (4.73) by (4.75) can be verified by direct substitution and making use of $\hat{\mathbf{P}}_{mn}^b = \hat{\mathbf{X}}_{mn}^b \hat{\mathbf{X}}_{mn}^{bT}$. From (4.72) we have $\mathbf{Y}_{mn} = \sqrt{\mathbf{A}_{mn}}$, and, if the positive symmetric square root is chosen, then (4.74) is satisfied. Thus we have as a possible solution

$$\mathbf{Y}_{mn} = (\mathbf{A}_{mn})^{1/2}. \tag{4.76}$$

It remains to show that (4.75) and (4.76) also satisfies (4.51). By (4.75) and (4.47) we have $\mathbf{A}_{mn}\mathbf{v} = \mathbf{v}$; i.e., $\mathbf{v}$ is an eigenvector of $\mathbf{A}_{mn}$ with eigenvalue one.

Since the positive square root is employed in (4.76) $\mathbf{v}$ is also an eigenvector of $\mathbf{Y}_{mn}$ with eigenvalue one. Hence $\mathbf{X}^b_{mn}\mathbf{Y}_{mn}\mathbf{v} = \mathbf{X}^b_{mn}\mathbf{v}$, which is identically zero by (4.47), thus satisfying (4.51).

Potter's expression for $\mathbf{A}_{mn}$ is obtained by using (4.28) and (4.30) in (4.74) and (4.76),

$$\mathbf{A}_{mn} = \mathbf{I} - \hat{\mathbf{X}}^{bT}_{mn}\hat{\mathbf{H}}^T_{mn}\left[\hat{\mathbf{H}}_{mn}\hat{\mathbf{P}}^b_{mn}\hat{\mathbf{H}}^T_{mn} + \mathbf{R}_{mn}\right]^{-1}\hat{\mathbf{H}}_{mn}\hat{\mathbf{X}}^b_{mn}. \tag{4.77}$$

For (4.76) and (4.77) the square root is taken of a $k'+1$ by $k'+1$ matrix, but the inverse is of an s by s matrix, where s is the dimension of the local observation space. An equivalent way to write (4.77) in our setting is

$$\mathbf{A}_{mn} = \mathbf{I} - \hat{\mathbf{X}}^{bT}_{mn}\hat{\mathbf{V}}_{mn}\hat{\mathbf{H}}^T_{mn}\mathbf{R}^{-1}_{mn}\hat{\mathbf{H}}_{mn}\hat{\mathbf{X}}^b_{mn}, \tag{4.78}$$

where

$$\hat{\mathbf{V}} = \left[\mathbf{I} + \hat{\mathbf{H}}^T_{mn}\mathbf{R}^{-1}_{mn}\hat{\mathbf{H}}_{mn}\hat{\mathbf{P}}^b_{mn}\right]^{-1}. \tag{4.79}$$

Now aside from $\mathbf{R}_{mn}$, we need only invert a k by k matrix. As previously discussed, although $\mathbf{R}_{mn}$ is s by s, its inverse is easily computed even when s is much larger than k.

We now ask whether different solutions of (4.48) for $\mathbf{Z}_{mn}$ have corresponding $\mathbf{Y}_{mn}$ such that $\mathbf{Z}_{mn}\hat{\mathbf{X}}^b_{mn}$ and $\hat{\mathbf{X}}^b_{mn}\mathbf{Y}_{mn}$ yield the same result for $\hat{\mathbf{X}}^a_{mn}$. To see that they do, we note that the matrix $\hat{\mathbf{X}}^b_{mn}$ (which consists of k rows and $k'+1$ columns) has a (nonunique) right inverse $\left(\hat{\mathbf{X}}^b_{mn}\right)^{-1}$ such that $\hat{\mathbf{X}}^b_{mn}\left(\hat{\mathbf{X}}^b_{mn}\right)^{-1} = \mathbf{I}_k$, where

$$\left(\hat{\mathbf{X}}^b_{mn}\right)^{-1} = \hat{\mathbf{X}}^{bT}_{mn}\left(\hat{\mathbf{X}}^b_{mn}\hat{\mathbf{X}}^{bT}_{mn}\right)^{-1} + \mathbf{E}_{mn} = \hat{\mathbf{X}}^{bT}_{mn}\left(\hat{\mathbf{P}}^b_{mn}\right)^{-1} + \mathbf{E}_{mn}, \tag{4.80}$$

and $\mathbf{E}_{mn}$ is any $k \times (k'+1)$ matrix for which $\hat{\mathbf{X}}^b_{mn}\mathbf{E}_{mn} = \mathbf{0}_{mn}$. Thus, from $\hat{\mathbf{X}}^a_{mn} = \mathbf{Z}_{mn}\hat{\mathbf{X}}^b_{mn}$, we have

$$\hat{\mathbf{X}}^a_{mn} = \hat{\mathbf{X}}^b_{mn}\left(\hat{\mathbf{X}}^b_{mn}\right)^{-1}\mathbf{Z}_{mn}\hat{\mathbf{X}}^b_{mn}. \tag{4.81}$$

From the definition of $\mathbf{Y}_{mn}$, $\hat{\mathbf{X}}^a_{mn} = \hat{\mathbf{X}}^b_{mn}\mathbf{Y}_{mn}$, we see that (4.81) and (4.80) yields

$$\mathbf{Y}_{mn} = \hat{\mathbf{X}}^{bT}_{mn}\left(\hat{\mathbf{P}}^b_{mn}\right)^{-1}\mathbf{Z}_{mn}\hat{\mathbf{X}}^b_{mn} + \mathbf{G}_{mn}, \tag{4.82}$$

where $\mathbf{G}_{mn}$ is any $(k'+1)$ by $(k'+1)$ matrix satisfying $\hat{\mathbf{X}}^b_{mn}\mathbf{G}_{mn} = \mathbf{0}$. Since we desire that $\mathbf{Y}_{mn} = \mathbf{I}_{k'+1}$, when $\mathbf{Z}_{mn} = \mathbf{I}_k$, a possible choice for $\mathbf{G}_{mn}$ is

$$\mathbf{G}_{mn} = \mathbf{I}_{k'+1} - \hat{\mathbf{X}}^{bT}_{mn}\left(\hat{\mathbf{P}}^b_{mn}\right)^{-1}\hat{\mathbf{X}}^b_{mn}. \tag{4.83}$$

(We note that $\mathbf{G}_{mn}$ given by (4.83) is a projection operator (i.e., an idempotent matrix), $(\mathbf{G}_{mn})^p = \mathbf{G}_{mn}$ for any integer exponent p.) Thus from (4.82) and (4.83), a $\mathbf{Y}_{mn}$ corresponding to any solution $\mathbf{Z}_{mn}$ (e.g., solution 1, 2 or 3) is

$$\mathbf{Y}_{mn} = \hat{\mathbf{X}}^{bT}_{mn}\left(\hat{\mathbf{P}}^b_{mn}\right)^{-1}(\mathbf{Z}_{mn} - \mathbf{I}_k)\hat{\mathbf{X}}^b_{mn} + \mathbf{I}_{k'+1}. \tag{4.84}$$

Using (4.84), (4.48), and (4.47) it can be verified that $\mathbf{Y}_{mn}\mathbf{Y}^T_{mn} = \mathbf{A}_{mn}$ with $\mathbf{A}_{mn}$ given by (4.75). Thus $\mathbf{Y}_{mn}\mathbf{Y}^T_{mn}$ is the *same* $(k'+1)\times(k'+1)$ matrix for all solutions $\mathbf{Z}_{mn}$ (e.g., solutions 1,2, and 3). The general solution of $\mathbf{Y}_{mn}\mathbf{Y}^T_{mn} = \mathbf{A}_{mn}$ is

$$\mathbf{Y}_{mn} = \sqrt{\mathbf{A}_{mn}} = (\mathbf{A}_{mn})^{1/2}\mathbf{O}_{mn}, \tag{4.85}$$

where $\mathbf{O}_{mn}$ is an arbitrary orthogonal matrix. However, to ensure that (4.51) is satisfied we also require that $\mathbf{O}_{mn}\mathbf{v} = \pm\mathbf{v}$; i.e., that $\mathbf{v}$ is an eigenvector of $\mathbf{O}_{mn}$. For example, $\mathbf{O}_{mn}$ can be any rotation about $\mathbf{v}$. Thus there is still a large family of allowed orthogonal matrices $\mathbf{O}_{mn}$. (Note that $\mathbf{O}_{mn}$ can depend on $\hat{\mathbf{P}}^a_{mn}$ and $\hat{\mathbf{P}}^b_{mn}$, and that for (4.74) to be satisfied, $\mathbf{O}_{mn}$ must be $\mathbf{I}$ whenever $\hat{\mathbf{P}}^a_{mn} = \hat{\mathbf{P}}^b_{mn}$.) Hence we can think of the various solutions for $\mathbf{Y}_{mn}$ either as being generated by (4.60) and (4.84) with different choices for the metric matrix $\mathbf{D}_{mn}$, or as being generated by (4.75) and (4.85) with different choices for the orthogonal matrix $\mathbf{O}_{mn}$.

In Appendix B we show that the solution for which $\mathbf{Y}_{mn}$ is symmetric, (4.76), is the same as that given by solution 3 for $\mathbf{Z}_{mn}$. That is, $\mathbf{Z}_{mn}\hat{\mathbf{X}}_{mn}^{b}$ with $\mathbf{Z}_{mn}$ given by (4.65) and $\hat{\mathbf{X}}_{mn}^{b}\mathbf{Y}_{mn}$ with $\mathbf{Y}_{mn}$ given by (4.75) and (4.76) both yield the same result for $\hat{\mathbf{X}}_{mn}^{a}$. Also, in Appendix C we show that $\mathbf{Y}_{mn}$ as given by (4.84) can be used to directly obtain the analysis $\mathbf{X}_{mn}^{a}$ (note the absence of the superscribed circumflex on $\mathbf{X}_{mn}^{a}$).

4.4.5   Construction of the global fields

Regardless of which of these solution methods for $\{\delta\hat{\mathbf{x}}_{mn}^{a(i)}\}$ is chosen, by use of (4.36) we now have $(k'+1)$ local analyses $\mathbf{x}_{mn}^{a(i)}$ at each point $\mathbf{r}_{mn}$, and it now remains to construct an ensemble of global fields $\left\{\mathbf{x}^{a(i)}(\mathbf{r},t)\right\}$ that can be propagated forward in time to the next analysis time. There are various ways of doing this. One method (used in our numerical example of section 4.5) takes into account all the atmospheric states at the point $\mathbf{r}_{mn}$ obtained from each of the $(2l+1)^2$ local vectors $\mathbf{x}_{m-m',n-n'}^{a(i)}$ $(|m'| \leq l, |n'| \leq l)$ that include the point $\mathbf{r}_{mn}$. In particular, these $(2l+1)^2$ states at $\mathbf{r}_{mn}$ are averaged to obtain $\mathbf{x}^{a(i)}(\mathbf{r},t)$.

4.4.6   Variance inflation

In past work on ensemble Kalman filters (Anderson and Anderson 1999; Whitaker and Hamill 2002) it was found that inflating the covariance ($\mathbf{P}^{a}$ or $\mathbf{P}^{b}$) by a constant factor on each analysis step, leads to more stable and improved analyses. One rationale for doing this is to compensate for the effect of finite sample size, which can be shown to, on average, underestimate the covariance. In addition, in Section 4.5 and Appendix D we will investigate the usefulness of enhancing the probability of error in directions that formally show only very small error

84

probability. Following such a modification of $\hat{\mathbf{P}}^a_{mn}$ $[\hat{\mathbf{P}}^b_{mn}]$, for consistency, we also make modifications to the ensemble perturbations $\delta\hat{\mathbf{x}}^{a(i)}_{mn}$ $[\delta\hat{\mathbf{x}}^{b(i)}_{mn}]$ so as to preserve the relationship (4.41) [(4.46)]. (Again, similar to the discussion in Section 4.4.2, the choice of these modifications is not unique.)

In our numerical experiments in section 4.5 we will consider two methods of variance inflation. One method, which we refer to as *regular variance inflation*, multiplies all background perturbations $\delta\hat{\mathbf{x}}^{b(i)}_{mn}$ by a constant $(1+\delta)$. This corresponds to multiplying $\hat{\mathbf{P}}^b_{mn}$ by $(1+\delta)^2$. This method has been previously used by Anderson and Anderson (1999) and by Whitaker and Hamill (2002). In addition to this method, in Appendix D we introduce a second variance inflation method, which, as our results of section 4.5 indicate, may yield superior performance. We refer to this method as *enhanced variance inflation*.

4.5 Numerical experiments

4.5.1 40-variable Lorenz model

The skill of the proposed local ensemble Kalman Filter scheme is demonstrated by Observing System Simulation Experiments (OSSE's) carried out with the 40-variable Lorenz (L40) model (Lorenz 1996; Lorenz and Emanuel 1998),

$$\frac{dx_m}{dt} = (x_{m+1} - x_{m-2})x_{m-1} - x_m + F. \tag{4.86}$$

Here, $m = 1, \cdots, 40$, where $x_{-1} = x_{39}$, $x_0 = x_{40}$, and $x_{41} = x_1$. This model mimics the time evolution of an unspecified scalar meteorological quantity, x, at 40 grid points along a latitude circle. We solve (4.86) with a fourth-order Runge-Kutta time integration scheme with a time step of 0.05 non-dimensional

unit (which may be thought of as nominally equivalent to 6-h in real world time assuming that the characteristic time scale of dissipation in the atmosphere is 5-days; see Lorenz 1996 for details).

For our chosen forcing, $F = 8$, the steady state solution, $x_m = F$ for $m = 1, \cdots, 40$, in (4.86) is linearly unstable. This instability is associated with unstable dispersive waves characterized by westward (i.e., in the direction of decreasing m) phase velocities and eastward group velocities. Lorenz and Emanuel (1998) demonstrated by numerical experiments for $F = 8$ that the x field is dominated by a wave number 8 structure, and that the system is chaotic; it has 13 positive Lyapunov exponents, and its Lyapunov dimension (Kaplan and Yorke 1979) is 27.1. It can be expected that, due to the eastward group velocities, growing uncertainties in the knowledge of the model state propagate eastward. A similar process can be observed in operational numerical weather forecasts, where dispersive short (longitudinal wave number 6-9) Rossby waves, generated by baroclinic instabilities, play a key role in the eastward propagation of uncertainties (e.g., Persson 2000; Szunyogh et al. 2002; and Zimin et al. 2003).

4.5.2   Rms analysis error

The 40-variable Lorenz model was also used by Whitaker and Hamill (2002) to validate their ensemble square root filter (EnSRF) approach. In designing our OSSE's we follow their approach of first generating the 'true state', $x_m^t(t)$, $m = 1, \cdots, 40$, by a long (40,000 time-step) model integration; then creating 'observations' of all model variables at each time step by adding uncorrelated normally distributed random noise with unit variance to the 'true state' (i.e., $\mathbf{R}_{mn} = \mathbf{I}$). (The rms random observational noise variance of 1.00 is to be com-

pared with the value 3.61 of the time mean rms deviation of solutions, $x_m(t)$, of (4.86) from their mean.) The observations are assimilated at each time step, and the accuracy of the analysis is measured by the time mean of the rms error,

$$E = \left(\frac{1}{40} \sum_{m=1}^{40} (\bar{x}_m^a - x_m^t)^2 \right)^{1/2}. \tag{4.87}$$

Thus, our 'true state' and observations are generated in the same way as in Whitaker and Hamill (2002). Also, we use the same ensemble size as Whitaker and Hamill ($k' + 1 = 10$). Hence our analysis error results and theirs can be directly compared.

4.5.3 The 'optimal' error

For the sake of comparison with our local Kalman filter results, we now establish a standard that can be regarded as the best achievable result that could be obtained given that computer resources placed no constraint on computations of the analysis. (In contrast with operational weather prediction, for our simple 40-variable Lorenz model, this is indeed the case.) For this purpose, we considered the state $\mathbf{x}(t) = \big(x_1(t), x_2(t), \cdots, x_{40}(t)\big)$ on the entire domain rather than on a local patch. Then several Kalman filter runs were carried out with different numbers of ensemble members. In these integrations, full (k') rank estimates of the covariance matrices were considered and the ensemble perturbations were updated using (4.74), (4.76), and (4.90) of Appendix C. (Section). We found that at about $k' + 1 = 100$ ensemble members the time mean of E converged to 0.20 and that no reduction of error occurred upon further increase of k'. We refer to this value as the *optimal error*, and we regard it as a comparison standard for our local Kalman filter method. We now describe the implementation of our method on the 40-variable Lorenz-model.

4.5.4 Analysis results

From (4.26) we know that for our OSSE's ($\mathbf{R}_{mn} = \mathbf{I}$, $\hat{\mathbf{H}}_{mn} = \mathbf{Q}_{mn}$), the analysis error covariance matrix is $\hat{\mathbf{P}}^a_{mn} = \left[(\mathbf{P}^b_{mn})^{-1} + \mathbf{I} \right]^{-1}$. Thus $\hat{\mathbf{P}}^a_{mn}$ and $\hat{\mathbf{P}}^b_{mn}$ commute. In this case (4.52), (4.56), and (4.65) are identical. Hence solutions 1, 2, and 3 are the same solution. We implement this solution using (4.74) and (4.90) of Appendix C.

In our experiments, the local analysis covariance matrix is computed by (4.27) and the local analysis is obtained by (4.25). The analysis ensemble is updated by (4.74) and (4.90) of Appendix C, and the variance of the background ensemble is increased by a factor of $1 + \varepsilon$ in each step using the enhanced variance inflation algorithm (see Appendix D for detail). The final analysis at each point m is computed by averaging the $(2l+1)$ local analyses whose local regions include the point m (a one-dimensional version of section 4.4.6).

The free parameters of our scheme are the dimensionality of the local regions (which is $2l + 1$), the rank of the covariance matrices (k), and the coefficient (ε) in the enhanced variance inflation algorithm. In what follows, the sensitivity of the data assimilation scheme to the tunable free parameters is investigated by numerical experiments (k' is held fixed at $k' = 9$).

In the first experiment the variance inflation coefficient is constant, $\varepsilon = 0.012$, while the dimension of the local vectors $(2l + 1)$ and the rank (k) of the background covariance matrix are varied. The results are shown in Table 4.1. The scheme seems to be stable and accurate for a wide range of parameters. The optimal size local region consists of $2l + 1 = 11, 13$ grid points, at which rank $k = 5, 6, 7, 8, 9$ estimates of the background covariance matrix provide similarly accurate analyses. Moreover, rank 3 and 4 estimates lead to surprisingly accu-

k / $2l+1$	3	4	5	6	7	8	9
5	0.24	0.23					
7	0.22	0.22	0.21	0.22			
9	0.22	0.21	0.21	0.21	0.21	0.21	
11	D	D	0.20	0.20	0.20	0.20	0.20
13	D	D	0.20	0.20	0.20	0.20	0.20
15	D	D	D	0.22	0.20	0.2	0.20

Table 4.1: Dependence of the time mean rms error on the box size $(2l+1)$ and the rank (k) of the background covariance matrix. The symbol D stands for time mean rms errors larger than one, which is the rms mean of the observational errors. The coefficient of the enhanced variance inflation is $\epsilon = 0.012$.

rate analyses for the smaller size $(2l+1 = 5, 7, 9)$ local regions. This indicates that the background uncertainty in a local region at a given time $(\hat{\mathbf{P}}^b_{mn})$ can be well approximated in a low (k) dimensional linear space. Our premise, that the dimension of this space can be significantly lower than the number of ensemble members $(k' + 1)$ needed to evolve the uncertainty, proved to be correct for the L40 model. (We note that the local dimensionality k is also much smaller than the "global" Lyapunov-dimension, 27.1, of the system). On the practical side, this result suggests that, at least for the L40 model, the efficiency of the analysis scheme can be significantly improved by using ranks that are smaller than the dimension of the local vectors and the number of ensemble members. We note that our best results are at least as good as the best results published in Whitaker and Hamill (2002) and attain the optimal value (0.20) from section 4.5.3.

ϵ \ k	4	5	6	7	8	9
0.008	D	D	0.44	0.20	0.20	0.20
0.010	D	D	0.2	0.2	0.20	0.20
0.012	D	0.20	0.20	0.20	0.20	0.20
0.014	D	0.20	0.20	0.20	0.20	0.20
0.016	D	0.20	0.20	0.20	0.20	0.20
0.018	D	0.20	0.20	0.20	0.20	0.20
0.020	0.21	0.20	0.20	0.20	0.20	0.20

Table 4.2: Dependence of the time mean rms error on the coefficient (ε) of the enhanced variance inflation scheme and the rank (k) of the background covariance matrix. The meaning of D is the same as in Table 4.1. The window size is 13.

In the second experiment, the dimension of the local regions is constant ($2l + 1 = 13$), while the rank and the variance inflation coefficient are varying. The results are shown in Table 4.2 . While lower rank estimates of the background error covariance matrix require somewhat stronger variance inflation, the results are not sensitive to the choice of ε once it is larger than a critical value. (By critical value we mean the smallest ε that provides the optimal error).

The second experiment was then repeated by using the regular variance inflation of Anderson and Anderson (1999) and Whitaker and Hamill (2002). In the regular variance inflation, all background ensemble perturbations are multiplied by $r = 1 + \delta$, where δ is small, $1 \gg \delta > 0$. This inflation strategy increases the total variance in the background ensemble by a factor of $(1 + \Delta) = 1 + \delta^2 + 2\delta \approx 1 + 2\delta$. It can be seen from Table 4.5.4 that, except for $k = 4$,
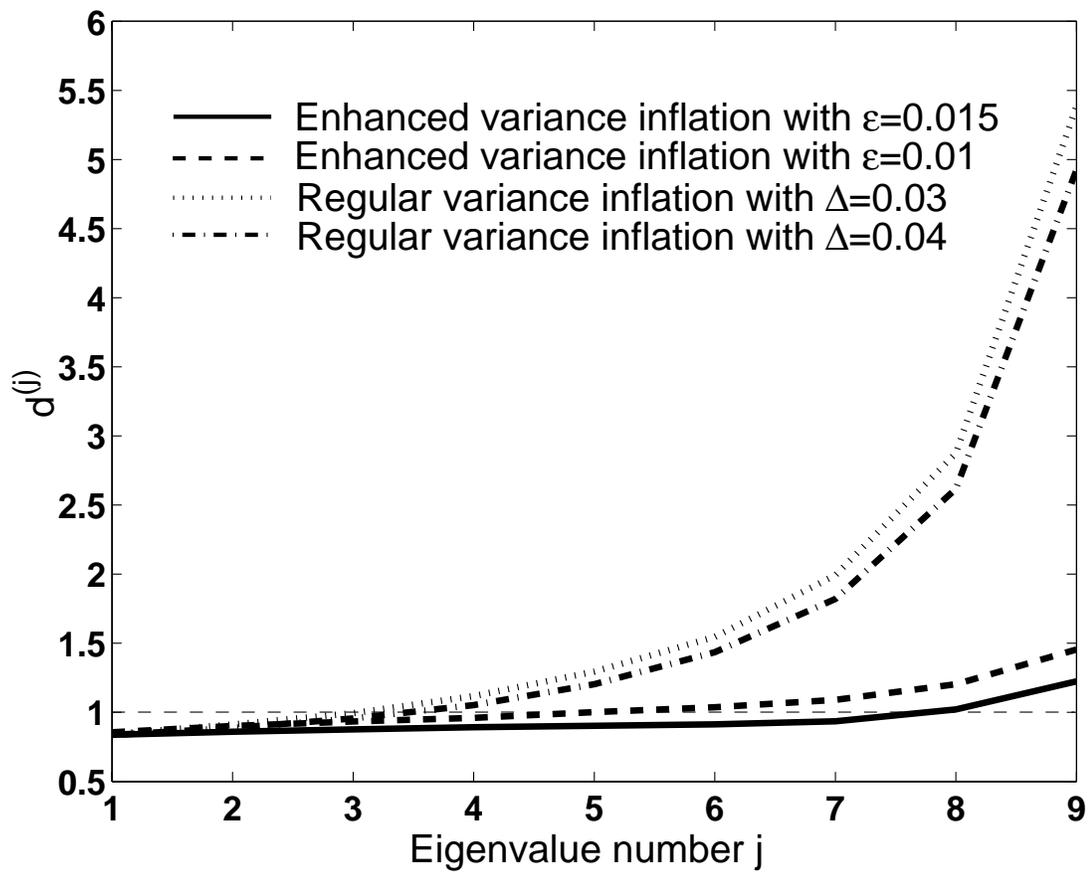
Figure 4.2: The ratio $d^{(j)}$ at $m = 1$ as function of j for two different values of ε and Δ.
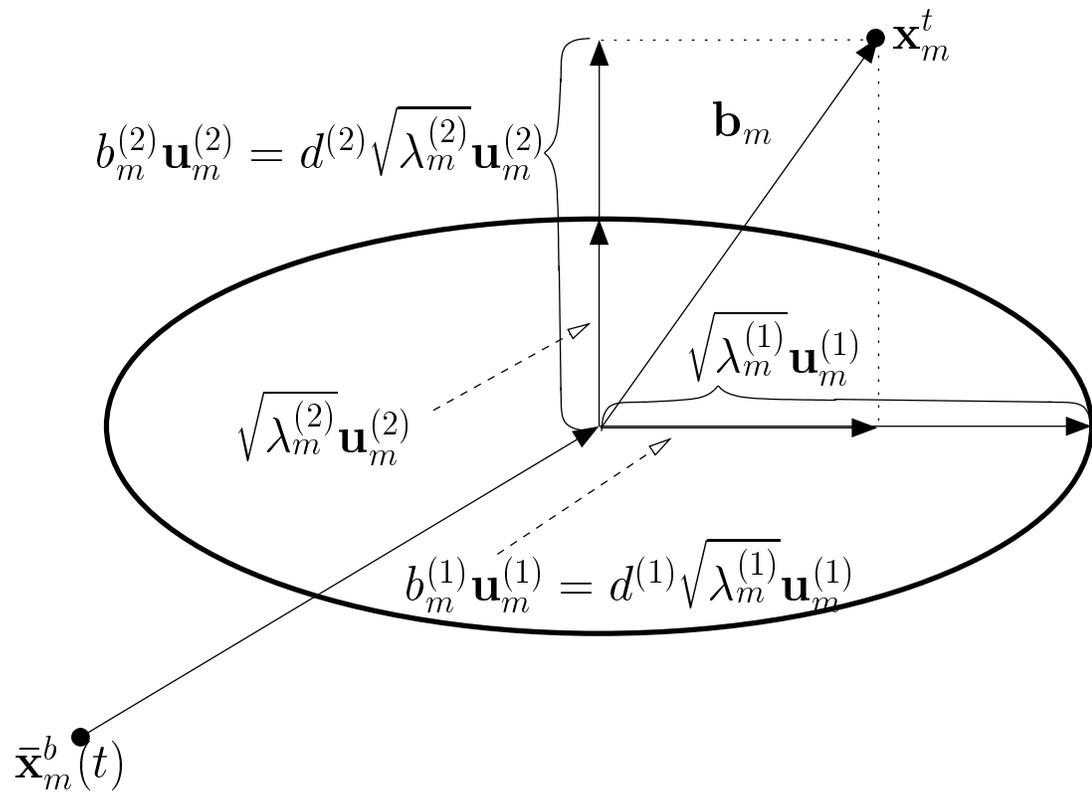
Figure 4.3: Projection of the true background error, $\mathbf{b}_m$ on the main axes of the probability ellipsoid.
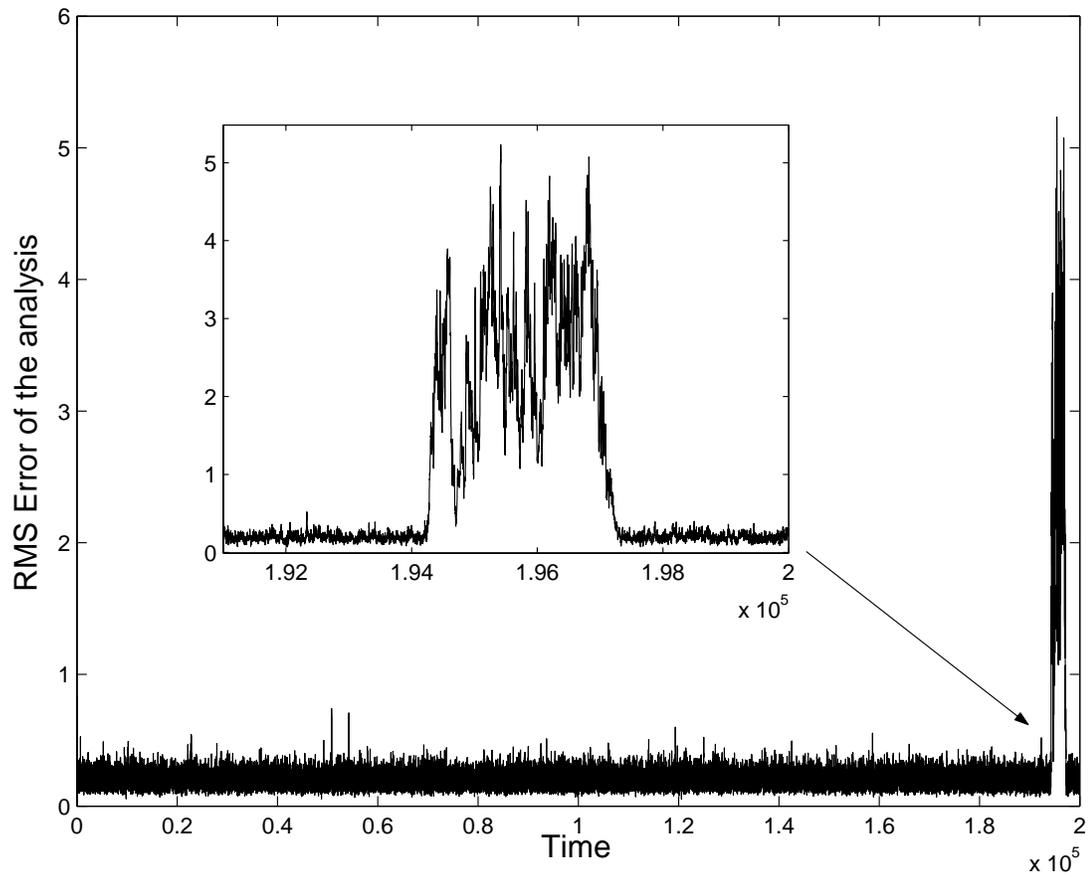
Figure 4.4: The time evolution of the rms analysis error. The inserted panel shows the time evolution of the spike in more detail.

the critical value of ε is less than half of the critical value of Δ. The main difference between the two inflation schemes is that the enhanced scheme inflates the dominant eigendirections of the background covariance matrix less aggressively, and the least dominant eigendirections more aggressively. The numerical results suggest that this feature of the scheme is beneficial, indicating that the ensemble-based estimate of the background error is more reliable in the more unstable directions than in the other directions. This is also well illustrated by the quantitative results shown in Figure 4.2. To explain this figure, we define the true background error, $\mathbf{b}_m = \mathbf{x}_m^t - \bar{\mathbf{x}}_m^b$ by the difference between the truth, $\mathbf{x}_m^t$ and the background mean, $\bar{\mathbf{x}}_m^b$. We also define $b_m^{(j)} = \mathbf{b}_m^T \mathbf{u}_m^{(j)}$, the component of $\mathbf{b}_m$ along the semi-axis of the probability ellipsoid, corresponding to the jth largest eigenvalue of $\mathbf{P}_m^b$, where $j = 1, 2, \cdots, k$. (The case $k = 2$ is illustrated in Figure 4.3.) For an ensemble that correctly estimates the uncertainty in each basis direction, the time means of

$$d_m^{(j)} = \left(b_m^{(j)^2}/\lambda_m^{(j)}\right)^{1/2}, \qquad j = 1, 2, \cdots, k, \tag{4.88}$$

should be close to one. When, for a given j, the ratio $d_m^{(j)}$ is smaller than one, the ensemble tends to overestimate the distance between the truth and the background in the $\mathbf{u}_m^{(j)}$ direction. When $d_m^{(j)}$ is larger than one, the ensemble underestimates this distance. Figure 4.2 shows that with the enhanced variance inflation the behavior of the ensemble is much better than with the regular variance inflation. This is especially true for the less dominant eigendirections, for which the ensemble with regular variance inflation significantly (by about a factor of 6) underestimates the distance between the truth and the mean background. We found that $\|\mathbf{b}_m\|^2 - \sum_{j=1}^{9} b_m^{(j)^2}$, the true background variance unexplained by the directions, $\mathbf{u}_m^{(j)}$, $j = 1, 2, \cdots, 9$; is about 3% of the true total variance ($\|\mathbf{b}_m\|^2$)

for all four cases shown in Figure 4.2. Thus the results indicate that the superior performance of the enhanced variance inflation is due to the better distribution of the variance between the resolved directions. We note that this advantage of the enhanced variance inflation could not be exploited if the analysis was not done in $\mathbb{S}_{mn}$ introduced in Section 4.2.

An interesting feature is the anomalously large error value of 0.29 at $\Delta = 0.036$, $k = 8$ in Table 4.5.4. An inspection of the data revealed that the higher time average is associated with a sudden and short-lived high amplitude spike in the rms analysis error. A further analysis of the problem revealed that spikes occur very rarely and they usually have small amplitude (smaller than 1). On rare occasions, however, the spikes can have large amplitude (sometimes larger than 5), and they can last a few thousand time steps. This phenomenon is illustrated by Figure 4.4, in which the large spike occurs after more than 194,000 time steps (equivalent to about 133 years, assuming that one time step is equivalent to 6 hours) and lasts about 3,000 time steps (2 years in real time). The severity of this problem was studied by carrying out several long term integrations with different combinations of the tunable parameters. An interesting feature is that the spikes do not destroy the overall stability of the cycle; the large errors always disappear after a finite time and the mean error is smaller than 0.3. (For the case shown in Figure 4.4 the time mean error is 0.23). Spikes occur regardless of the size of the local region, and the type of the variance inflation scheme. They become less frequent, however, as the rank and the variance inflation are increased. In particular, no spikes were observed for $\varepsilon \geq 0.022$. This suggests that the easiest way to prevent the occurrence of spikes is to choose a large enough variance inflation coefficient.

k	4	5	6	7	8	9
Δ						
0.020	D	D	0.50	0.30	D	D
0.024	D	0.87	0.42	0.21	0.21	0.21
0.028	0.21	0.36	0.20	0.20	0.20	0.20
0.032	0.20	0.20	0.20	0.20	0.20	0.20
0.036	0.20	0.20	0.20	0.20	0.29	0.20
0.040	0.20	0.20	0.20	0.20	0.20	0.20
0.044	0.20	0.20	0.20	0.20	0.20	0.20
0.048	0.20	0.20	0.20	0.20	0.20	0.20

Table 4.3: Dependence of the rms analysis error on Δ in the regular variance inflation scheme and the rank (k) of the background error covariance matrix. The meaning of D is the same as in Table 4.1. The window size is 13.

All results shown so far were obtained using (4.90) to generate the analysis ensemble, $\mathbf{X}_m^a$. This scheme results in analyis pertrubations of the form $\delta\mathbf{x}_m^{a(i)} = \delta\mathbf{x}_m^{a(i)(\|)} + \delta\mathbf{x}_m^{a(i)(\perp)}$ as required by (4.32)-(4.35). In order to test the importance of including the small $\delta\mathbf{x}_m^{a(i)(\perp)} = \mathbf{x}_m^{b(i)(\perp)}$ component, the first experiment was repeated by using solution 1 [(4.52)] for $\mathbf{Z}_m$ and $\delta\mathbf{x}_m^{a(i)(\perp)} = \mathbf{0}$ instead of (4.35). (Using solution 1 and (4.35) would give the same result as (4.90) for our choice of $\mathbf{R}_{mn} = \mathbf{I}$.) This modified scheme, restricting the analysis perturbations to the k dimensional space $\mathbb{S}_m$, is clearly inferior (compare Tables 4.2 and 4.4 and Tables 4.5.4 and 4.5). More precisely, the constrained scheme provides stable analysis cycles only if both k and ε are relatively large. This is not unexpected, since setting the component $\delta\mathbf{x}_m^{a(i)(\perp)}$ to zero artificially reduces the total variance, $\|\delta\mathbf{x}_m^{a(i)}\|^2$. Increasing k decreases the reduction in the total variance, while increasing ε compensates for an increasing part of the lost variance. Also, the constrained scheme is more stable when the enhanced variance inflation is used, indicating that correcting the distribution of the variance is not less important than increasing the total variance.

Finally, we note that we have also performed tests of our data assimilation method using the Lorenz model (4.86), again with $F = 8$, but with periodicity length L greater than 40. Choosing l, k, and ε values that, for $L = 40$, gave the lowest mean error (0.2), we found that the mean error does not change with increasing L up to $L = 400$. This is roughly consistent with the supposition of an effective correlation length for the dynamics that is less than L. Thus our method appears to be effective on large systems of this type. Moreover, the (non-parallelized) analysis computational time scales linearly with the number of local regions (i.e., with L). This favorable scaling is to be expected, since the analysis

	k	4	5	6	7	8	9
inflation coefficient ϵ							
0.010		D	D	D	0.41	0.20	0.20
0.012		D	D	D	0.27	0.20	0.20
0.014		D	D	D	0.21	0.20	0.20
0.016		D	D	D	0.21	0.20	0.20
0.018		D	D	0.46	0.21	0.20	0.20
0.020		D	D	0.28	0.21	0.20	0.20
0.022		D	D	0.23	0.21	0.20	0.21
0.024		D	D	0.22	0.21	0.21	0.21

Table 4.4: Same as Table 4.2 except that Solution 1 and $\delta\mathbf{x}_m^{a(i)(\perp)} = \mathbf{0}$ is used (instead of 4.90) to obtain the analysis ensemble.

k	4	5	6	7	8	9
Δ						
0.020	D	D	D	D	D	D
0.024	D	D	D	D	0.75	0.21
0.028	D	D	D	D	0.22	0.21
0.032	D	D	D	D	D	0.20
0.036	D	D	D	D	0.21	0.25
0.040	D	D	D	0.22	0.21	0.20
0.044	D	D	D	D	0.20	0.20
0.048	D	D	D	0.25	0.20	0.20

Table 4.5: Same as Table 4.3 except that Solution 1 and $\delta\mathbf{x}_m^{a(i)(\perp)} = \mathbf{0}$ is used (instead of 4.90) to obtain the analysis ensemble.

computation size in each local region is independent of L.

4.6 Summary and conclusions

In this work, we have introduced a local method for assimilating atmospheric data to determine best-guess current atmospheric states. Our method, which we call a Local Ensemble Kalman Filter is motivated by recent studies (Patil et al. 2001, 2002) that have shown that vectors of the forecast uncertainties in local regions of moderate size tend to lie in subspaces whose dimension is relatively low. The main steps in our method are the following.

- The global analysis perturbations are advanced by the atmospheric model to obtain the global background perturbations at the next analysis time.

- In each local region, each background perturbation from the ensemble mean is used to construct a 'local vector'.

- The observations are assimilated in each local region.

- The local analyses are used to determine the global analysis and the global analysis perturbations. The cycle is then repeated.

Numerical tests of the our method using the Lorenz model, (4.86), have been performed. These tests indicate that the method is potentially very effective in assimilating data. Other potential favorable features of our method are that only low dimensional matrix operations are required, and that the analyses in each of the local regions are independent, suggesting the use of efficient parallel computation. These features should make possible fast data assimilation in operational settings. This is supported by preliminary work (not reported here) in which we

have implemented the our method on the T62, 28-level version of the National Centers for Environmental Prediction Medium Range Forecasting Model (NCEP MRF). The assimilation of a total number of 1.5×10^6 observations (including wind, temperature, and surface pressure observations) at $k' = k = 9$ takes about 20 minutes CPU time on a single 1 GHz Intel processor and about 12 minutes CPU time on a single SP2 Winterhawk node.

Acknowledgments

Appendix A: Global continuity of matrix square roots

Not all matrix square root definitions yield global continuity. One particular important mechanism for non-global-continuity of matrix square roots is that the eigenvectors of a globally continuous, symmetric, non-negative matrix, $\mathbf{M}(\mathbf{r})$, may not be definable in a globally continuous manner. In particular, for smooth variation of $\mathbf{M}(\mathbf{r})$ in two dimensions, it can be shown that there will generically be isolated points in space where two of the eigenvalues of $\mathbf{M}(\mathbf{r})$ are equal. Following previous terminology in the field of quantum chaos (e.g., Ott 2002), we call such points "diabolical points" (e.g., Berry 1983). Assume that two eigenvalues of $\mathbf{M}(\mathbf{r})$ denoted $\xi_1(\mathbf{r})$ and $\xi_2(\mathbf{r})$, are equal at the diabolical point $\mathbf{r} = \mathbf{r}_d$, and de-

note their associated orthonormal eigenvectors by $\mathbf{v}_1(\mathbf{r})$ and $\mathbf{v}_2(\mathbf{r})$. Now consider starting at a point $\mathbf{r}_o \neq \mathbf{r}_d$ and following a continuous path C that encircles $\mathbf{r}_d$ and return to $\mathbf{r}_o$. Then it can be shown that, with continuous variation of $\mathbf{v}_1(\mathbf{r})$ and $\mathbf{v}_2(\mathbf{r})$ along the path, their directions are flipped by 180° upon returns to $\mathbf{r}_o$. This presents no contradiction, since orthonormal eigenvectors are arbitrary up to within a change of sign, but it shows that $\mathbf{v}_1(\mathbf{r})$ and $\mathbf{v}_2(\mathbf{r})$ cannot be defined in a globally continuous manner. The positive symmetric square root $\big(\mathbf{M}(\mathbf{r})\big)^{1/2}$,

$$\big(\mathbf{M}(\mathbf{r})\big)^{1/2} = \sum_j \xi_j^{1/2}(\mathbf{r})\mathbf{v}_j(\mathbf{r})\mathbf{v}_j^T(\mathbf{r}),$$

is globally continuous because $\mathbf{v}_j(\mathbf{r})\mathbf{v}_j^T(\mathbf{r})$ returns to itself upon circuit around a diabolical point, even though $\mathbf{v}_j(\mathbf{r})$ may flip by 180°. Thus the solutions for $\mathbf{Z}_{mn}$ given in (4.52), (4.56), and (4.65) will be globally continuous, since positive symmetric square roots are used. The Cholesky square root will also yield global continuity. On the other hand, as an example of one of the choices that is unsatisfactory, the matrix square root choice,

$$\sqrt{\mathbf{M}(\mathbf{r})} = \big(\mathbf{M}(\mathbf{r})\big)^{1/2}\big[\mathbf{v}_1(\mathbf{r}) \mid \mathbf{v}_2(\mathbf{r}) \mid \cdots\big]^T$$

is clearly not globally continuous if diabolical points are present.

Appendix B: Equivalence of (4.75) and (4.76) for $\mathbf{Y}_{mn}$ and Solution 3 for $\mathbf{Z}_{mn}$

In this Appendix we show that $\mathbf{Z}_{mn}]\hat{\mathbf{X}}_{mn}^b$ with $\mathbf{Z}_{mn}$ given by (4.65) and $\hat{\mathbf{X}}_{mn}^b\mathbf{Y}_{mn}$ with $\mathbf{Y}_{mn}$ given by (4.75) and (4.76) both yield the same result for $\hat{\mathbf{X}}_{mn}^a$. To see this we assume that $\mathbf{Z}_{mn}$ and $\mathbf{Y}_{mn}$ yield the same $\hat{\mathbf{X}}_{mn}^a$ as would be the case for

(4.84). Thus $\mathbf{Z}_{mn}\hat{\mathbf{X}}_{mn}^{b} = \hat{\mathbf{X}}_{mn}^{b}\mathbf{Y}_{mn}$. Premultiplying both sides of this equation by $\hat{\mathbf{X}}_{mn}^{bT}$ we obtain

$$\left(\mathbf{P}_{mn}^{b}\right)^{1/2}\left[\left(\mathbf{P}_{mn}^{b}\right)^{-1/2}\mathbf{Z}_{mn}\left(\mathbf{P}_{mn}^{b}\right)^{1/2}\right]\left(\mathbf{P}_{mn}^{b}\right)^{1/2} = \hat{\mathbf{X}}_{mn}^{b}\mathbf{Y}_{mn}\hat{\mathbf{X}}_{mn}^{bT}. \qquad (4.89)$$

Thus by (4.89) our choice of the symmetric square root, (4.76), implies that $\left(\mathbf{P}_{mn}^{b}\right)^{-1/2}\mathbf{Z}_{mn}\left(\mathbf{P}_{mn}^{b}\right)^{1/2}$ is also symmetric, which by the discussion in section 4.4.2 implies solution 3.

Appendix C: $\mathbf{X}_{mn}^{a}$ obtained directly from $\mathbf{Y}_{mn}$

In this Appendix we show that $\mathbf{Y}_{mn}$ as given by (4.84) can be used to obtain the analysis $\mathbf{X}_{mn}^{a}$ directly. That is, our specification (4.84) has so far been shown to yield the analysis component, $\delta\mathbf{x}_{mn}^{a(i)(\|)} = \mathbf{Q}_{mn}\delta\hat{\mathbf{x}}_{mn}^{a(i)}$, in the low dimensional subspace $\mathbb{S}_{mn}$, but we now claim that the same $\mathbf{Y}_{mn}$ can be applied to yield the full analysis,

$$\mathbf{X}_{mn}^{a} = \mathbf{X}_{mn}^{b}\mathbf{Y}_{mn}. \qquad (4.90)$$

(The crucial difference between (4.90) and (4.45) is the absence of the super-scribed circumflexes in (4.90)). First we note that premultiplication of (4.90) by $\mathbf{Q}_{mn}$ returns (4.45), so that $\delta\mathbf{x}_{mn}^{a(i)(\|)} = \mathbf{Q}_{mn}^{T}\mathbf{Q}_{mn}\mathbf{X}_{mn}^{b}\mathbf{Y}_{mn}$ as required for the analysis component in $\mathbb{S}_{mn}$. Operating on both sides of (4.90) with $\mathbf{\Lambda}_{mn}^{(\perp)}$ and using $\hat{\mathbf{X}}_{mn}^{bT} = \mathbf{X}_{mn}^{bT}\mathbf{Q}_{mn}$ in (4.84), we have

$$\mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^{a} = \mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^{b}\mathbf{X}_{mn}^{bT}\mathbf{Q}_{mn}\hat{\mathbf{P}}^{b-1}(\mathbf{Z}_{mn} - \mathbf{I})\hat{\mathbf{X}}_{mn}^{b} + \mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^{b}. \qquad (4.91)$$

Now we recall from section 4.2 that $\mathbb{S}_{mn}$ and $\bar{\mathbb{S}}_{mn}$ are consructed from spanning vectors that are eigenvectors of $\mathbf{P}_{mn}^{b'}$. Thus $\mathbb{S}_{mn}$ and $\bar{\mathbb{S}}_{mn}$ are invariant under $\mathbf{P}_{mn}^{b'}$.

Since $\mathbf{P}_{mn}^{b'} = \mathbf{X}_{mn}^b\mathbf{X}_{mn}^{bT}$ (see equation 4.10), we have that $\mathbf{X}_{mn}^b\mathbf{X}_{mn}^{bT}$ commute with the projection operators $\mathbf{\Lambda}_{mn}^{(\|)}$ and $\mathbf{\Lambda}_{mn}^{(\perp)}$. Thus

$$\mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^b\mathbf{X}_{mn}^{bT}\mathbf{Q}_{mn} = \mathbf{X}_{mn}^b\mathbf{X}_{mn}^{bT}\mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{Q}_{mn} = \mathbf{0}, \qquad (4.92)$$

where the second equality follows because $\mathbf{Q}_{mn}\hat{\mathbf{w}}$ is in $\mathbb{S}_{mn}$ for any k-dimensional column vector $\mathbf{w}$, thus yielding $\mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{Q}_{mn} = \mathbf{0}$. From (4.91) and (4.92) we have $\mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^a = \mathbf{\Lambda}_{mn}^{(\perp)}\mathbf{X}_{mn}^b$ or $\delta\mathbf{x}_{mn}^{a(i)(\perp)} = \delta\mathbf{x}_{mn}^{b(i)(\perp)}$, as required by (4.35). This establishes (4.90). We find that use of (4.90) can be potentially advantegous for efficient parallel implementation of our method. We plan to further discuss this in a future publication applying our local ensemble Kalman filter to the operational global model of the National Centers for Environmental Prediction.

Appendix D: Enhanced Variance Inflation

In section 4.4.6 we mentioned the modification of $\mathbf{P}_{mn}^a$ or $\mathbf{P}_{mn}^b$ to prevent the occurrence of small eigenvalues in these matrices. Furthermore, we noted the possibility of an accompanying modification of the corresponding ensemble perturbations, so as to preserve the relation,

$$\hat{\mathbf{P}}_{mn} = \frac{1}{k'}\sum_{i=1}^{k'+1}\delta\hat{\mathbf{x}}_{mn}^{(i)}\big(\delta\hat{\mathbf{x}}_{mn}^{(i)}\big)^T. \qquad (4.93)$$

In the above equation we have suppressed the superscript a or b with the understanding that (4.93) can apply to either the analysis or background.

We consider the example where $\hat{\mathbf{P}}_{mn}$ is changed to a new covariance matrix by addition of a small perturbation in the form,

$$\hat{\mathbf{P}}_{mn}^* = \hat{\mathbf{P}}_{mn} + \frac{\varepsilon\Lambda}{k}\mathbf{I}_k, \qquad \varepsilon > 0, \qquad (4.94)$$
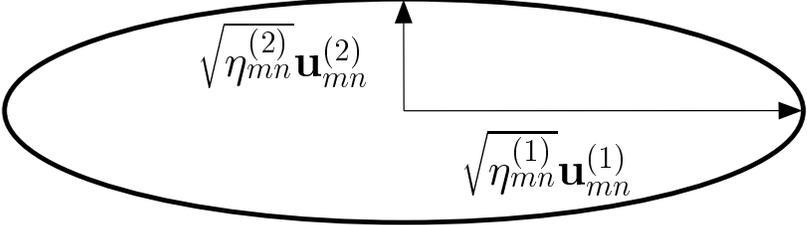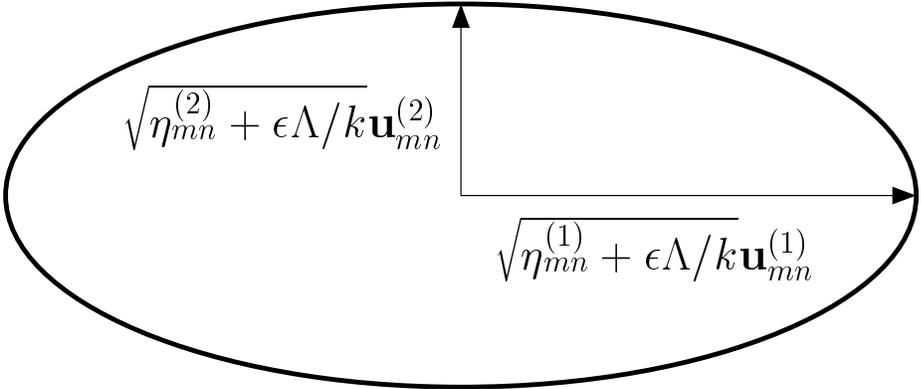
Figure 4.5: The effect of the enhanced variance inflation (equation 4.94) on the probability ellipsoid. For the special case $\hat{\mathbf{P}}_{mn} = \hat{\mathbf{P}}_{mn}^b$, $\eta_{mn}^{(1)} = \lambda_{mn}^{(1)}$ and $\eta_{mn}^{(2)} = \lambda_{mn}^{(2)}$.

where $\mathbf{I}_k$ denotes the $k \times k$ unit matrix, and Λ is the trace of $\hat{\mathbf{P}}_{mn}$; i.e., it is the sum of its eigenvalues, and thus represents the total variance of the ensemble. (The case $k = 2$ is illustrated in Figure 4.5.) Hence (4.94) increases the total variance by the factor $(1 + \varepsilon)$, where we regard ε as small, $1 \gg \varepsilon > 0$. More importantly, for small ε, the additional variance represented in (4.94) results in a relatively small change in the largest eigenvalues of $\hat{\mathbf{P}}_{mn}$, but prevents any eigenvalue from dropping below $\varepsilon\Lambda/k$, thus effectively providing a floor on the variance in any eigendirection. Having modified $\hat{\mathbf{P}}_{mn}$ to $\hat{\mathbf{P}}^*_{mn}$ via (4.94), we now consider the modification of the ensemble perturbations, $\delta\hat{\mathbf{x}}^{(i)}_{mn}$, to another set of ensemble perturbations, $\delta\hat{\mathbf{x}}^{(i)*}_{mn}$, with the perturbed covariance,

$$\hat{\mathbf{P}}^*_{mn} = \frac{1}{k'} \sum_{i=1}^{k'+1} \delta\hat{\mathbf{x}}^{(i)*}_{mn} \left(\delta\hat{\mathbf{x}}^{(i)*}_{mn}\right)^T. \tag{4.95}$$

We use the result of sections 4.4.2 and 4.4.4 to choose the $\delta\hat{\mathbf{x}}^{(i)*}_{mn}$ to minimize the difference with $\delta\hat{\mathbf{x}}^{(i)}_{mn}$. This result is the same for all metrics $\mathbf{D}_{mn}$ that commute with $\hat{\mathbf{P}}_{mn}$ (equivalently $\hat{\mathbf{P}}^*_{mn}$). (Note that the solutions in (4.60) are all the same if $\mathbf{D}_{mn}$, $\hat{\mathbf{P}}^a_{mn}$ and $\hat{\mathbf{P}}^b_{mn}$ commute.) Adopting this solution for $\delta\hat{\mathbf{x}}^{(i)*}_{mn}$, we introduce the orthogonal eigenvectors of $\hat{\mathbf{P}}_{mn}$, which we denote $\mathbf{w}^{(j)}_{mn}$. The result for $\delta\hat{\mathbf{x}}^{(i)*}_{mn}$ is then

$$\delta\hat{\mathbf{x}}^*_{mn} = \mathbf{Z}^*_{mn}\delta\hat{\mathbf{x}}_{mn}, \tag{4.96}$$

where

$$\mathbf{Z}^*_{mn} = \sum_{j=1}^{k} \xi^{(j)}_{mn}\mathbf{w}^{(j)}_{mn}\left(\mathbf{w}^{(j)}_{mn}\right)^T \tag{4.97}$$

with

$$\xi^{(j)}_{mn} = \sqrt{1 + \varepsilon\Lambda/k\eta^{(j)}_{mn}}, \tag{4.98}$$

and $\eta^{(j)}_{mn}$ is the eigenvalue of $\hat{\mathbf{P}}_{mn}$ corresponding to $\mathbf{w}^{(j)}_{mn}$; that is, $\hat{\mathbf{P}}^{(j)}_{mn}\mathbf{w}^{(j)}_{mn} = \eta^{(j)}_{mn}\mathbf{w}^{(j)}_{mn}$.

Recalling that $\hat{\mathbf{P}}^b_{mn}$ is diagonal (see (4.23)), we see that in the case $\hat{\mathbf{P}}_{mn} = \hat{\mathbf{P}}^b_{mn}$ (which is employed in section 4.5) the ith component of the vector $\mathbf{w}^j_{mn}$ is δ_{ij}. Consequently, for this case, (4.97) and (4.98) imply that $\mathbf{Z}^*_{mn}$ is diagonal,

$$\mathbf{Z}^*_{mn} = diag(\xi_1, \xi_2, \cdots, \xi_k). \tag{4.99}$$

In the case $\hat{\mathbf{P}}_{mn} = \hat{\mathbf{P}}^a_{mn}$, one could combine variance inflation and a procedure for obtaining the analysis ensemble $\{\delta\hat{\mathbf{x}}^{a(i)}_{mn}\}$ (e.g., solutions 1, 2,, or 3 of section 4.4.2): First inflate $\hat{\mathbf{P}}^a_{mn}$,

$$\hat{\mathbf{P}}^{a*}_{mn} = \hat{\mathbf{P}}^a_{mn} + \hat{\mathbf{G}}^a_{mn},$$

where $\hat{\mathbf{G}}^a_{mn}$ is any chosen inflation; and, second, replace $\hat{\mathbf{P}}^a_{mn}$ by $\hat{\mathbf{P}}^{a*}_{mn}$ in the chosen algorithm for determining the analysis ensemble (e.g., section 4.4.2).

BIBLIOGRAPHY

M. Abramowitz and I. A. Stegun *Handbook of Mathematical Functions*, (Dover Books, New York, 1965).

J. C. Alexander, I. Kan, J. Yorke and Z. You., Int. J. Bifurcation Chaos **2**, 795 (1992).

Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.,* **129,** 2884-2903.

Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.,* **127,** 2741-2758.

P. Ashwin, J. Biescu, and I. Stewart, Phys. Lett. A **193**, 126 (1994).

Bergot, T., G. Hello, and A. Joly, 1999: Adaptive observations: A feasibility study. *Mon. Wea. Rev.,* **127,** 743-765.

Berry, M. V., 1983: Semiclassical mechanics of regular and irregular motion. In *Chaotic Behavior of Deterministic Systems,* R. H. G. Helleman and G. Ioos, eds., North-Holland, Amsterdam.

Bierman, G. J., 1977: *Factorization Methods for Discrete Sequential Estimation.* Mathematics in Science and Engineering, Academic Press, New York.

Bishop, C. H., B. J. Etherton, and S. Majumdar, 2001: Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420-436.

Bloomfield, P., 2000: *Fourier Analysis of Time Series: An Introduction.* 2d ed. Wiley-Interscience, 261 pp.

Buizza, R., and P. Chessa, 2002: Prediction of the U.S. Storm of 24-26 January 2000 with the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.,* **130**, 1531-1551.

Chang, E. K. M., 2000: Wave packets and life cycles of troughs in the upper troposphere: Examples from the Southern Hemisphere Summer Season of 1984/85. *Mon. Wea. Rev.,* **126,** 25-50.

Chang, E. K. M. and D. B. Yu, 1999: Characteristics of wave packets in the upper troposphere. Part I: Northern Hemisphere Winter. *J. Atmos. Sci.,* **56,** 1708-1728.

Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press, New York.

Dee, D., S. Cohn, A. Dalcher, and M. Ghil, 1985: An efficient algorithm for estimating noise covariances in distributed systems. *IEEE Trans. Automatic Control,* **30,** 1057-1065.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99**(C5), 10 143-10 162.

Evensen, G., and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman Filter with a quasi-geostrophic model. *Mon. Wea. Rev.,* **124,** 85-96.

Fisher, M., 1998: Development of a simplified Kalman Filter. ECMWF Research Department Tech. Memo. 260, 16 pp. [Available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire, RG2 9AX, United Kingdom.]

H. Fujisaka and H. Yamada, Prog. Theor. Phys. **75**, 1088 (1986).

Gabor, D. 1946: Theory of communication. *J. IEEE*, London, **93,** pages 429–457.

D. J. Gauthier and J. C. Bienfang, Phys. Rev. Lett. **77**, 1751 (1996).

Gelaro, R., R. H. Langland, G. D. Rohaly, T. E. Rossmond, 1999: An assessment of the singular-vector approach to target observations using the FASTEX dataset. *Quart. J. Roy. Meteor. Soc.*, **125**, 3299-3328.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. In *Dynamic meteorology: data assimilation methods,* L. Bengtsson, M. Ghil, and E. Kallen, eds. Springer-Verlag, New York, 139-224.

J. Hale and H. Kocak *Dynamics and Bifurcations*, (Springer-Verlag, New York, 1991).

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman Filter-3D variational analysis scheme. *Mon. Wea. Rev.,* **128,** 2905-2919.

Hamill, T. M., J. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an Ensemble Kalman Filter. *Mon. Wea. Rev.,* **129,** 2776-2790.

J. F. Heagy, T. L. Carroll, and L. M. Pecora, Phys. Rev. E **52** R1253 (1995).

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 796-811.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman Filter technique. *Mon. Wea. Rev.,* **126,** 796-811.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123-137.

Hovmöller, E., 1949: The trough-and-ridge diagram. *Tellus,* **1,** 62-66.

B. R. Hunt and E. Ott, Phys. Rev. Lett. **76**, 2254 (1996).

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential, and variational. *J. Meteor. Soc. Japan,* **75**(1B), 181-189.

A. Jalnine and S.-Y. Kim, Phys. Rev. E **65**, 026210 (2002).

Jones, R., 1965: An experiment in nonlinear prediction. *J. Appl. Meteor.,* **4,** 701-705.

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D, J. Basic Eng.,* **82,** 35-45.

Kalman, R., R. Bucy, 1961: New results in linear filtering and prediction theory. *Trans. ASME, Ser. D, J. Basic Eng.,* **83,** 95-108.

Kalnay, E., 2002: *Atmospheric modeling, data assimilation, and predictability.* Cambridge University Press, Cambridge, 314 pp.

Kalnay, E., and Z. Toth, 1994: Removing growing errors in the analysis. Preprints, *10th AMS Conference on Numerical Weather Prediction,* Portland, OR, 212-215.

S.-Y. Kim, W. Lim, A. Jalnine and S. P. Kuznetsov, to be published.

S.-Y. Kim, W. Lim and Y. Kim, Prog. Theor. Phys. **107**, 2 (2002).

S.-Y. Kim and W. Lim, Phys. Rev. E **63**, 026217 (2001).

S.-Y. Kim, W. Lim and Y. Kim, Prog. Theor. Phys. **105**, 187 (2001).

S.-Y. Kim and W. Lim, Phys. Rev. E **64**, 016211 (2001).

Y. C. Lai *et al.*, Phys. Rev. Lett. **77**, 55 (1996).

Laine, A., and Fan, J., 1996: Frame Representations for Texture Segmentation. *IEEE Transactions on Image Processing,* **5,** 5, 771-780.

Langland, R. H., M. A. Shapiro, and R. Gelaro, 2002: Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm. *Mon. Wea. Rev.*, **130**, 957-974.

O. Lanford *Nonlinear Problems in the Physical Sciences and Biology, Lecture Notes in Mathematics*, (Springer-Verlag, New York, 1973).

Lee, S., and I. M. Held, 1993: Baroclinic wave packets in models and observations. *J. Atmos. Sci.,* **50**, 1413-1428.

Lorenc, A., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177-1194.

Lorenz, E. N., K. A Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399-414.

Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proc. Seminar on Predictability,* Vol. 1, ECMWF, Reading, Berkshire, UK, 1-18.

Morss, R. E., J. Woolen, and T. Marchok, 1999a: Ensemble-based targeting experiments during FASTEX: the effect of dropsonde data from the Lear jet. *Quart. J. Roy. Meteor. Soc.,* **125,**, 3189-3218.

Oppenheim, A., and Shafer, R., 1975: *Digital Signal Processing* Prentice Hall, 585 pp.

Orlanski, I. and J. P. Sheldon, 1993: A case of downstream baroclinic development over western North America. *Mon. Wea. Rev.,* **121,** 2929-2950.

Ott, E., 2002: *Chaos in dynamical systems,* (second edition). Cambridge University Press, Cambridge, chapter 11.

E. Ott and J. C. Sommerer, Phys. Lett. A **188**, 39 (1994).

E. Ott, J. C. Sommerer, J. C. Alexander, I. Kan, and J. A. Yorke, Phys. Rev. Lett. **71**, 4143 (1993).

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

Patil, D. J., I. Szunyogh, B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2002: Identification of local low dimensionality of atmospheric dynamics. To be submitted to Tellus.

Patil, D. J., B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. *Phys. Rev. Lett.,* **86,** 5878-5881.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* 2d ed. Springer-Verlag, 710 pp.

L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **64**, 821 (1990).

Persson, A., 2000: Synoptic-dynamic diagnosis of medium range weather forecast systems. *Proceedings of the Seminars on Diagnosis of models and data assimilation systems.* 6-10 September 1999, ECMWF, Reading, U.K., 123-137.

Petersen, D., 1973: A comparison of the performance of quasi-optimal and conventional objective analysis schemes. *J. Appl. Meteor.,* **12,** 1093-1101.

Phillips, N. A., 1990: *Dispersion processes in large-scale weather prediction.* WMO-No. 700, World Meteorological Organization, pp. 126.

A.S Pikovsky, Z. Phys. B **55**, 149 (1984).

N. Platt, E. A. Spiegel and C. Tresser, Phys. Rev. Lett. **70**, 279 (1993).

Pu, Z.-X., and E. Kalnay, 1999: Targeting observations with the quasi-inverse linear and adjoint NCEP global models: Performance during FASTEX. *Quart. J. Roy. Meteorol. Soc.,* **125,** 3329-3337.

Rossby, C.-G., 1945: On the propagation of frequencies and energy in certain types of oceanic and atmospheric waves. *J. Meteor.,* **2,** 187-204.

Rossby, C.-G., 1949: On the dispersion of planetary waves in a barotropic atmosphere. *Tellus,* **1,** 54-58.

Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Am. Meteorol. Soc.,* **77,** 953-961.

J. C. Sommerer and E. Ott, Nature (London) **365**, 136 (1993).

Swanson, K., and R. T. Pierrehumbert, 1994: Nonlinear wave packet evolution on a baroclinically unstable jet. *J. Atmos. Sci.,* **51,** 384-396.

D. Sweet, Physics of Plasmas, to appear (2002).

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.,* **130,** 1144-1165.

Szunyogh, I., Z. Toth, R. E. Morss, S. J. Majumdar, B. J. Etherton, and C. H. Bishop, 2000: The effect of targeted dropsonde observations during the 1999 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.,* **128,** 3520-3537.

Talagrand, O., 1981: A study of the dynamics of four-dimensional data assimilation. *Tellus*, **33,** 43-60.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2002: Ensemble square-root filters. (under review).

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **127,** 1374-1377.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteorol. Soc.,* **74,** 2317-2330.

S. C. Venkataramani, B. R. Hunt, E. Ott, D. J. Gauthier and J. C. Bienfang, Phys. Rev. Lett. **77**, 27 (1996).

S. C. Venkataramani, B. R. Hunt and E. Ott, Phys. Rev. E **54**, 2 (1996).

S.C. Venkataramani, T. M. Antonsen, E. Ott and J. C. Sommerer, Phys. Lett. A **207**, 173 (1995).

Y.-H. Wan SIAM J. Appl. Math. **34**, 1 (1978).

Wang, X., and C. H. Bishop, 2002: A comparison of breeding and Ensemble Transform Kalman Filter ensemble forecast schemes. Preprints, *Symposium on Observations, Data Assimilation, and Probabilistic Prediction*, Orlando, FL., Amer. Meteor. Soc., J28-J31.

Whitaker, J. S., and T. H. Hamill, 2002: Ensemble Data Assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913-1924.

S. Yanchuk, Y. Maistrenko and E. Mosekilde, Physica D: Nonlinear Phenomena **154**, 26 (2001).

Yeh, T., 1949: On energy dispersion in the atmosphere. *J. Meteor.,* **6,** 1-16.

Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the "Surprise" Snowstorm of the 24-25 January 2000. *Mon. Wea. Rev.,***130**, 1617-1632.

Zimin A. V., I. Szunyogh, D. J. Patil, B. R. Hunt, and E. Ott, 2003: Extracting envelopes of Rossby wave packets. *Mon. Wea. Rev.,* **131,** (in print).