Analysis of chaotic multi-variate time-series from spatio-temporal dynamical systems

Odd-Halvdan Sakse Ørstavik

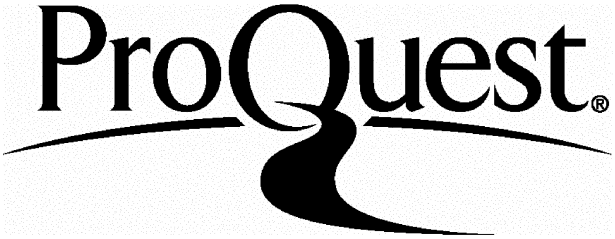
Thesis submitted for the degree of

Doctor of Philosophy

Centre for Nonlinear Dynamics and its Applications

University College London

19.08.99



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

This work would have not been possible without the guidance and encouragement given to me by my supervisor Dr. Jaroslav Stark. Thank you very much.

Thanks are also due to Dr. Ricardo Carretero-González which I worked closely with on most of my thesis. A lot of collaboration was done with Professor David Broomhead and Dr. Jerry Huke at University of Manchester Institute of Science and Technology from which I benefited greatly. I also enjoyed discussions with Dr. Thomas Schreiber.

I would also like to thank the Professor Michael Thompson and Professor Steven Bishop and all my fellow students at the Centre for Nonlinear Dynamics and its Applications for creating such a pleasant environment while I undertook this work.

Last, but not least, I would like to give a huge thank you to Sarah for putting up with me throughout the last 3 years.

Abstract

This work concerns the analysis of chaotic multi-variate time-series from spatio-temporal dynamical systems (STS). Such systems can be thought of as consisting of a collection of sub-systems at different spatial locations coupled together into one large system. These arise in many applications throughout science and engineering including most types of fluid flow, pattern formation in chemical and biological systems, dynamics of ecosystems, road traffic, vibration of structures such as beams, plates and shells, and many others. In many situations there is a desire to analyse data from STS in situations where little is known about the system generating the data. In particular one may have no idea of the system's structure, or even its state space. It has until now been an open question how to characterise, control and predict the future evolution of STS in these circumstances. To answer these questions this thesis builds on the chaotic time-series analysis framework that has been successfully developed for the analysis of low-dimensional systems. Coupled map lattices (CML) are used as model systems since these feature many of the characteristics of STS. Several new results that apply to spatio-temporal systems are presented and can be summarised as follows.

By using a mix of temporal and spatial embedding techniques one is able to carry out reconstruction and cross-prediction on a time-series generated by a CML and the results show that spatio-temporal delay reconstructions give better predictability than standard methods using either time delays or spatial delays only. A framework for embedding spatio-temporal systems is proposed.

Results also show that by using spatio-temporal embedding techniques with local observations one cannot detect the presence of spatial extent in CML's thus suggesting the impossibility of reconstructing the whole system from localised information.

New methods for calculating Lyapunov spectra for STS, and for extracting related quantities such as KS entropy density and Lyapunov dimension density, have been developed both for the case where the underlying dynamics is known and directly from time-series.

Contents

List of Figures

List of Tables

Publications of results

The results presented in chapter 3 have been published in S. Ørstavik & J. Stark, "Reconstruction nad cross-prediction in coupled map lattices using spatio-temporal embedding techniques," Phys. Lett. A **247**, 145-160 (1998).

The results presented in chapter 4 have been accepted for publication in R. Carretero-González, S. Ørstavik, J. Huke, D.S. Broomhead & J. Stark, "Thermodynamic limit from small lattices of coupled maps," Phys. Rev. Lett., to appear (1999).

The results presented in chapter 5 have been published in R. Carretero-González, S. Ørstavik, J. Huke, D.S. Broomhead & J. Stark, "Scaling and interleaving of subsystem Lyapunov exponents for spatio-temporal systems," Chaos **9(2)**, 466-482 (1999).

The results presented in chapter 6 are about to be submitted to Physica D in S. Ørstavik, R. Carretero-González & J. Stark, "Estimation of intensive quantities in spatio-temporal systems from time-series".

Chapter 1

Introduction

Historically, the vast majority of the theory of nonlinear dynamical systems was developed under the assumption that one knew the state space and evolution equations of the system under consideration. In principle the application of these results to real problems thus required the simultaneous observation of all the state space variables. This is obviously not possible in many practical applications, where all that one can manage to do is to make a sequence of repeated measurements of one or more observables whose relationship to the state variables is at best uncertain. It is then of fundamental importance to understand how much information about the original dynamical system can be extracted from such a time-series of measurements.

Remarkably enough, it turns out that in the case of systems with low dimensional state space it is possible to reconstruct the whole system just from the observed time-series, using the so called method of delays. The systematic use of this technique was first suggested by Packard et.al. (Packard, Crutchfield, Farmer & Shaw 1980), who attributed the basic idea to Ruelle, though a number of other authors around that time were beginning to experiment with it for specific systems (e.g. (Ott, Sauer & Yorke 1994)). The theoretical justification for this approach is given by Takens Embedding Theorem (Takens 1981). This implies that for generic finite dimensional systems and observables, essentially no dynamical information is lost in passing from the state space to the time-series. This result has stimu-

lated a vast range of applications in fields ranging from fluid dynamics, through electrical engineering to biology, medicine and economics. It has led to both a re-examination of old data sets, and to the construction of new experiments, with the aim of detecting and perhaps even taking advantage of deterministic behaviour in time-series which were previously thought to be random. In particular, algorithms now exist to characterize and predict such time-series, to remove noise or detect weak signals and more recently to control the original dynamical system (e.g. (Eckmann & Ruelle 1985; Abarbanel, Brown, Sidorowich & Tsimring 1993; Ott, Sauer & Yorke 1994)). In appropriate circumstances, such algorithms are capable of achieving levels of performance which are far superior to those obtained using classical linear signal processing techniques.

Despite the obvious success of this approach, its relevance is currently confined to relatively simple systems whose asymptotic dynamics is low dimensional. There is therefore a need to extend such techniques to more complex systems and in particular to spatio-temporal systems. These can be thought of as consisting of a collection of sub-systems at different spatial locations coupled together into one large system. These arise in many applications throughout science and engineering including most types of fluid flow, pattern formation in chemical and biological systems, dynamics of ecosystems, road traffic, vibration of structures such as beams, plates and shells, and many others. In some cases (such as weather prediction) there is a reasonably good understanding of the underlying deterministic processes, and one is able to directly observe the necessary state variables. One then has an *a priori* model to which the observations can be fitted in order to estimate parameters and/or initial conditions. In such cases, there is a sound theoretical basis for carrying out the data analysis, and usually a great deal of experience with numerical algorithms.

Increasingly, however, there is a desire to analyse data from spatio-temporal systems in situations where little is known about the system generating the data. In particular one may have no idea of the system's structure, or even its state space. It is then natural to try approaches analogous to those for low-dimensional systems and in particular to attempt to use the method of delays to reconstruct

the unknown dynamics. In many such situations we are able to measure several observables simultaneously giving rise to a multivariate time-series. Often each component of such a time-series corresponds to the measurement of the same observable at different spatial locations. Unfortunately there is almost no theoretical framework in this case; the standard Takens Theorem rarely being applicable since measurement functions depending on only a single spatial location are not generic. Furthermore, it is far from clear that embedding the whole state space, or even the whole attractor is the correct approach, since these objects will often be very high-dimensional. There is currently therefore little understanding of what delay reconstruction methods are actually doing, what their fundamental limits are, and how much information they preserve.

Theoretical considerations apart, there are also many practical questions posed by this approach. How much data do we need to collect and how often should we sample in both time and space? How do we choose amongst the different reconstructions that are possible? Can we analyse parts of the system separately, or do we need to always consider it in its entirety? These have considerable relevance to many real problems such as the analysis of electrocardiograms (ECG), electroencephalograms (EEG), remote sensing satellite data, and turbulent fluids.

In this thesis we try to answer some of the questions posed above. In particular we will focus on how we can effectively reconstruct a spatio-temporal system to optimise predictability and to estimate invariant quantities, such as the Lyapunov spectrum.

1.1 Thesis Outline

In chapter 2 we introduce the necessary background, theory and tools we need, and which this thesis is extending; chaotic time-series analysis and embedding techniques. We explore their application to prediction and the characterisation of chaotic systems through the Lyapunov spectrum. We also introduce coupled map lattices as the model system we will use throughout this thesis.

In chapter 3 we use a mix of temporal and spatial delay embedding techniques to carry out reconstruction and cross-prediction on time-series generated by a coupled map lattice. We find that spatio-temporal delay reconstructions give better predictability than standard methods using either time delays only or spatial delays only. We also observe that in all these cases it is completely infeasible to rigorously embed the original spatio-temporal system since this would require impractically large embedding dimensions. Despite this, it proves possible to make good short term predictions in embedding dimensions as low as 4. We discuss a possible explanation of this apparent paradox and also describe a tentative theoretical framework for reconstructing high-dimensional systems that this suggests.

In chapter 4 we extend the results from the previous chapter. We compare the behaviour of a small truncated coupled map lattice with random inputs at the boundaries with that of a large deterministic lattice essentially at the thermodynamic limit. We find exponential convergence for the probability density, predictability, power spectrum, and two-point correlation with increasing truncated lattice size. This suggests that spatio-temporal embedding techniques using local observations cannot detect the presence of spatial extent in such systems and hence they may equally well be modelled by a local low dimensional stochastically driven system.

The most common and useful tool for the characterisation of chaos is given by the Lyapunov exponents. From the Lyapunov spectrum it is possible to estimate bounds for the effective number of degrees of freedom of the system (i.e. the dimension of the attractor). The computation of the Lyapunov spectrum involves matrix manipulation techniques that soon become prohibitive (in terms of both computing time and memory storage) as the original number of system variables gets large (e.g. a few hundred). In chapter 5 we study the possibility of reconstructing the Lyapunov spectrum of a spatio-temporal system by using information from a small sub-system, thereby reducing considerably the computer resources involved in the computations. We propose a new rescaling method leading to better estimates of

the Lyapunov spectrum and examine the interlacing properties of Lyapunov spectra for consecutive sub-system sizes. In the process we propose a natural method for constructing the Jacobian of systems on high-dimensional lattices.

Chapter 6 is a further investigation of the results described in the previous chapter, but here we explore whether the same characteristics can be found when the only information we have is the time-series themselves. We find that by using spatial delay reconstrucion we are able to extract the Lyapunov spectrum of the whole system and from this we calculate related measures such as Lyapunov dimension density and KS entropy density.

In chapter 7 we discuss the results presented in this thesis and suggest ways in which this work can be taken forward.

Chapter 2

Background and theory

2.1 Chaotic time-series analysis

Given a time-series from a dynamical system we wish to use this information to model, characterise and predict the system's behaviour. However, most real-life systems involve complex interactions of numerous sub-systems, which makes attempts at such an analysis very difficult. If, as usually is the case, the system (through the time-series) features consistent aperiodicity and its deviations from periodicity cannot be explained by conventional linear models, we usually assume either that the system is perturbed by stochastic noise or that it is deterministic and chaotic. Let us start by considering a (discrete) dynamical system given by the equation

$$z_{k+1} = f(z_k) \tag{2.1}$$

where $z_{k+1} \in \mathbb{R}^m, m > 0$ describes the state of the system after $k+1$ units of time and f is the evolution law, determining exactly the state of the system at any point in time, given its state in the previous time period. This means that given an initial condition z_0, any future state of the system can be exactly predicted and is given recursively by the values $f(z_0), f^2(z_0), \cdots, f^k(z_0)$.

However, from the seminal paper by Lorenz (Lorenz 1963) we know that simple deterministic dynamical systems such as eq. (2.1) can generate time-series with a degree of randomness equal to those generated by purely stochastic pro-

cesses. This randomness-generating feature is shared by a wide family of systems called chaotic dynamical systems which often generate orbits that in the stationary regime wander over complex geometric structures called strange attractors. Thus, an empirically collected time-series which seemingly comes from a stochastic source could be generated by a chaotic (and deterministic) dynamical system. This could give us the opportunity to utilise the methods of nonlinear dynamics theory to predict, model and characterise the system. This involves the concepts of reconstruction and embedding which we will discuss in the next section.

2.1.1 Reconstruction and embedding

Consider, as above, a deterministic finite-dimensional dynamical system which at each point in time is determined by a point $\mathbf{z}$ lying on a m-dimensional manifold $M \subset \mathbb{R}^m$. The time evolution is given by a map $f^t : M \to M$ such that if the system is in state $\mathbf{z}$ at t_0 it is in state $f^t(\mathbf{z})$ at time $t_0 + t$. We usually do not have access to the state $\mathbf{z}$ and can only measure some function $h : M \to \mathbb{R}$. The evolution of this quantity is given by $h(f^t(\mathbf{z}))$. In practice we only observe this at discrete time intervals, and when these are the same we in fact observe the sequence $x^n = h(\mathbf{z}^n)$ for $n = 1, 2, \ldots$ where $\mathbf{z}^n = f^{n\tau}(\mathbf{z})$ and $\tau > 0$ is the sampling interval. By rescaling time if necessary, we may take without loss of generality $\tau = 1$ and write $f = f^\tau$. Takens (Takens 1981) showed that generically it is possible to use the time-series $\{x^n\}$ to reconstruct f^t up to some unknown smooth coordinate change. More precisely, if we fix some d (called the embedding dimension) and consider the map $\Phi : M \to \mathbb{R}^d$ given by $\Phi(\mathbf{z}) = (h(f^{-d}(\mathbf{z})), h(f^{-(d-1)}(\mathbf{z})), \ldots, h(f^{-1}(\mathbf{z})))$ then under suitable smoothness and genericity assumptions on f and h, the map Φ is an embedding for $d \geq 2m + 1$. In particular Φ is smooth and has a smooth inverse on its image and hence we can define the map $F = \Phi \circ f \circ \Phi^{-1}$ on the image of Φ. F can be seen as the same dynamical system as f under the coordinate change given by Φ. In particular all the coordinate independent properties (invariant sets and geometric invariants such as the correlation dimension and Lyapunov spectrum (which we will introduce in the next sections) are the same.

The advantage of this is that F can be directly expressed in terms of the observed time-series $x^n = h(\mathbf{z}^n)$. Define the point $\mathbf{x}^n \in \mathbb{R}^d$ by the delay coordinates $\mathbf{x}^n = (x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$. Clearly $\mathbf{x}^n = \Phi(\mathbf{z}^n)$ and hence $\mathbf{x}^n$ is in the image of Φ so we can apply F to it. $F(\mathbf{x}^n) = \Phi \circ f \circ \Phi^{-1}(\mathbf{x}^n) = \Phi \circ f(\mathbf{z}^n) = \Phi(\mathbf{z}^{n+1}) = \mathbf{x}^{n+1}$. Thus F describes the dynamics of $\mathbf{x}^n$ in $\mathbb{R}^d$ and this dynamics is just the image under the invertible coordinate change Φ of the dynamics of $\mathbf{z}^n$ under f, i.e. the dynamics of $\mathbf{x}^n$ is equivalent to the dynamics of $\mathbf{z}^n$ except for a smooth invertible coordinate transformation. We can rewrite the map $F(\mathbf{x}^n) = \mathbf{x}^{n+1}$ in terms of x^n as $F(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) = (x^{n-d+1}, x^{n-d+2}, \ldots, x^n)$. Thus, unlike the map f we can estimate the map F from the time series $\{x^n\}$. Furthermore, if we restrict F to its last coordinate we get a map $G : \mathbb{R}^d \to \mathbb{R}$ that determines x^n in terms of the values $x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}$, i.e.

$$G(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) = x^n.$$

Thus, Takens Theorem allows us to determine x^n from some finite number d of previous observations. In practice G will not be known explicitly. However, given a sample of the time-series, a variety of methods can be used to find an approximation $\hat{G}$ to G. This can then be used for prediction and characterisation, or as the basis of more sophisticated processing of the time-series, e.g. noise reduction. In the next section we describe how to approximate G and use it in prediction.

2.1.2 Local predictors

Given $\mathbf{x}^n$, how do we approximate G to get an estimate $\hat{x}^n$ of x^n. One way is to use local predictors. The idea is based on the fact that points close together in the embedding space map to points that are still close together. Thus, we search for the k nearest neighbours of $\mathbf{x}^n$, $\mathbf{x}^n(j), j = 1, \ldots, k$. Denote the first entry in $\mathbf{x}^{n+1}(j)$ as $x^n(j)$. We expect $x^n(j)$ to be close to x^n. In fact, Lorenz (Lorenz 1969) suggested to use $x^n(1)$, the nearest neighbour, as an estimate of x^n (the method of analogoues). We could also use the average of $x^n(j)$ as our estimate. A more refined approach, which we will use in this thesis, is to suppose that the local

relationship G is linear or quadratic. In the linear case we then have

$$\hat{x}^n = \hat{G}(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) = a_0 + \mathbf{a}\mathbf{x}^n. \tag{2.2}$$

To fit the parameters we again consider the k nearest neighbours $\mathbf{x}^n(j)$ of $\mathbf{x}^n$. Define

$$\mathbf{x} = \begin{pmatrix} x^n(1) \\ \vdots \\ x^n(k) \end{pmatrix}, \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix}, \mathbf{\Psi} = \begin{pmatrix} 1 & \mathbf{x}^n(1) \\ \vdots & \\ 1 & \mathbf{x}^n(k) \end{pmatrix}.$$

Then we can rewrite eq. 2.2 in matrixform

$$\mathbf{x} = \mathbf{\Psi}\mathbf{a}.$$

To determine $\hat{G}$ we solve for $\mathbf{a}$, $\mathbf{a} = \mathbf{\Psi}^{-1}\mathbf{x}$. $\mathbf{\Psi}$ might be ill-conditional so we use singular value decomposition to find $\mathbf{\Psi}^{-1}$ (the pseudo-inverse). The fitted $\mathbf{a}$ then gives us an estimate of x^n by eq. 2.2. One question that has to be addressed in the practical implementation of the method is the size of the neighbourhood that should be used and this will be discussed later.

The accuracy of the predictions is measured by the root-mean-square error $\sigma_{\text{rms}} =< (\hat{x^n}(n) - x^n)^2 >^{1/2}$ and normalised by dividing by the standard deviation σ_x of the time-series, that is

$$E = \frac{\sigma_{\text{rms}}}{\sigma_x} \tag{2.3}$$

When $E = 1$ the prediction error is no different from predicting the mean, while $E = 0$ is perfect prediction.

2.1.3 Lyapunov spectrum and related quantities

The Lyapunov exponents are an important invariant of nonlinear dynamical systems and are closely related to other quantities of interest. Consider a discrete dynamical system with m state variables. For such an m-dimensional system there exist m Lyapunov exponents corresponding to the rates of expansion and/or contraction of nearby orbits in the tangent space in each dimension. Let us look at this in more detail.

Assume a dynamical system as in eq. (2.1). Consider an orbit displaced from the original orbit by an infitesimal vector $\delta z_k, z_k \to z_k + \delta z_k$. The evolution of δz_k is described by differentiating eq. (2.1), i.e.

$$\delta z_{k+1} = Df(z_k)\delta z_k \tag{2.4}$$

where $Df(x)$ denotes the $m \times m$ Jacobian matrix (of partial derivatives of $f(z)$ with respect to the m components of z).

Let $y_k = \delta z_k/|\delta z_0|$, (i.e. we normalise for convenience). The vector y_k is called a tangent vector and the space in which y_k lies is called the tangent space. The evolution equation for y_k is

$$y_{k+1} = Df(z_k)y_k. \tag{2.5}$$

Clearly the evolution of y_k depends on both the orbit $\{z_k\}$ determined by the initial condition z_0, and on the initial orientation of the unit tangent vector y_0. We are interested in the exponential rate at which the magnitude of y grows or shrinks per iterate of the map. Let

$$\lambda(z_0, y_0) = \lim_{k\to\infty} \lambda(z_0, y_0, k) \tag{2.6}$$

where

$$\lambda(z_0, y_0, k) = \frac{1}{k} \ln |y_k|. \tag{2.7}$$

We call $\lambda(z_0, y_0)$ the *Lyapunov exponents*. Let $L = e^\lambda$. L represents an average factor by which the magnitude of the infitesimal vector displacement is multiplied on each iterate. Combining eqs. (2.5) and (2.7) we get

$$\lambda(z_0, y_0, k) = \frac{1}{k} \ln |Df^k(z_0)y_0|$$

where $Df^k(z_0)$ is the Jacobian matrix of the map f^k. Applying $Df^k(z_0)$ to all unit tangent vectors y_0 results in an ellipsoid, whose m principal radii are given by $L_i(z_0, k)$ for $i = 1, 2, \ldots, m$. The principal directions of the ellipsoid are the m perpendicular eigenvectors of the real symmetric matrix $[Df^k(z_0)] \cdot [Df^k(z_0)]^T$ where $[\cdot]^T$ denotes the transpose. The principal radii of the ellipsoid are the square

roots of the m eigenvalues (the singular values) of $Df^k(z_0)$. Taking the $k \to \infty$ limit in eq. (2.6), we get m possible values of the Lyapunov exponents depending on the initial orientation of the vector y_0. Denote these values $\lambda_i(z_0)$ and assign the index i such that

$$\lambda_1(z_0) \geq \lambda_2(z_0) \geq \cdots \geq \lambda_m(z_0). \tag{2.8}$$

Then Oseledec's Multiplicative Ergodic Theorem (Oseledec 1968) guarantees that if the orbit from z_0 generates an ergodic probability measure then the limit in eq. (2.6) exists and the values of $\lambda_i(z_0)$ are the same for almost every z_0 with respect to the ergodic measure. Since the m numbers $\lambda_i(z_0)$ will be the same for all z_0 on the attractor (except for a possible set of natural measure zero) we can drop the z_0 dependence of the Lyapunov exponent and $\lambda_i(z_0) \to \lambda_i$ (understanding that λ_i is that value of $\lambda_i(z_0)$ obtained for almost every z_0). The set of all the Lyapunov exponents arranged in decreasing order is called the Lyapunov spectrum (LS).

Thus, a chaotic attractor can now be defined by the condition that there is a net average stretching for at least one orientation of y_0, and almost every z_0: $\lambda_1 > 0$. In general the $\lambda_i, i \geq 2$ can be positive or negative. if $\lambda_1 \geq \cdots \geq \lambda_j > 0 > \lambda_{j+1} \geq \ldots \geq \lambda_m$ we have j stretching directions and $(m-j)$ contracting directions. Think of an infitesimal m-dimensional ball in phase space mapped forward in time by k iterates resulting in an infitesimal ellipsoid with m principal radii. The ratio of these to the initial radius will be of the order $e^{k\lambda_i}$.

The Lyapunov exponents characterise the dynamics on the attractor, in that they say something about stability. An alternative characterisation which by contrast is purely a geometric measure are the attractor dimension. There is a range of different definitions of dimensions in dynamical systems, such as Hausdorff dimension, box-counting dimension, information dimension and correlation dimension (see (Ott, Sauer & Yorke 1994)). The correlation dimension, introduced by Grassberger and Procaccia (1983), is the one most used when calculating the dimension of an underlying system directly from time-series.

There is though a close relationship between these geometric measures and the

Lyapunov spectrum as conjectured by Kaplan and Yorke (Kaplan & Yorke 1979; Schuster 1988)).

Kaplan and Yorke define the Lyapunov dimension, D_L, as

$$D_L = j + \frac{1}{|\lambda_{j+1}|} \sum_{i=1}^{j} \lambda_i, \tag{2.9}$$

where j is the largest integer for which $\sum_{i=1}^{j} \lambda_i > 0$.

Another useful invariant that can be derived from the LS is the so called *Kolmogorov-Sinai* (KS) *entropy*, h, that can be bounded from above by the sum of the positive Lyapunov exponents λ_i^+ and that in many cases can be well approximated by (Eckmann & Ruelle 1985)

$$h = \sum \lambda_i^+. \tag{2.10}$$

The KS entropy quantifies the mean rate of information production in a system, or alternatively the mean rate of growth of uncertainty in a system subjected to small perturbations.

If we know the map f then the calculation of the LS is a straightforward task. But it has been shown that it is also possible to calculate the LS of a system when all the information available is a time-series sampled from it. To do this we need to reconstruct the dynamical system underlying the time-series.

2.1.4 Lyapunov spectrum from time-series

The Lyapunov exponents of a dynamical system can be found by applying the algorithm introduced by Eckmann and Ruelle [1985]. The method consists of finding the Jacobians along a trajectory and reorthogonalising using QR decomposition as described below. In the case where only a time-series of the system is known to us, then this require us to estimate the Jacobian from the time-series at each time step.

Assume that we have given a time-series $\{x^n\}, n = 1, 2, \ldots, N$, sampled from a dynamical system as described above. From the data we construct d-dimensional time delay vectors $v^n = (x^{n-1}, x^{n-2}, \ldots, x^{n-d})$. Let us denote by $v_j^n, j = 1, \ldots, k$

the k nearest neighbours of v^n, using the l_2-norm as distance metric. We can then define k distance vectors $z_j^n = v^n - v_j^n$. The neighbourhood matrix B_n at time n is then defined by

$$B_m = (z_1^n, z_2^n, \ldots, z_k^n)^T.$$

The image of this matrix is simply given by

$$B_{n+1} = (z_1^{n+1}, z_2^{n+1}, \ldots, z_k^{n+1})^T,$$

and the Jacobian J_n at time n is then estimated by

$$J_n = B_{n+1}^+ B_m$$

where $(\cdot)^+$ denotes the pseudo-inverse.

Now, in order to find the Lyapunov exponents recall that we look at the growth of d-volumes under the action of $J_n(v^0)$. Let Q_0 be a unit d-cube (matrix) defined by orthogonal vectors $\{q_1, \ldots, q_d\}$. The Lyapunov exponents are then given by the logarithms of the eigenvalues of $J_n(v^0)Q_0$ as $n \to \infty$. A problem is that this matrix is ill-conditioned. To avoid this we use the trick suggested in [Eckmann and Ruelle, 1985] and reorthogonalize at each step as described below.

A matrix A can be uniquely factorised into $A = QR$ where Q is an orthonormal matrix and R is upper triangular with positive values on the diagonal. The diagonal elements of R, R_{ii}, are the eigenvalues of A. Let $T_i = J^i(v^0), T_0Q_0 = Q_1R_1, T_iQ_i = Q_{i+1}R_{i+1}$ which by the chain rule gives

$$\begin{aligned} T_nQ_0 &= T_{n-1}T_{n-2}\cdots T_1T_0Q_0 \\ &= T_{n-1}T_{n-2}\cdots T_1Q_1R_1 \\ &= T_{n-1}T_{n-2}\cdots Q_2R_2R_1 \\ &= T_nR_nR_{n-1}\cdots R_2R_1 \\ &= Q_n\prod_{i=1}^{n} R_i. \end{aligned}$$

Without loss of generality, the matrix Q_0 can be chosen to be the identity matrix. We then have $r_{ii} = \prod_{j=1}^{n} r_{ii}^{(j)}$, where $R_{ii}^{(j)}$ is the ith diagonal element of R_j

and the Lyapunov exponents are then given by

$$\lambda_i = \lim_{n\to\infty} \log|r_{ii}| = \lim_{n\to\infty} \sum_{j=1}^{n} \log|r_{ii}^{(j)}|$$

Since the amount of data is limited, we do not take the limit to infinity but hope that by including enough iterations the convergence to the Lyapunov exponents is achieved. Moreover, the method require us to find nearby neighbours in d dimension thus the data requirements soon become prohibite for large d. Therefore the method has usually been applied to time-series from low-dimensional systems where $d < 5 \sim 6$.

2.2 Spatio-temporal dynamical systems

Spatio-temporal systems give rise to a wide range of interesting phenomena that cannot occur in dynamical systems with only a few degrees of freedom.

An interesting feature of spatio-temporal systems is that it can exhibit spatio-temporal chaos. This is characterised by exponentially decaying correlations both in time and space directions.

The most common approach to modelling complex spatio-temporal behaviour is through the use of partial differential equations (PDE's). The analysis and even the numerical integration of PDE's is usually quite intricate. Thus, if one desires to study the full range of complex spatio-temporal behaviour whilst conserving a relatively simple dynamical framework, a better approach is to consider discrete spatio-temporal systems. By this we mean a collection of coupled simple low-dimensional dynamical units arranged on a spatial lattice. The coupling is usually (but not always) restricted to a finite neighbourhood. An immediate advantage of such systems is their straightforward computational implementation. Another possible advantage is that the local dynamics at each lattice site in the uncoupled limit can be thoroughly analysed. The knowledge of such local dynamics in the uncoupled limit can help to provide some insight of the complexity of the coupled system.

2.2.1 Coupled map lattices

To formalise this, we will consider coupled map lattices (CML's) (Kaneko 1983; Kaneko 1984). A CML is a discrete space-time dynamical system with a continuous state space, in contrast to cellular automata where the state space is discrete. Let us denote by x_i^n the state of the ith site at time n, where the integer index i runs from 1 to N. The CML dynamics is defined by

$$x_i^{n+1} = (1-\varepsilon)f(x_i^n) + \sum_{k=-l}^{r} \varepsilon_k f(x_{i+k}^n), \tag{2.11}$$

where typically we use periodic boundary conditions, f is a real function and we ask $\sum \varepsilon_k = \varepsilon$ as a conservation law. The general CML (2.11) couples $l \geq 0$ left neighbours and $r \geq 0$ right neighbours with coefficients ε_k.

The appeal of CML's is due on one hand to their computational simplicity and on the other to the fact that they display a wide variety of spatio-temporal phenomena ranging from spatio-temporal periodic states (Gade & Amritkar 1993; Zhilin, Gang, Benkun & Gang 1994) and travelling interfaces (Carretero-González, Arrowsmith & Vivaldi 1997; Kapral, Livi, Oppo & Politi 1994) to intermittency (Keeler & Farmer 1986) and turbulence (Beck 1994; Willeboordse & Kaneko 1995).

2.3 Summary

In this chapter we have desribed the background and theory for how to predict and characterise a low-dimensional dynamical system when the only information available is a time-series sampled from it.

In the remainder of the thesis we will study more complex systems, that is spatio-temporal systems as examplified by coupled map lattices. These systems will typically be high-dimensional and we will investigate to which extent the above methods still apply and how we can overcome the problems when they do not.

Chapter 3

Reconstruction and cross prediction using spatio-temporal embedding techniques

This chapter aims to initiate a systematic study of the kinds of questions asked in the introduction, and in particular investigate how different choices affect the quality of a reconstruction and of resulting predictions in spatio-temporal dynamical systems. In itself, we do not expect this to deepen our theoretical understanding of the structure of spatio-temporal chaos. Instead, we intend to develop and evaluate techniques which can be used to analyse data from real spatio-temporal systems and relate them to existing models and theories of spatio-temporal behaviour. We shall restrict ourselves to a simple model of complex spatio-temporal behaviour, namely a coupled map lattice (CML). As mentioned in chapter 2, these have been widely studied as a paradigm spatio-temporal system, and have been shown to exhibit a wide range of interesting properties (see (Kaneko 1993) for a comprehensive review). More specifically we study a 1D coupled map lattice of size L with nearest neighbour coupling

$$x_i^{n+1} = (1-\epsilon)f(x_i^n) + \frac{\epsilon}{2}(f(x_{i-1}^n) + f(x_{i+1}^n)) \tag{3.1}$$

where x_i^n denotes the state of site i at time n (with $1 \leq i \leq L$). We use periodic boundary conditions ($x_i^n = x_{i+L}^n$). The map creating the local dynamics was chosen

to be the logistic map

$$f(x) = 1 - ax^2$$

since this is the most widely studied example. The resulting CML has been shown to exhibit a very complex range of behaviour as ϵ and a are varied, ranging from periodic solutions in space and time through intermittency to fully developed spatio-temporal chaos (Zhilin & Gang 1994; Kaneko 1993). For simplicity, we shall use the state of each lattice site as an observable giving rise to L simultaneous time-series.

The main difficulty introduced by the spatial degrees of freedom (through the coupling) is the high dimension of the full system. One could in principle take only one observable from one space location and attempt to reconstruct the dynamics of the system in the usual way using delay variables. Strictly speaking the current version of Takens Theorem does not apply to this case, since measurement functions depending on only one site are far from generic. However, an appropriate theorem could almost certainly be proved using the techniques developed in (Stark 1999). This would ensure that as long as the system is totally coupled any measurement should reveal the full structure of all the degrees of freedom. However, given the high dimensionality, such an approach is not always feasible and is likely to require unrealistically large quantities of very high quality data to be available at one spatial location (e.g. see (Lorenz 1991; Kantz & Olbrich 1997)).

It seems intuitive that rather than treating the system as purely temporal it would be better to include the spatial information in the reconstruction process. Indeed several different approaches have been attempted e.g. by using time delays of spatial averages (Little, Ellner, Pascual, Neubert, Kaplan, Sauer, Caswell & Solow 1996; Rand & Wilson 1995) or reconstructing in space instead of time (Essex, Lookman & Nerenberg 1987; Pritchard & Theiler 1994; Afraimovich, Ezersky, Rabinovich, Shereshevsky & Zheleznyak 1992; Rombouts, Keunen & Stam 1995). These studies have shown that spatial embedding techniques perform equally well as time delay embeddings. In this chapter we build further on these ideas and use spatio-temporal reconstructions, i.e. we utilise measurements at different spatial

locations as well as time delays. A similar idea was used by Zheleznyak and Chua (Zheleznyak & Chua 1994) to calculate spatio-temporal correlation dimensions, though note that they assumed that one can observe the state space variables directly. In that case the issue of reconstruction does not arise, and hence a Takens type embedding theorem is not required.

The hope is that this technique will yield better reconstructions compared to purely temporal or spatial approaches. One can also ask what spatial information should be included, i.e. what is the effect of increasing the distance between locations used in a spatial or spatio-temporal delay reconstruction. This is analogous to the choice of time delay in the temporal embedding case.

As a measure of the relative quality of different reconstructions we use predictability. Other possible measures exist, but prediction error appears to give best overall performance, is less data intensive, and more robust in the presence of noise (Schreiber & Schmitz 1997). For an interesting discussion of the relation between Lyapunov exponents and predictability in spatio-temporal systems see (Boffetta, Giuliani, Paladin & Vulpiani 1998).

Another question that arises naturally in the context of spatially extended systems is whether we can use time-series from an observable in one part of the system to say something about the evolution of observables in other parts of the system. Here we will look at cross-prediction, i.e. we use data from one location to either reconstruct or predict data at other locations. As mentioned above, this depends on the coupling between different parts of the system. If the system can be split into two noninteracting (uncoupled) sub-systems and we only observe one of these sub-systems, the other sub-system is unobservable and nothing can be deduced about its behaviour. Another possibility is two sub-systems where one is forcing the other. Given the nature of coupled sub-systems there are in effect a whole architecture of different possibilities. This raises the question if the embedding theorem still applies to these situations and indeed Stark (Stark 1999; Stark, Broomhead, Davies & Huke 1997) has extended Takens Embedding Theorem to forced systems. Cross-prediction has been investigated by Muldoon et.al. (Muldoon, Broomhead & Huke 1994) who studied 32 channels of laboratory data from a

rotating annulus experiment and were able to use data from one channel to reconstruct all the other channels. Schiff et.al. (Schiff, So, Chang, Burke & Sauer 1996) used mutual prediction to characterize the dynamical interdependence between systems.

Lastly we look at how predictability scales with different prediction parameters. Earlier work has shown that a decay in predictability with prediction time suggests a low-dimensional attractor (Sugihara & May 1990; Wales 1991; Tsonis & Elsner 1992). We want to see if this still holds for our high-dimensional coupled map lattice system and if the scaling has the same form.

3.1 Spatio-temporal embeddings

We are interested in the case where we have more than one measurement at each time interval sampled from the coupled map lattice. Suppose we have l measurements. Now each of these measurements will be a component of a l-dimensional vector given by the measurement function $v : M \to \mathbb{R}^l$. Thus we have l time-series $\{x_i^n\}$ with $i = 1, \ldots, l$ where $v(f^n(\mathbf{z})) = (x_1^n, x_2^n, \ldots, x_l^n) = \mathbf{x}^n$. It has to be borne in mind, that as mentioned in the introduction, Takens Theorem will not necessarily apply to this case, since the functions v found in many applications will not be generic observations of M. We believe, however, that appropriate generalisations of Takens Theorem could be proved and hence that this approach to spatio-temporal data analysis is a reasonable one.

A more serious difficulty is raised by the high dimensionality of the systems we seek to study. Thus in our numerical work we typically use lattice sizes of 100, or even larger. Takens Theorem thus gives an embedding dimension of 201, which is clearly impractical. Of course, the condition $d \geq 2m + 1$ in the theorem is only a sufficient condition (see section 2.1.1), and may not be necessary. However, in any case we must have at least $d \geq m$ for an embedding, which is still unusably large. One might thus instead only attempt to reconstruct the attractor. Sauer et.al. (Sauer, Yorke & Casdagli 1991) give a theorem appropriate to this case, which shows that if $d > 2D_B$, where D_B is the box-counting dimension of the

attractor of f then Φ restricted to this attractor is 1-1 and an embedding on any compact manifold contained in the attractor (such as a bounded part of an unstable manifold). Unfortunately, this does not help a great deal either, since as we shall see in chapter 5 there is good numerical evidence that the Lyapunov dimension of the attractor in our case is around 70. It is thus likely that the box-counting dimension is also large and hence there is little hope of embedding even just the attractor. We thus conclude that for the systems under consideration in this chapter, the delay map Φ defined above is never an embedding in any rigorous sense. The same holds for its spatial and spatio-temporal analogues. Note that, despite this, the literature often refers to time delay or spatial embeddings, even when it is clear that the relevant Φ is not an embedding. To avoid such confusion, we shall refer to Φ as a reconstruction or delay map.

The previous paragraph seems to lead to the pessimistic conclusion that the delay reconstruction methods that have developed out of Takens Theorem are simply not relevant to spatio-temporal systems. Yet, as indicated, some of the attempts to apply such techniques to spatio-temporal data have met with reasonable success. As an example, we shall see in this chapter that time-series generated by the lattice described above exhibit a great deal of predictability in embedding dimensions as low as 4. This suggests that the extent of the attractor in many of its dimensions is small and hence that in some sense the system is reasonably well approximated by a low dimensional one. We thus believe that although the framework based on Takens Theorem described above is strictly speaking not relevant to spatio-temporal systems, it does provide the correct intuition. We thus expect that the delay reconstruction techniques that it has inspired should have a valid basis even for spatio-temporal systems, at least in some approximate sense. More precisely, it ought to be possible to generalize Takens Theorem to systems which are only "approximately low dimensional". We suggest a possible way of doing this in the discussion of this chapter.

3.2 Cross prediction

We will also consider cross-prediction between observations at different sites. This term is actually somewhat of a misnomer, since often we wish to estimate the current value of an observation at a different site, rather than forecasting its future evolution. However, the term cross-prediction has now been adopted in the literature, and hence we shall conform to this usage. The procedure may be motivated as follows. Given simultaneous measurements v_i at more than one site let us assume that delay embedding of each v_i gives rise to an embedding Φ_i. This gives the relation $x_i^n = G_i(\mathbf{x}_i^n)$ for the ith measurement site. But since each Φ_i gives an embedding of M we can think of the measurements x_j^n at site j $(j \neq i)$ as functions on the reconstructed object $\Phi_i(M)$, i.e. given measurements from one site we should be able to reconstruct and predict signals at other measurement sites:

$$x_j^{n+T} = G_{ij}(\mathbf{x}_i^n) \tag{3.2}$$

Even when embedding is rigorously not feasible, it seems reasonable to look for a relationship of this form. To approximate G_{ij} we use local linear and quadratic approximations as described in chapter 2.1.2.

3.3 Results of predictions

We first used the logistic map to create the local dynamics with $a = 2.0$ and strong coupling $\epsilon = 0.45$. The lattice size was $L = 100$. For this choice of ϵ, a we are in the chaotic region. We used random initial conditions and discarded the first 10^5 steps. We then recorded the next 20000 data at each site. This was then divided into two sets of equal length to give a sample set of length $N_s = 10000$ and a testing set of length $N_t = 10000$. The data was normalised to $[0, 1]$. We used linear fitting in all calculations unless otherwise specified.

Since this is a homogeneous system, in the sense that every site has the same dynamics, it does not matter which site we study. Here we have chosen an arbitrary

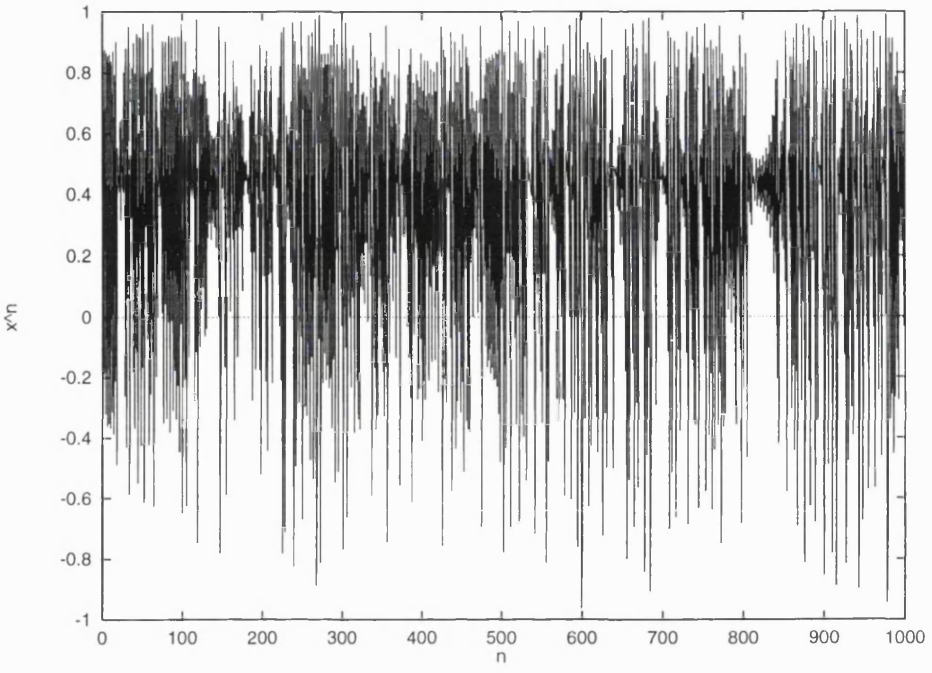


Figure 3.1: Time history of the first 1000 points of the sample set at $j = 0$.

reference site ($j = 0$) and will refer to other sites relative to this.

Figure 3.1 gives the time history of the first 1000 points of the sample set from one site. We clearly see the irregular behaviour. Note that the system often spends long periods around the point $1/2$ which is the unstable fixed point for the single map ($\epsilon = 0$) at $a = 2.0$.

In figure 3.2 we have plotted two consecutive values of x_0 against each other. We see that much of the structure given in the case of a single map (superimposed as the graph of $1 - 2x^2$) is retained.

We now turn to the problem of doing predictions on the system. Since we have past values in both space and time available we can (as described in the introduction) include both in the reconstruction. Three natural choices are either the usual time delay map

$$\mathbf{x}_0^n = (x_0^n, x_0^{n-1} \ldots, x_0^{n-(\tau_n d_n - 1)}) \tag{3.3}$$

where $d = d_n$ is the embedding dimension or, secondly, a spatial delay map given by

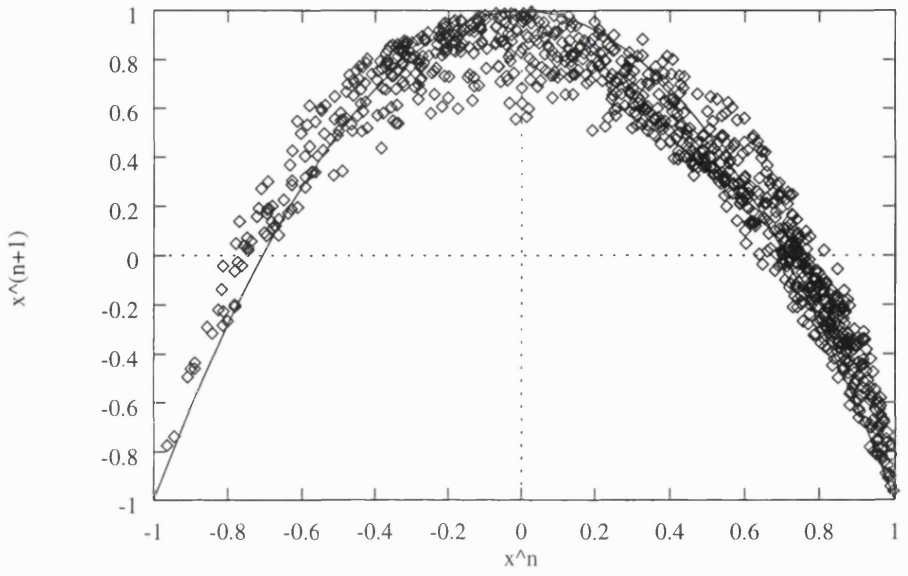


Figure 3.2: Time delay reconstruction at one node.

$$\mathbf{x}_0^n = (x_0^n, x_{-1}^n, \ldots, x_{-(\tau_s d_s - 1)}^n) \tag{3.4}$$

where now the embedding dimension is $d = d_s$. Thirdly, we can think of mixing the two types, creating a spatio-temporal delay map

$$\begin{aligned} \mathbf{x}_0^n = (x_0^n, x_0^{n-1}, &\quad \ldots \quad, x_0^{n-(\tau_n d_n - 1)}, \\ x_{-1}^n, x_{-1}^{n-1}, &\quad \ldots \quad, x_{-1}^{n-(\tau_n d_n - 1)}, \\ &\vdots \qquad\qquad (3.5) \\ x_{-(\tau_s d_s - 1)}^n, x_{-(\tau_s d_s - 1)}^{n-1}, &\quad \ldots \quad, x_{-(\tau_s d_s - 1)}^{n-(\tau_n d_n - 1)}) \qquad (3.6) \end{aligned}$$

Here the embedding dimension is $d = d_n d_s$ (we set the time delay $\tau_n = 1$ and spatial delay $\tau_s = 1$ unless otherwise specified).

Note that we are using spatial coordinates only from the left of x_i^{n+1} (i.e. x_j^n such that $j \leq i$). An obvious choice of spatio-temporal delay would be a symmetric one such as $\mathbf{x}_i^n = (x_i^{n+1}, x_i^n, x_i^{n-1})$. However, this would give artificially good results (for both the full and truncated lattices) since x_i^{n+1} depends only on these variables (c.f. eq. (3.1)). This is an artefact of the choice of coupling and observable and could not be expected to hold in general. Therefore, we use the delay map as

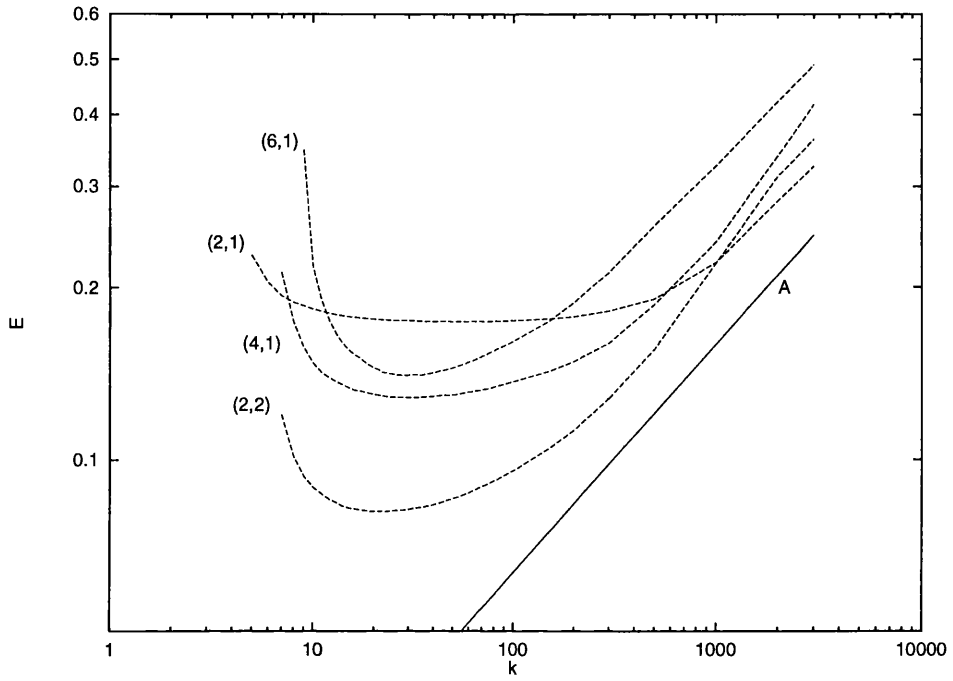


Figure 3.3: One-step prediction error (E) for increasing number of nearest neighbours ($k = 1, \ldots, 3000$) and for different choices of delay maps (d_n, d_s). The top three curves are for usual time delay map giving a minimum error for $d_n = 4$. When we instead use $d_n = d_s = 2$ i.e. the same embedding dimension size but including spatial information the predictions are substantially better. We see that the error is minimized for small values of k. The line A is a fit of the scaling law in eq.(3.8).

defined above in order to "hide" some dynamical information affecting the future state and hence make the prediction problem a non-trivial one.

In figure 3.3 we have plotted the prediction error E as we increase the number of neighbours k for some different embedding dimensions. The neighbours were found using the box-search algorithm described in (Schreiber 1995). The plot of E against k has been suggested as a way of distinguishing between low-dimensional chaos and stochastic noise in the sense that a low-dimensional deterministic system will have a minimum prediction error for a low number of nearest neighbours (Casdagli 1991). The figure shows that for time delay reconstruction the minimum prediction error is found using $d = d_n = 4$. For all reconstructions the minimum error is found using a low number of nearest neighbours, typically $k = 20 - 50$. The figure

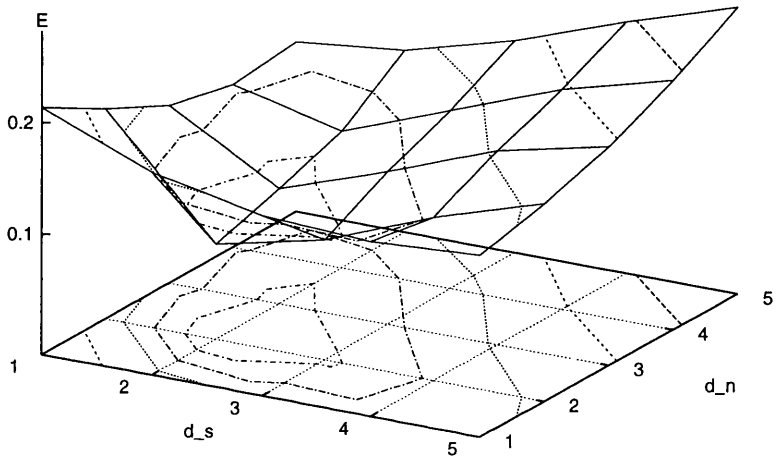


Figure 3.4: One-step prediction error for different reconstruction designs using 50 nearest neighbours.

clearly shows the positive effect of using a spatio-temporal reconstruction with $d_n = d_s = 2$ (i.e. the same embedding dimension $d = 4$ as for the best time delay reconstruction).

In figure 3.4 we have employed all three types of reconstruction for a one-step prediction while varying d_n and d_s. We chose the number of nearest neighbours as $k = 50$ which is close to the optimum for the different embedding dimensions (see figure 3.3). From figure 3.4 we see that we are able to predict even in one dimension ($d_n = d_s = 1$). This is obvious given the shape of the time-delay reconstruction in fig.3.2. This predictability is a result of having such simple local dynamics and becomes less apparent if replace the logistic map by a higher dimensional one. Note also that the error is much worse than it would be for the uncoupled case.

The prediction gets better when we employ time delay reconstruction ($d_n > 1, d_s = 1$) or spatial delay reconstruction ($d_n = 1, d_s > 1$) up until $d_n = 4$ and $d_s = 4$ respectively. However, the predictability is clearly increased if we allow both time and spatial delays together in the delay map. In this case the predictions are

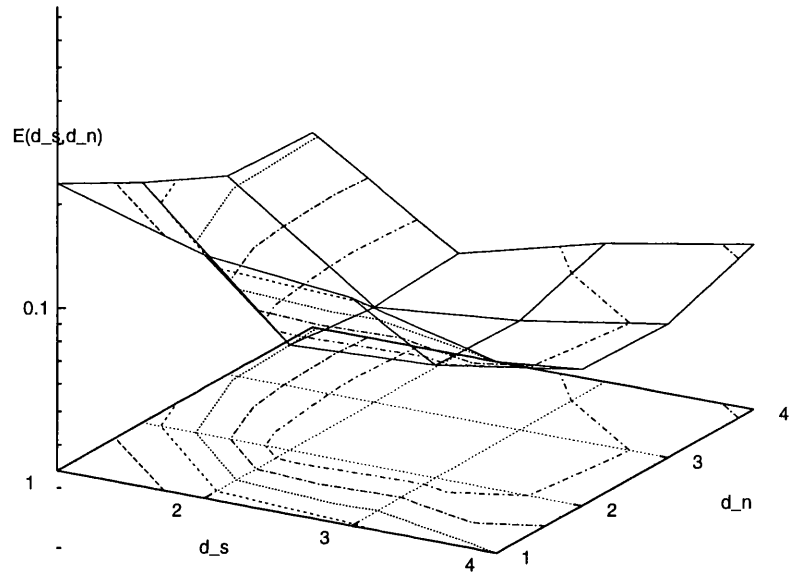


Figure 3.5: One-step prediction error for different reconstruction designs using quadratic fit and 150 nearest neighbours.

best for $d_n = d_s = 2$, i.e. an embedding dimension of $d = 4$. We see that for $d_n \geq d_s \geq 3$ or $d_s \geq d_n \geq 3$ the predictability gets worse.

In figure 3.5 we have repeated this but now using a quadratic fit. Such a fit requires more data and hence we had to increase the number of neighbours to $k = 150$. This also meant that we were unable to perform a quadratic fit at the largest embedding dimensions, i.e. $d_s = 5$ or $d_n = 5$. We see that the qualitative conclusions are the same, and we again obtain the least prediction error for $d_n = d_s = 2$. As we would expect from using a more flexible set of fitting functions, quantitatively the predictions are slightly better than for linear fits. However it is dangerous to make direct quantitative comparisons, since quadratic fits are much more data intensive, particularly at larger embedding dimensions. Given the qualitative agreement between the two types of fit we restrict ourselves to linear fits for the remainder of the chapter.

The extent of predictability is clearly shown in figure 3.6 where we have plotted the predicted values together with the actual values for the first 150 points in the

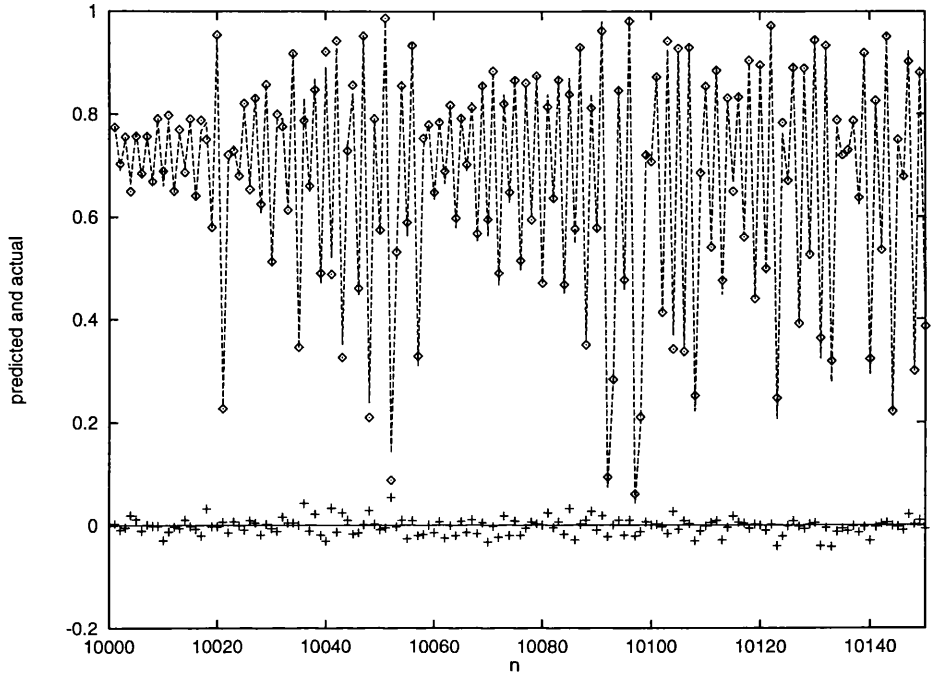


Figure 3.6: Predicted ($\diamond$) and actual (—) values for the first 150 points in the test set, $k = 50, p = 2$. (+) denotes predicted minus actual values.

fitting set ($d_n = d_s = 2, k = 50$). We see that the errors are generally less than 1 percent.

When we let $\tau_s > 1$ i.e. we reconstruct using sites whose distance apart is greater than one, the error for a one-step prediction at the reference node gets worse as shown in figure 3.7 for the cases $(d_n, d_s) = (1,2), (2,2), (3,2)$. We see that as the distance between the nodes τ_s increases the prediction error grows. For sufficiently large τ_s the error actually exceeds that for pure time delay reconstructions ($d_s = 1, d_n = 1,2,3$).

Now we turn to the question of cross-prediction, that is to what extent can we predict the evolution at other locations than the ones used in the reconstruction. In particular, given a delay vector $\mathbf{x}_0^n$ we want to find x_j^{n+T} for $T \geq 0$ and $j \geq 0$. In the case of $T = 0, j \geq 1$ what we are actually doing is reconstructing values at other locations, while for $T \geq 1, j \geq 1$ we are predicting values in the future at such locations.

The results are given in figure 3.8 where we have used in all cases $d_n = 2, d_s =$

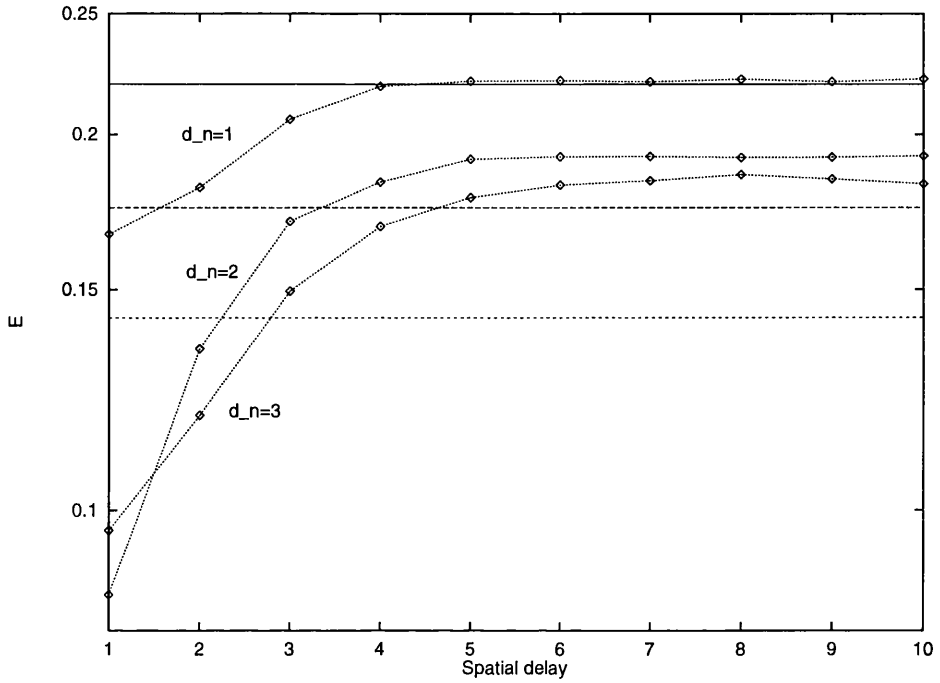


Figure 3.7: Prediction error as we increase the spatial delay. Here we have plotted the cases $(d_n, d_s) = (1,2), (2,2), (3,2)$. We used $k = 50$ neighbours. The horizontal lines depict the error when we have reconstruction in time only (the top one is for $d_n = 1$, the middle is $d_n = 2$ and the bottom is for $d_n = 3$).

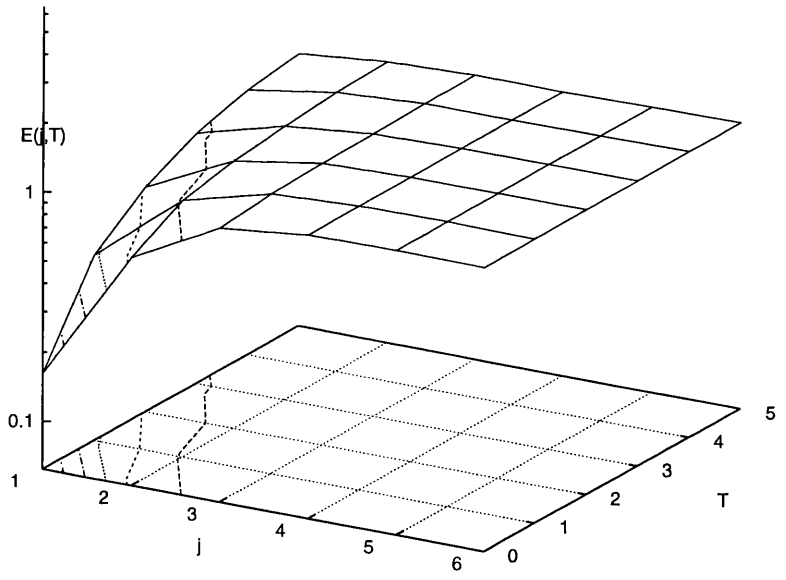


Figure 3.8: Prediction error for $j = 1 - 6, T = 0 - 5$ using $d_n = 2, d_s = 2, k = 19$.

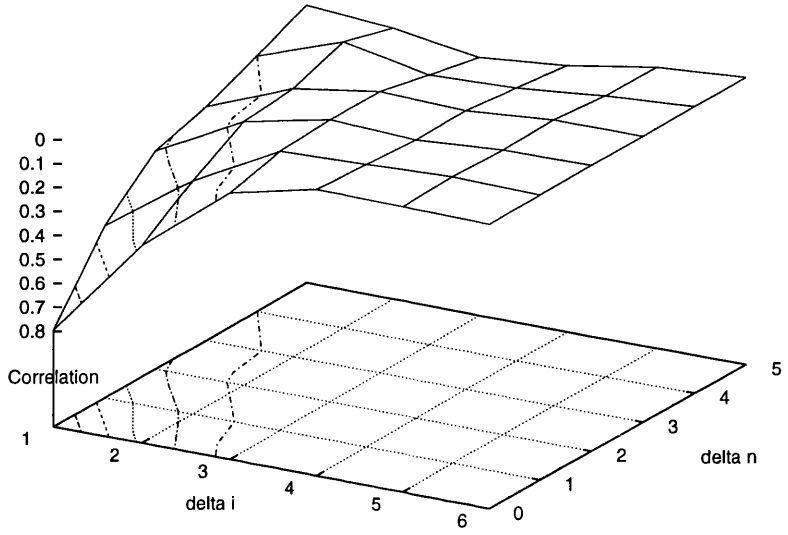


Figure 3.9: Absolute value of space-time two-point correlations. Actually we have plotted $1 - |C|$ against Δj and τ.

$2, k = 19$. We see that for cross prediction at $j = 1$ the predictability decreases as we increase T and is lost for $T \geq 5$. For $T = 0$ the ability to reconstruct is lost for $j \geq 3$. This then defines a triangle $0 \leq T \leq 4, 0 \leq j \leq 2$ where reconstruction and cross-prediction is possible.

It is interesting to compare this to the two-point space-time correlation i.e. the correlation C between two time-series separated by a time τ and space ξ defined by

$$C(\tau, \xi) = \frac{< uv > - < u >< v >}{\sigma_u \sigma_v} \tag{3.7}$$

where $u_i^n = x_i^n$ and $v_i^n = x_{i+\xi}^{n+\tau}$, σ_u and σ_v are the standard deviations of u and v. This relationship is given in figure 3.9. We see that the absolute value of C has the same form as the prediction errors in figure 3.8.

Finally we have repeated the above experiments using a variety of other coupling strengths. The results remain qualitatively the same, i.e. spatio-temporal reconstructions give the best predictions for all coupling strengths considered.

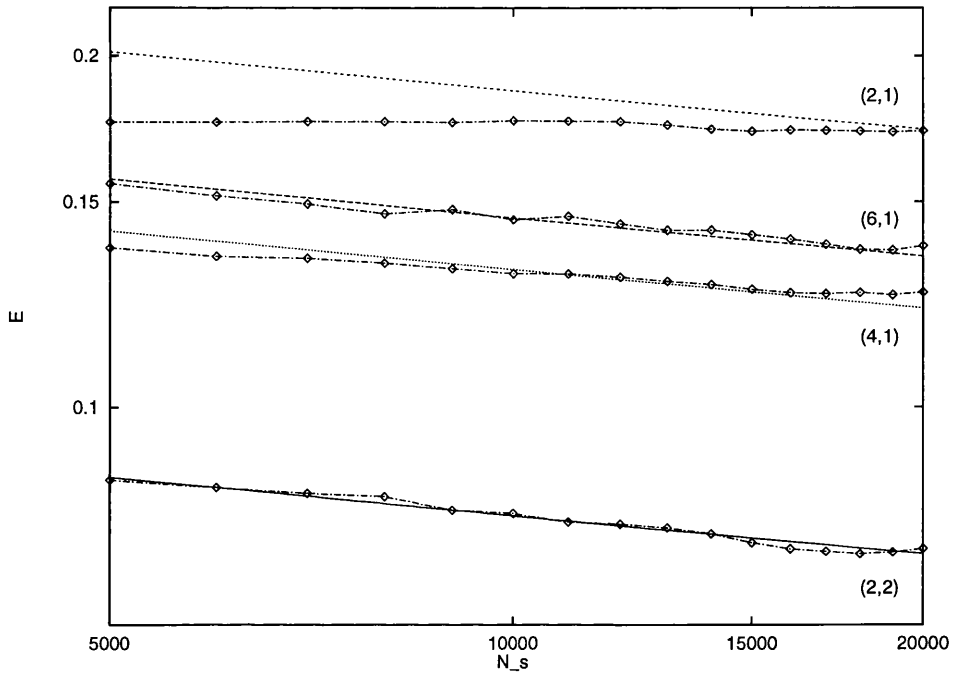


Figure 3.10: One-step prediction error for increasing N_s and for different choices of (d_n, d_s). 20 nearest neighbours were used in all cases. The imposed lines all have slope -0.11.

3.4 Scaling Laws

It has been suggested (Casdagli 1991) that the prediction error E scales with k, d and N_s as

$$E_d(k, N_s) \sim s(k/N_s)^{(p+1)/D} \tag{3.8}$$

for $k/N_s \to 0$. Here s is a 'curvature' constant, p is the order of the fitting, and D denotes the dimension of the attractor. Thus an estimate of D is given by the slope in a log-log plot of E against k or N_s. The line (A) in figure 3.3 is an approximation to such a fit and has slope 0.4 suggesting an attractor size of $D = 5$. In figure 3.10 we plot the prediction error against the number of points N_s in the sample set. The figure shows that for $(d_n \geq 4, d_s = 1)$ and for $(d_n = 2, d_s = 2)$ the prediction error follows the scaling law (3.8) with slope -0.11 suggesting an attractor size of approximately 18. Thus this scaling law gives inconclusive answers.

The second important dependence of the prediction error on the prediction

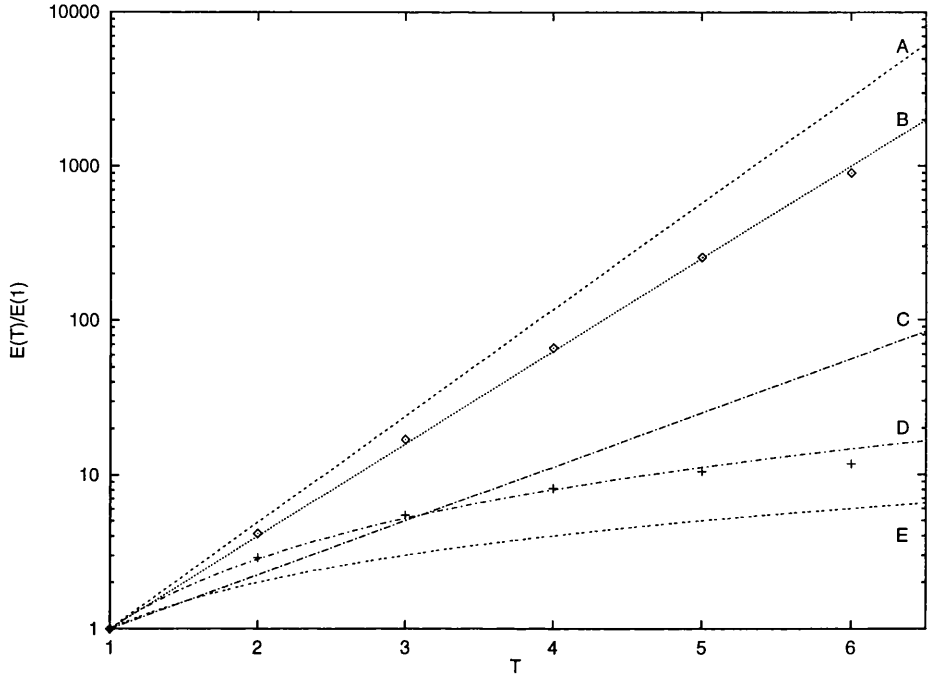


Figure 3.11: $E(T)/E(1)$ versus T on a semi-log scale. (+): $\epsilon = 0.45, k = 20, d_n = 2, d_s = 2$. ($\diamond$): Single logistic map ($\epsilon = 0.0, a = 2.0$) using $k = 20, d_n = 4, d_s = 1$. The lines depict: (A) scaling (3.10) with $h = 0.7$, (B) scaling (3.10) with $h = 0.6$, (C) scaling (3.10) with $h = 0.35$, (D) scaling (3.11) with $H = 0.75$, (E) Scaling (3.11) with $H = 0.5$.

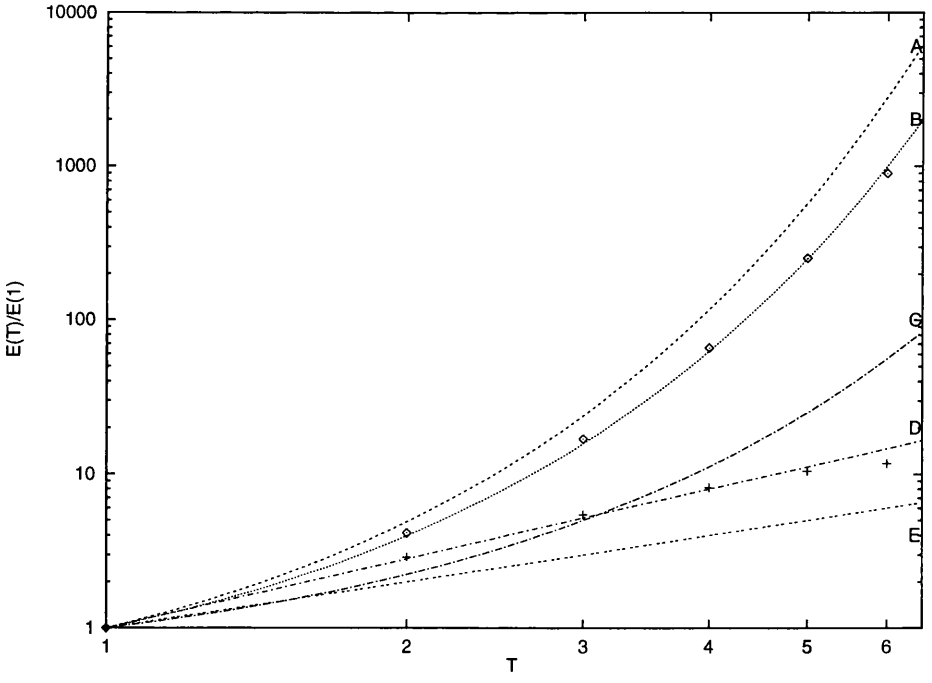


Figure 3.12: Same as figure 3.11 but with $E(T)/E(1)$ versus T on a log-log scale.

parameters is the prediction time T. Farmer and Sidorowich (Farmer & Sidorowich 1987) suggest that for direct prediction E scales with T as

$$E(T) \sim N_s^{-(p+1)/d} e^{(p+1)hT} \tag{3.9}$$

where h is the metric entropy. Thus for $h > 0$ the error E grows exponentially with T.

Sugihara and May (Sugihara & May 1990) noticed that the decay in $E(T)$ for increasing T did not occur for uncorrelated noise and hence saw the decay as a sign of determinism. However, Tsonis and Elsner (Tsonis & Elsner 1992) showed that coloured noise also exhibits this decay, but that it scales differently. More specifically, we expect the error in a deterministic system to scale with T as $E(T) = ce^{hT}$ where $c = N_s^{-(p+1)/d} e^{p+1}$. This implies

$$\log \frac{E(T)}{E(1)} = h(T-1) \tag{3.10}$$

Coloured noise, on the other hand, has the scaling

$$E(T) = E(1)T^{2H} \tag{3.11}$$

for small T where H is a scaling exponent related to the decay of the power spectrum, i.e. to the degree of correlation, or 'colour' in the noise (see (Osborne & Provenzale 1989)). This suggests that deterministic chaos can be detected by a linear correlation in a semi-log plot of $E(T)/E(1)$ against T rather that by a log-log plot where linear correlation would imply a stochastic process.

In figure 3.11 we have plotted $E(T)/E(1)$ versus T on a semi-log scale. The prediction errors for the CML are depicted by $(+)$ and we have also included the prediction errors for a single logistic map (with $a = 2.0$) for comparison as $(\diamond)$. We see that the errors for the single map seem to fit a straight line, while this is not the case for the CML. However, in the corresponding log-log plot in figure 3.12 we see the opposite, here the CML has a straight line fit whilst this is not the case for the single map. The correlation coefficient (r^2) was calculated for the pairs $(T, E(T)/E(1))$ and is shown in table 3.1.

	the semi-log plot r^2	the log-log plot r^2
single map	1.000	0.945
CML	0.891	0.992

Table 3.1: Degrees of linearity (r^2) of the semi-log and the log-log plots for the single logistic map ($\epsilon = 0.00$) and the coupled map lattice with $\epsilon = 0.45$. See figures 3.11 and 3.12

We thus see that the prediction error of the CML scales with T in the same way as coloured noise. Since we know that the system is actually deterministic, this strongly suggests that it must be high dimensional and provides an excellent illustration of the widely held view that high dimensional deterministic systems and stochastic systems are indistinguishable. We also remark that the power law scaling that we observe is a little reminiscent of that seen by Keeler and Farmer (Keeler & Farmer 1986).

The scaling eq. (3.10) can be used to estimate the largest Lyapunov exponent (Wales 1991). We know that the Lyapunov exponent for the single logistic map is $\lambda = \ln 2$ while we will see in chapter 5 that the CML have $\lambda_{\max} \approx 0.35$. Looking again at figure 3.11 the lines A and C have slopes 0.7 and 0.35 respectively. The fit through the first 6 points for the single map (line B) in figure 3.11 gives $h = 0.6$ i.e. an estimate of the Lyapunov exponent for the single map close to the real value. The CML however does not follow the scaling law (3.10) but instead follows the scaling law (3.11) with $H = 0.75$ (given by line D in figure 3.12). We are not thus able to extract the Lyapunov exponents in this case.

3.5 Discussion

We have shown that spatio-temporal reconstructions can increase predictability when compared to standard temporal or spatial delay maps. This seems natural because spatio-temporal techniques better capture the dynamics in a system with

local couplings. We would therefore like to encourage their use in the practical analysis of real spatio-temporal time-series. We have also found that increasing the spatial sampling interval gives decreasing predictability and demonstrated the close relation between cross predictability and space-time correlations.

Perhaps the most surprising aspect of our results is that the best predictions were obtained for low embedding dimensions (i.e. $d = 4$). This is despite the fact that there is considerable numerical evidence that the attractor of the system is high dimensional. In order to investigate this apparent discrepancy further we have computed a number of other measures of dimensionality commonly used in the analysis of chaotic time-series. The false nearest neighbour algorithm (Kennel, Brown & Abarbanel 1992) suggests an attractor dimension of 4. Given that this algorithm essentially tests for optimal prediction using a locally constant predictor, this is not unexpected given our predictability results. The correlation dimension (Grassberger & Procaccia 1983), on the other hand, does not converge to give a low-dimensional estimate. The scaling of prediction error with k and N_s, given above, yields dimension estimates of 5 and 18 respectively, whilst the Tsonis and Elsner criterion based on scaling with prediction time suggests a coloured stochastic process (i.e. an infinite dimensional system). Finally, the Lyapunov dimension density calculated using the sub-space Lyapunov spectrum (see chapter 5) is $\rho_L = 0.707$ giving an estimate for the Lyapunov dimension of approximately 70 for a lattice size of $L = 100$.

Our interpretation of these apparently contradictory results is that the attractor of the system is indeed high dimensional, and we would expect the value obtained from the Lyapunov dimension density to be the most reliable estimate of this. However, the vast majority of these dimensions are relatively small in extent, and the preponderance of the significant dynamics occurs in only a few dimensions. This permits us to make relatively good predictions using only a small embedding dimension.

This raises the question of an appropriate framework for embedding and reconstructing such systems. As indicated, we have no practical hope of rigorously embedding a 70 dimensional object, no matter how clever we are in the choice of em-

bedding variables. Yet clearly, in some approximate sense, reconstruction using delay variables does work. Our intuitive picture of this is that we can think of the system as a low dimensional local system, weakly coupled to a high dimensional global system, representing the rest of the CML. If we regard the latter as essentially stochastic, we can envisage the local system as a stochastically forced low dimensional deterministic system. A version of Takens Theorem applicable to this case has recently been proved (Stark, Broomhead, Davies & Huke 1997). More precisely let f_ω be a dynamical system on a k dimensional manifold M driven by a stochastic process $\omega = ..., \omega_{n-1}, \omega_n, \omega_{n+1}, ...$, so that $\mathbf{z}^{n+1} = f_{\omega_n}(\mathbf{z}^n) = f(\mathbf{z}^n, \omega_n)$. As before (see chapter 2.1.1), let $h : M \to \mathbb{R}$ be an observable on M, and construct a delay map $\Phi_\omega(\mathbf{z}) = (h(f^{-1}_{\omega_{-d}} \circ \ldots \circ f^{-1}_{\omega_{-1}}(\mathbf{z})), h(f^{-1}_{\omega_{-(d-1)}} \circ \ldots \circ f^{-1}_{\omega_{-1}}(\mathbf{z})), \ldots, h(f^{-1}_{\omega_{-1}}(\mathbf{z})))$. This gives $\mathbf{x}^n = \Phi_\omega(\mathbf{z}^n)$ where $\mathbf{x}^n = (x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$ is a delay vector reconstructed from the observed time-series. Under reasonable technical conditions it turns out that generically Φ_ω is an embedding for almost every ω if $d \geq 2m + 1$. We can then define a random map on the embedding space by $F_\omega = \Phi_{\sigma(\omega)} \circ f \circ \Phi_\omega^{-1}$ where $\sigma : \Omega \to \Omega$ is the shift $[\sigma(\omega)]_n = \omega_{n+1}$ and Ω is the space of all possible noise realisations (i.e. all possible sequences ω). Unravelling the definitions gives $F_\omega(\mathbf{x}^n) = \mathbf{x}^{n+1}$ and the last component of F can be written as $G_\omega(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) = x^n$ giving a time-series model appropriate to this case. Where ω is known (e.g. for input-output systems, or where it represents irregular sampling times) this directly leads to a prediction algorithm. Where it is unknown, it can potentially be estimated using Bayesian techniques, or if the noise level is low, the dependence of G on ω can be ignored, leading to an "average prediction" function:

$$G(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) = \int G_\omega(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1}) d\omega$$

Returning to our case of a CML, it is not strictly true that the high-dimensional global system is a random one. But suppose that we can divide the whole system into a low dimensional M, which we observe via $h : M \to \mathbb{R}$ and a high dimensional Ω, so that the dynamics is given by

$$\begin{aligned} \mathbf{z}^{n+1} &= f(\mathbf{z}^n, \omega_n) \\ \omega_{n+1} &= \sigma(\mathbf{z}^n, \omega_n) \end{aligned}$$

Note that something similar appears to be suggested in the last two sentences of (Kantz & Olbrich 1997). As usual, one can construct a delay map $\Phi_\omega : M \to \mathbb{R}^d$, which is now defined by

$$\Phi_\omega(\mathbf{z}) = (h(\pi(f,\sigma)^{-d}(\mathbf{z},\omega))), h(\pi(f,\sigma)^{d-1}(\mathbf{z},\omega))), \ldots, h(\pi(f,\sigma)^{-1}(\mathbf{z},\omega)))),$$

where $\pi(\mathbf{z},\omega) = \mathbf{z}$. Working through the definitions (see chapter 2.1.1) shows that we have $\Phi_{\omega_n}(\mathbf{z}^n) = \mathbf{x}^n$, where $\mathbf{x}^n$ is the reconstructed delay vector $\mathbf{x}^n = (x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$. Note that the dynamics on Ω is now no longer a simple shift, and more significantly it is no longer independent of $\mathbf{z}$. Nevertheless, in view of the results in ((Stark, Broomhead, Davies & Huke 1997; Stark 1999)), it seems reasonable to conjecture that under suitable technical assumptions, Φ_ω is an embedding for generic ω if $d \geq 2m+1$. This would then imply essentially the same conclusions as for the stochastic theorem described above. In particular, the time-series should satisfy an equation of the form $x^n = G_{\mathbf{x}^n,\omega_n}(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$. Note the double dependence of G on $\mathbf{x}^n = (x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$ through its arguments and through the subscript. The latter is due to the dependence of the dynamics of ω on $\mathbf{z}$. Assuming the interaction between ω and $\mathbf{z}$ is weak, we can ignore the subscripts on G in an analogous fashion to eq. (3.12) to yield an "approximate" prediction function $x^n = G(x^{n-d}, x^{n-d+1}, \ldots, x^{n-1})$. The existence of such a function is presumably what allows us to carry out reasonable predictions on the example presented in this chapter, and provides a rigorous framework for delay reconstruction approaches to spatio-temporal time-series.

Chapter 4

Truncated lattices and the thermodynamic limit

If we have a physical process whose evolution is governed by a large number of variables, whose precise interactions are *a priori* unknown, then we may be unable to decide on the basis of observed data whether the system is fundamentally deterministic or not. This has led to an informal classification of dynamical systems into two categories: low-dimensional deterministic systems and all the rest. The main aim of this chapter is to demonstrate that using observed data we may be unable to distinguish a spatio-temporal system from a local low-dimensional system perturbed by noise. Since the latter is much simpler, it may in many cases provide a preferable model of the observed data. On one hand this suggests that efforts to reconstruct the spatio-temporal dynamics of extended systems may be misplaced, and we should instead focus on developing methods to locally embed observed data. A preliminary framework for this was described in chapter 3 of this thesis. On the other hand, these results may help to explain why standard time-series methods sometimes work surprisingly well on data generated by high-dimensional spatio-temporal systems, where a priori they ought to fail: in effect such methods only see the "noisy" local system, and as long as the "noise level" is reasonably low can still perform at an acceptable level. Overall we see that we add a third category to the above informal classification: namely that of low di-

mensional systems driven by noise and we need to adapt our approach to observed time-series to take account of this.

Again consider the one-dimensional array of diffusively coupled logistic maps which we studied in chapter 3:

$$x_i^{n+1} = (1-\varepsilon)f(x_i^n) + \frac{\varepsilon}{2}(f(x_{i-1}^n) + f(x_{i+1}^n)), \tag{4.1}$$

where x_i^n denotes the discrete time dynamics at discrete locations $i = 1, \ldots, L$, $\varepsilon \in [0, 1]$ is the coupling strength and the local map f is the fully chaotic logistic map $f(x) = 4x(1-x)$. Recent research has focused on the thermodynamic limit, $L \to \infty$, of such dynamical systems (Pikovsky & Kurths 1994). Many interesting phenomena arise in this limit, including the rescaling of the Lyapunov spectrum (which we will discuss in chapters 5 and 6) and the linear increase in dimension (Lyapunov dimension (Parekh, Kumar & Kulkarni 1998) and fractal dimension (Puccioni, Torcini, Politi & G.D'Alessandro 1991)). The physical interpretation of such phenomena is that a long array of coupled systems may be thought of as a concatenation of small-size sub-systems that evolve almost independently from each other (Ruelle 1982; Kaneko 1989). As a consequence, the limiting behaviour of an infinite lattice is extremely well approximated by finite lattices of quite modest size. In our numerical work, we thus approximate the thermodynamic limit by a lattice of size $L = 100$ with periodic boundary conditions.

We mentioned in chapter 3 that numerical evidence suggests that the attractor in such a system is high-dimensional (Lyapunov dimension approximately 70). If working with observed data it is clearly not feasible to use an embedding dimension of that order of magnitude. On the other hand, we have shown that it is possible to make quite reasonable predictions of the evolution of a site using embedding dimensions as small as 4. This suggests that a significant part of the dynamics is concentrated in only a few degrees of freedom and that a low-dimensional model may prove to be a good approximation of the dynamics at a single site. In order to investigate this we further introduce the following truncated lattice. Let us take N sites ($i = 1, \ldots, N$) coupled as in equation (4.1) and consider the dynamics at the boundaries x_0^n and x_{N+1}^n to be produced by two independent driving inputs.

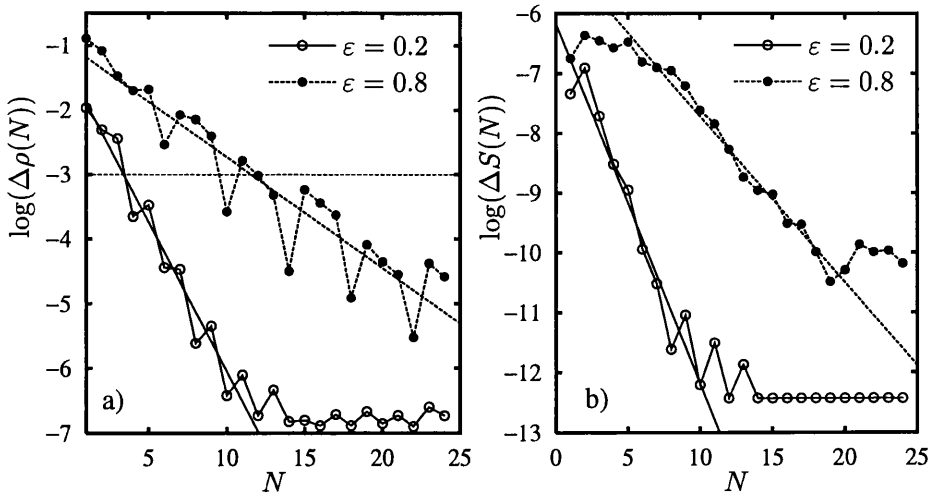


Figure 4.1: Distance between (a) the probability density and (b) the power spectra in the thermodynamic limit and its truncated lattice counterpart as the number of sites N in the latter is increased.

The driving input is chosen to be white noise equally distributed in the interval $[0, 1]$. We are interested in comparing the dynamics of the truncated lattice to the thermodynamic limit case.

4.1 Results

We begin the comparison between the two lattices by examining their respective invariant probability density at the central site (if the number of sites is even, either of the two central sites is equivalent). For a semi-analytic treatment of the probability density of large arrays of coupled logistic maps see (Lemaître, Chaté & Manneville 1997). Let us denote by $\rho_\infty(x)$ the single site probability density in the thermodynamic limit and $\rho_N(x)$ the central site probability density of the truncated lattice of size N. We compare the two densities in the $\mathcal{L}_1$ norm by computing

$$\Delta\rho(N) = \int_0^1 |\rho_\infty(x) - \rho_N(x)|\, dx \tag{4.2}$$

for increasing N. The results are summarised in figure 4.1.a where $\log(\Delta\rho(N))$ is plotted for increasing N for different values of the coupling. The figure suggests that the difference between the densities decays exponentially as N is increased

(see straight lines for guidance). Similar results were obtained for intermediate values of the coupling parameter. The densities used to obtain the plots in figure 4.1.a were estimated by a box counting algorithm by using 100 boxes and 10^8 points (10^2 different orbits with 10^6 iterations each). The maximum resolution typically achieved by using these values turns out to be around $\Delta\rho(N) \simeq \exp(-6.5) \simeq 0.0015$. This explains the saturation of the distance corresponding to $\varepsilon = 0.2$. For $\varepsilon = 0.8$ the saturation would occur for approximately $N = 35$ given enough computing power (more refined boxes and more iterations). Nonetheless, densities separated by a distance of approximately $\exp(-3) \simeq 0.05$ (see horizontal threshold in figure 4.1.a), or less, capture almost all the structure. Therefore, one recovers the essence of the thermodynamic limit probability density with a reasonable small truncated lattice (see figures 4.2.a,b).

Next we compare temporal correlations in the truncated lattice with those in the full system. Denote by $S_\infty(\omega)$ the power spectrum of the thermodynamic limit and $S_N(\omega)$ its counterpart for the truncated lattice. Figure 4.1.b shows the difference $\Delta S(T)$ in the $\mathcal{L}_1$ norm between the power spectra of the truncated lattice and of the thermodynamic limit for $\varepsilon = 0.2$ and 0.8 (similar results were obtained for intermediate values of ε). As for the probability density, the power spectra appear to converge exponentially with the truncated lattice size. Note that for large N, particularly for small ε, the difference tends to saturate around $\exp(-12) \approx 10^{-6}$, this is because the accuracy of our power spectra computations reaches its limit (with more iterations one can reduce the effects of the saturation). Our results were obtained by averaging 10^6 spectra ($|\mathrm{DFT}|^2$) of 1024 points each. In figures 4.2.c,d we depict the comparison between the spectra corresponding to the thermodynamic limit and to the truncated lattice. As can be observed from the figure, the spectra for the truncated lattice give a good approximation to the thermodynamic limit. It is worth mentioning that the spectra depicted in figures 4.2.c,d are plotted in logarithmic scale so to artificially enhance the discrepancy of the distance between the thermodynamic limit and the truncated lattice. The distance corresponding to these plots lies well below $\Delta S(T) < \exp(-7.5) \approx 5 \times 10^{-4}$. The convergence of the power spectrum is much faster than the one for the

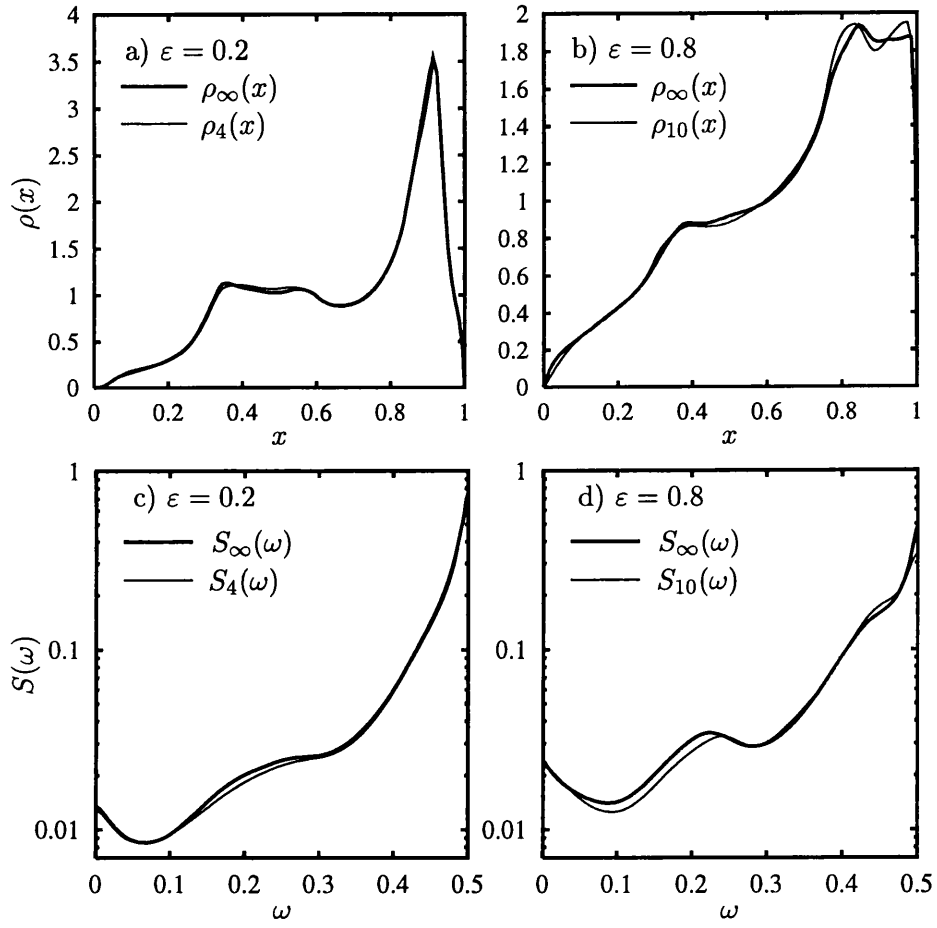


Figure 4.2: Approximating (a,b) the probability density and (c,d) the power spectra of the thermodynamic limit (thick lines) using a truncated lattice (thin lines).

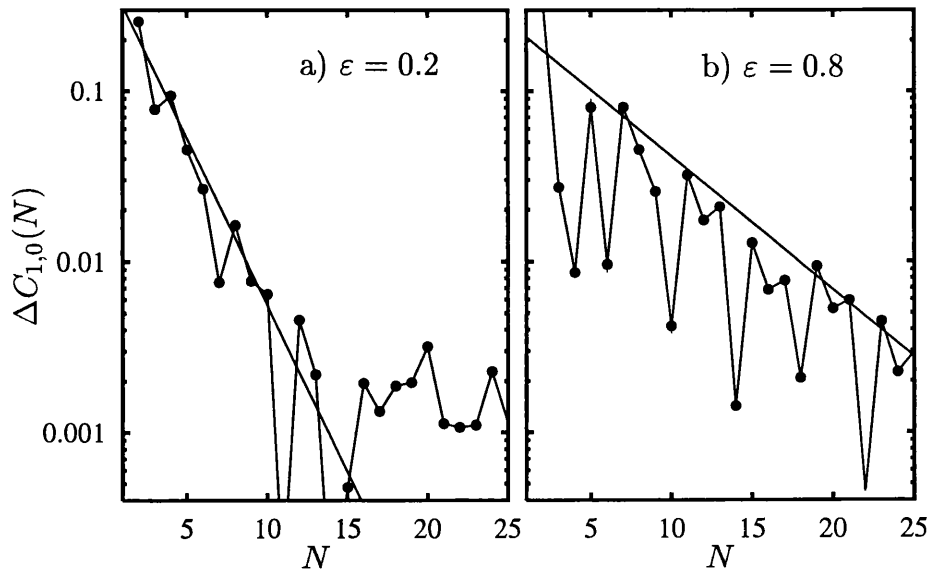


Figure 4.3: Difference of the two-point correlation between the truncated lattice and the thermodynamic limit for two neighbours at the same iteration ($C(\xi = 1, \tau = 0)$).

probability density (compare both scales in figures 4.1).

To complete the comparison picture we compute the two-point correlation which we defined in eq. (3.7). $C(\xi, \tau)$ corresponds to the correlation of two points in the lattice dynamics separated by ξ sites and τ time steps. To obtain the two-point correlation for the truncated lattice we consider the two points closest to the central site separated by ξ. We then compute $\Delta C_{\xi,\tau}(N)$ defined as the absolute value of the difference of the correlation in the thermodynamic limit with that obtained using the truncated lattice of size N. In figure 4.3 we plot $\Delta C_{1,0}(N)$ as a function of N for $\varepsilon = 0.2$ and 0.8. For $\varepsilon = 0.2$, due to limited accuracy of our calculations, the saturation is reached around $N = 10$. Nonetheless it is possible to observe an exponential decrease (straight lines in the linear-log plot) before the saturation. For larger values of ε the exponential convergence is more evident (see figure 4.3.b). Similar results were obtained for intermediate ε-values. Note that because the correlation oscillates, it is not possible to have a point by point exponential decay for $\Delta C_{1,0}(N)$, however, the upper envelope clearly follows an exponential decay (see straight lines for guidance). Similar results were obtained

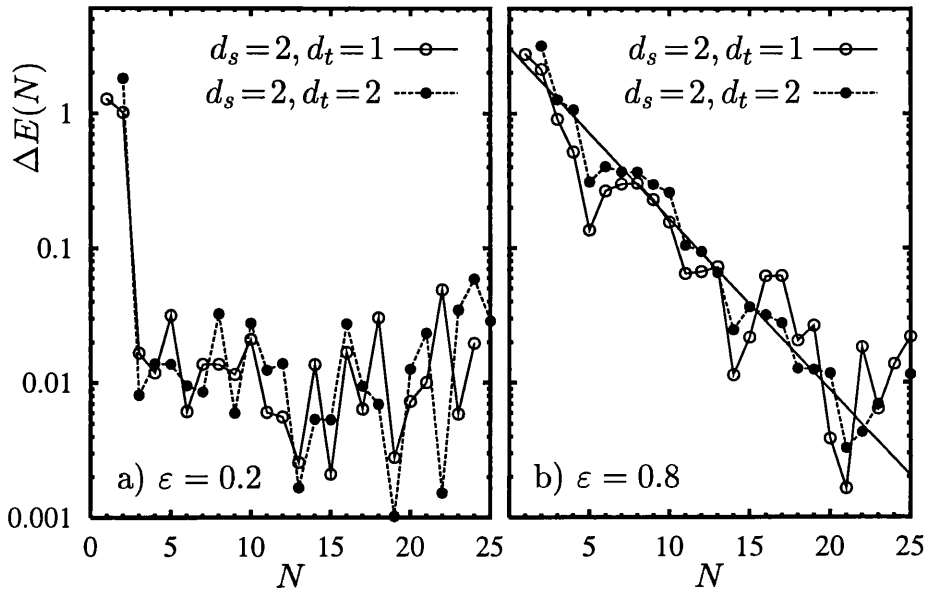


Figure 4.4: Normalised one-step prediction error difference (4.3) between a truncated lattice and the thermodynamic limit for two spatio-temporal embeddings ($(d_s, d_n) = (2, 1)$ and $(d_s, d_n) = (2, 2)$) and different couplings strengths.

for different values of (ξ, τ).

The above comparisons were carried out by using the data produced by the known system (4.1). Often, in practice, one is deprived of the evolution laws of the system. In such cases, the only way to analyse the system is by using time-series reconstruction techniques. This is particularly appropriate when dealing with real spatio-temporal systems where, typically, only a fraction of the set of variables can be measured or when the dynamics is only indirectly observed by means of a scalar measurement function. In the following we suppose that the only data available about the system is provided by the time-series of a set of variables in a small region of space. We would like to study the effects on predictability when using a truncated lattice instead of the thermodynamic limit.

Instead of limiting ourselves to one-dimensional time-series (temporal embedding) we use the mix of temporal and spatial delay embeddings (spatio-temporal embedding) that we introduced in chapter 3. Therefore we use the delay map eq. (3.6) (with $\tau_m = \tau_s = 1$). The delay map is used to predict x_i^n. Recall from chapter 3 that the best one-step predictions using the delay map (3.6) are typi-

cally obtained for $d_s = d_n = 2$. Here we use the two cases $(d_s, d_n) = (2, 1)$ and $(d_s, d_n) = (2, 2)$; almost identical results are obtained for higher dimensional embeddings ($(d_s, d_n) \in [1, 4]^2$). Denote by $E(N)$ the normalised root-mean square error for the one step prediction using the delay map at the central portion of the truncated lattice of size N. The comparison between $E(N)$ and $E(N \to \infty)$ is shown in figure 4.4 where we plot the absolute value of the normalised error difference

$$\Delta E(N) = |(E(N) - E(\infty))/E(\infty)| \tag{4.3}$$

for increasing N and for different spatio-temporal embeddings and coupling strengths. The figure shows a rapid decay of the prediction error difference for small N and then a saturation region where the limited accuracy of our computation hinders any further decay. For $\varepsilon = 0.2$ the drop to the saturation region is almost immediate while for the large coupling value $\varepsilon = 0.8$ the decay is slow enough to observe an apparently exponential decay (see fitted line corresponding to $d_s = d_n = 2$ for $N = 1, \ldots, 20$), thereafter the saturation region is again reached. For intermediate values of the coupling ε, the saturation region is reached between $N = 5$ and 20. Before this saturation it is possible to observe a rapid (exponential) decrease of the normalised error difference. This corroborates again the fact that it seems impossible in practice to differentiate between the dynamics of the relatively small truncated lattice and the thermodynamic limit.

All the results in this chapter where obtained from the simulation of a truncated lattice with white noise inputs at the boundaries. Other kinds of inputs did not change our observations in a qualitative way. It is worth mentioning that a truncated lattice with random inputs with the *same* probability density as the thermodynamic limit ($\rho_\infty(x)$) produces approximatively the same exponential decays as above with just a downward vertical shift (i.e. same decay but smaller initial difference).

The numerical results shown here correspond to locally coupled map lattices. It is clear that the nature of the coupling plays an important role in the phenomenology hereby presented. In order to check the effects of including a more global

coupling we also studied the dynamics of large lattices of coupled maps with an exponentially decreasing coupling: $x_i^{n+1} = (1-\beta)/(1+\beta)\sum_{k=-\infty}^{\infty}\beta^{|i-k|}f(x_{i-k}^n)$, where $\beta \in (0,1)$ measures the decay of the coupling. We found that for small $\beta(\beta < 0.3)$ it is possible to model the dynamics at a single site for the thermodynamic limit with a relatively small truncated lattice. However, as the coupling becomes more global for larger values of β, a subtle collective coherence emerges and we were unable to obtain promising results from approximating a large lattice by a truncated one. It is well known that globally coupled maps are prone to a subtle collective behaviour even though coherence of individual sites is not present (Kaneko 1992; Sinha, Biswas, Azam & Lawande 1992). In such cases, the idea of replacing a potentially infinite lattice with a truncated lattice with random inputs breaks down. In particular, the violation of the law of large numbers reported in (Kaneko 1992; Sinha, Biswas, Azam & Lawande 1992) will not occur for the truncated lattice.

4.2 Discussion

The properties of the thermodynamic limit of a coupled logistic lattice we have considered here (probability densities, power spectra, two-point correlations and predictability) were approximated remarkably well (exponentially close) by a truncated lattice with random inputs. Therefore, when observing data from a limited spatial region, *given a finite accuracy in the computations* and a reasonably small truncated lattice size, it would be impossible to discern any dynamical difference between the thermodynamic limit lattice and its truncated counterpart. The implications from a spatio-temporal systems time-series perspective are quite strong and discouraging: even though in theory one should be able to reconstruct the dynamics of the *whole* attractor of a spatio-temporal system from a local time-series (Takens theorem), it appears that due to the limited accuracy (due to machine precision, CPU time and memory limitations, measurement errors, limited amount of data) it would be impossible to test for definite high-dimensional determinism in practice.

The evidence presented here suggests the impossibility of reconstructing the state of the whole lattice from localised information. It is natural to ask whether we can do any better by observing the lattice at many (possibly all) different sites. Whilst in principle this would yield an embedding of the whole high-dimensional system, it is unlikely to be much more useful in practice. This is because the resulting embedding space will be extremely high-dimensional and any attempt to characterise the dynamics, or fit a model will suffer from the usual "curse of high dimensionality". In particular, with any realistic amount of data, it will be very rare for typical points to have close neighbours. Hence, for instance, predictions are unlikely to be much better than those obtained from just observing a localised part of the lattice.

If one actually wants to predict the behaviour at many or all sites, the results suggest that the best approach is to treat the data as coming from a number of uncoupled small noisy systems, rather than a single large system. Then one would have to decide the optimal size of the local system. One way of doing this would be to choose the embedding dimension for which the predictability is highest (cfr. chapter 3). Of course, if one has good reason to suppose that the system is spatially homogeneous, one should fit the same local model simultaneously at all spatial locations, thereby substantially increasing the amount of available data. If one had data from l neighbouring sites and had found that the optimal local size to be d sites, then the potential neighbours would increase up by a maximum multiple of $(l - d + 1)$.

Chapter 5

Interleaving and rescaling of the Lyapunov spectrum

The computation of the entire Lyapunov spectrum for spatio-temporal dynamical systems is a very time consuming task. The high number of variables, and even the number of effective degrees of freedom, often leads to severe difficulties because of the large amount of resources (computing time and memory space) required for many computations. Therefore it is useful, and often crucial, to develop techniques that derive information about the whole system by analyzing a comparatively small sub-system.

We saw in chapter 2 that for dynamical systems with only a few degrees of freedom the computation of the LS is a straightforward task; however, when the number of degrees of freedom gets large (e.g. a few hundred) it becomes a painstaking process (Grassberger 1989; Puccioni, Torcini, Politi & G.D'Alessandro 1991; Bauer, Heng & Martienssen 1993). In particular, any algorithm to compute the LS must contain two fundamental procedures; one to multiply by the Jacobian at each time step and the other to perform some kind of reorthonormalisation (Geist, Parlitz & Lauterborn 1990). The latter is required to prevent the Jacobian matrix progressively getting more ill-conditioned, until the largest Lyapunov exponent swamps all the others. Such orthogonalisation procedures are based upon the factorisation of the Jacobian matrix into a product of an orthogonal matrix Q and an

upper triangular matrix R. The two most widespread methods for achieving such orthogonalisation are based upon modified Gram-Schmidt (MGS) orthogonalisation and the so-called HQR decomposition that uses Householder transformations. The MGS-based methods are widely used because of their quite simple numerical implementation though they are known to introduce small errors due to the fact that the orthogonality of the matrix Q may fail. The HQR-based methods are more difficult to implement but they give a better approximation of the LS (von Bremen, Udwadia & Proskurowski 1997) since they do not have the problem of losing orthogonality of the matrix Q. The difficulty in using any of these methods for computing the LS of systems with a high number of degrees of freedom N is that they require $\mathcal{O}(N^3)$ operations (von Bremen, Udwadia & Proskurowski 1997). The usual naive algorithm for matrix multiplication is also $\mathcal{O}(N^3)$, so that overall computing the full LS is an $\mathcal{O}(N^3)$ process (in principle matrix multiplication can be done faster than $\mathcal{O}(N^3)$ using specialized techniques, but this hardly seems worth doing under the circumstances). As an example, the computation of the LS using a HQR method for a logistic coupled map lattice with $N = 100$ takes a few hours on a standard workstation. When the system size is an order of magnitude larger (e.g. for two or more spatial dimensions) and/or the convergence of the Lyapunov exponents is rather slow, the task quickly becomes infeasible. Therefore one must rely on other techniques to approximate the LS for large systems.

One such technique to estimate the LS in a fully spatio-temporal chaotic regime is to take a principal sub-matrix of the Jacobian and compute the LS for this sub-system. It has been observed in a wide range of spatio-temporal systems that such a sub-system LS converges to the spectrum of the whole system under appropriate rescaling. In a number of specialized cases e.g. turbulent Navier-Stokes flows (Ruelle 1982) and hard sphere gases (Sinai & Chernov 1982; Sinai 1996) there are rigorous results for this phenomenon. However is seems difficult to prove its occurrence more generally, and certainly there are many systems where it is observed numerically but no rigorous analysis exists. These include coupled logistic maps (Grassberger 1989), chaotic neural networks (Bauer & Martienssen 1991), coupled map lattices (Kaneko 1989; Kaneko 1986), reaction-diffusion systems (Parekh,

Kumar & Kulkarni 1996; Parekh, Kumar & Kulkarni 1997) (lattice of ODEs), turbulent fluids (Ruelle 1983), the Kuramoto-Sivashinsky model (Manneville 1985) (PDE's), and others.

Such a rescaling approach consists of evolving the *whole* N-dimensional system under the equations of motion, taking a *sub-system* of size N_s to compute the LS and then rescaling it to obtain an estimate of the original LS. This method relies on the linear increase of Lyapunov dimension D_L and KS entropy h (see chapter 2 for definitions) with the sub-system size (see above references). A physical interpretation of this phenomenon can be given in terms of the thermodynamic limit of the system. A spatio-temporal system in a fully chaotic regime will possess a typical correlation length ξ such that elements further apart than ξ evolve almost independently from each other. The whole system can then be thought of in some sense as the union of several almost independent sub-systems of size ξ. In the limit when these sub-systems are completely uncoupled the LS repeats itself in each one of them. If an interaction between the sub-systems is introduced, one may expect the overall LS not to be significantly altered. Thus in the limit of a large number of degrees of freedom, a number of Lyapunov exponents per ξ-volume may be defined. One expects such an intuitive picture to become more accurate in the limit of a large number of degrees of freedom and a small correlation length.

As we will see shortly, when examining closer the Lyapunov spectra in the fully chaotic regime for several spatio-temporal systems we found that the Lyapunov exponents of two consecutive sub-system sizes N_s and $N_s + 1$ were interleaved. In other words, the ith Lyapunov exponent for the sub-system N_s lies between the ith and $(i+1)$th Lyapunov exponents of the sub-system N_s+1. The interleaving of the eigenvalues for a single matrix is a well-known fact (Cauchy's interlace theorem) and is common in many areas such as Sturm sequences of polynomials (Parlett 1980). Unfortunately there appears to be no obvious generalisation which would imply the same fact for sub-system Lyapunov Spectra.

5.1 Interleaving and rescaling for homogeneous states

In order to gain some insight into interleaving and rescaling behaviour of the Lyapunov spectrum in extended dynamical systems let us start with the simplest case of all: homogeneous evolution. Recall the coupled map lattice equation 2.11. We define *homogeneous states* as states of the form $X_n = \{x_i^n\}_{i=1}^N$ where $x_i^n = x^n$ is the same for all i. It is trivial that by setting the initial state of the lattice to a homogeneous state $x_i^0 = x^0$ one has that $X_n = \{f^n(x^0)\}$ for all i at any future time n. In other words the homogeneity of the initial state is preserved under iteration by (2.11).

Let us take a simple form for the coupling by using the most widespread model of a CML, the so called *diffusive* CML:

$$x_i^{n+1} = (1-\varepsilon)f(x_i^n) + \frac{\varepsilon}{2}\left(f(x_{i-1}^n) + f(x_{i+1}^n)\right), \tag{5.1}$$

where the coupling is symmetric and only between nearest neighbours. We shall perform a linear stability analysis of homogeneous states in this system. Such an analysis for more general CML's has also served as the starting point for the study of signal propagation (Carretero-González 1997) and pattern formation (Gade & Amritkar 1993). Since (5.1) preserves homogeneity under iteration it is natural to ask whether the stability of f completely determines the stability of the homogeneous state. The answer turns out to be yes.

Recall from chapter 2 that the Lyapunov exponents λ_i are given by the logarithms of the eigenvalues of the matrix

$$\Gamma = \lim_{n\to\infty} \left[P(n)^{\mathrm{tr}} \cdot P(n)\right]^{1/2n} \tag{5.2}$$

where

$$P(n) = J(n) \cdot J(n-1) \cdots J(2) \cdot J(1)$$

and where $J(s)$ is the Jacobian matrix of the CML dynamics at time s and $(\,\cdot\,)^{\mathrm{tr}}$ denotes matrix transpose. For the homogeneous lattice

$$J(n) = \mu_n \cdot M \tag{5.3}$$

where $\mu_n = f'(x^n)$ is the multiplier of the local map and M is the constant matrix

$$M = \begin{pmatrix} 1-\varepsilon & \varepsilon/2 & 0 & \cdots & \varepsilon/2 \\ \varepsilon/2 & 1-\varepsilon & \varepsilon/2 & \cdots & 0 \\ 0 & \varepsilon/2 & 1-\varepsilon & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \varepsilon/2 & \cdots & 0 & \varepsilon/2 & 1-\varepsilon \end{pmatrix}.$$

The matrix M is not only symmetric but also *circulant.* Recall that a matrix is circulant if in each successive row the elements move to the right one position (with wrap–around at the edges) (Davis 1979). It is straightforward to prove (Bellman 1960) that the eigenvalues of a circulant matrix

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{k-1} \\ c_{k-1} & c_0 & \cdots & c_{k-2} \\ \vdots & \ddots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix}$$

are given by $c_0 + c_1 r_j + \cdots + c_{k-1} r_j^{k-1}$, where $r_j = \exp(2\pi \mathrm{i} j/N)$ is an Nth root of unity. Thus, the eigenvalues $\beta_j(n)$ of $J(n)$ are given by

$$\begin{aligned} \beta_j(n) &= \mu_n \left((1-\varepsilon) + \frac{\varepsilon}{2}(r_j + r_j^{N-1}) \right) \\ &= \mu_n \, \phi_j(\varepsilon, N), \end{aligned}$$

where

$$\phi_j(\varepsilon, N) = (1-\varepsilon) + \varepsilon \cos\left(\frac{2\pi j}{N}\right). \tag{5.4}$$

It is important to notice that $\phi_j(\varepsilon, N)$ does not depend on the iteration n: the time dependence has been decoupled (factorized) into μ_n. The Lyapunov exponents are

then given by

$$\begin{aligned}\lambda_i &= \lim_{t\to\infty} \ln \prod_{n=1}^{t} |\beta_i(n)|^{1/t} \\ &= \lim_{t\to\infty} \ln \left(|\phi_i(\varepsilon, N)| \prod_{n=1}^{t} |\mu_n|^{1/t} \right) \\ &= \ln |\phi_i(\varepsilon, N)| + \lim_{t\to\infty} \frac{1}{t} \sum_{n=1}^{t} \ln |\mu_n|.\end{aligned}$$

Thus by defining λ_0 to be the Lyapunov exponent of a typical orbit of a single, uncoupled, local map, starting at x^0: $\lambda_0 = \lim_{t\to\infty}(1/t) \sum_{n=1}^{t} \ln |\mu_n|$, one obtains the following expression for the Lyapunov exponents of a homogeneous evolution:

$$\lambda_i = \lambda_0 + \ln |\phi_i(\varepsilon, N)|. \tag{5.5}$$

Note that the Lyapunov exponents defined by (5.5) are not arranged in decreasing order. Re-indexing them in decreasing order they become

$$\lambda_k = \begin{cases} \lambda_0 + \ln |\phi_{\frac{k}{2}}(\varepsilon, N)| & k \text{ even} \\ \\ \lambda_0 + \ln |\phi_{\frac{k-1}{2}}(\varepsilon, N)| & k \text{ odd} \end{cases} \tag{5.6}$$

where $k = 1$ to N. It is clear that $\lambda_k = \lambda_{k+1}$ when k is even, so most of the exponents occur in degenerate pairs, apart from the largest, and, if N is even, the smallest. The linear stability of a homogeneous orbit is then characterized by the Lyapunov exponent λ_0 of a single site in the uncoupled case ($\varepsilon = 0$). In particular, if the local map is not chaotic then the homogeneous evolution is not chaotic either since $\lambda_k \leq \lambda_0$ ($|\phi_k(\varepsilon, N)| \leq 1$ for all k).

It is interesting to notice that the same shape for the LS of a homogeneous CML (*cf.* (5.5)) is obtained for a lattice of coupled Bernoulli shifts (Vannitsem & Nicolis 1996) for *any* orbit. There is however an important difference: while in the CML the LS dependence on the actual orbit was decoupled thanks to the homogeneity, in the case of coupled Bernoulli shifts, the LS is decoupled from the orbit because the derivative of the local map at any point is constant. Examining further this similarity, if one takes the fully chaotic logistic map ($4\,x\,(1-x)$) as the local map for the diffusive CML, the LS for the homogeneous evolution is

$$\lambda_i = \ln 2 + \ln |\phi_i(\varepsilon, N)|. \tag{5.7}$$

In fact, any one-dimensional map whose Lyapunov exponent is $\lambda_0 = \ln 2$ gives rise to the LS (5.7) under homogeneous evolution. The LS (5.7) corresponds exactly to the LS of a lattice of coupled Bernoulli shifts and thus the results described below for the rescaling of the sub-system LS are valid for the case of a lattice of coupled Bernoulli shifts.

Now let us perform the LS analysis for a sub-system of the original CML. Thus instead of taking all the sites $i = 1, \ldots, N$ we take N_s sites starting at any position j. The choice of j is not important since we are dealing with periodic boundary conditions and because the state is homogeneous; from now on we choose $j = 1$. Thus, we take a principal sub-matrix J' of size $N_s \times N_s$ from the whole Jacobian J. In matrix terms $J' = \pi(J)$ where π is the following projection

$$\pi(J) = \Pi_l \cdot J \cdot \Pi_r \tag{5.8}$$

with the left (Π_l) and right (Π_r) projection matrices defined as

$$\Pi_l = \left(\begin{array}{c|c} I & Z \end{array} \right)$$

$$\Pi_r = \left(\begin{array}{c} I \\ \hline Z^{\mathrm{tr}} \end{array} \right)$$

where, from now on, I is the $N_s \times N_s$ identity matrix and Z is the $N_s \times (N - N_s)$ null matrix. Therefore, in order to compute the Lyapunov exponents for the truncated system one has to compute the following product of projected matrices

$$\begin{aligned} P'(n) &= (\Pi_l J(n) \Pi_r) \cdots (\Pi_l J(2) \Pi_r)(\Pi_l J(1) \Pi_r) \\ &= \Pi_l J(n) \Pi_c J(n-1) \cdots J(2) \Pi_c J(1) \Pi_r, \end{aligned} \tag{5.9}$$

where

$$\Pi_c = \Pi_r \cdot \Pi_l = \left(\begin{array}{c|c} I & Z \\ \hline Z^{\mathrm{tr}} & 0 \end{array} \right).$$

Multiplying equation (5.9) from the left by the identity matrix obtained by $\Pi_l \cdot \Pi_r = I$ yields

$$\begin{aligned} P'(n) &= \Pi_l \Pi_r \Pi_l J(n) \Pi_c J(n-1) \cdots J(2) \Pi_c J(1) \Pi_r \\ &= \Pi_l \Pi_c J(n) \Pi_c J(n-1) \cdots J(2) \Pi_c J(1) \Pi_r \\ &= \pi\left(\tilde{P}(n)\right), \end{aligned} \tag{5.10}$$

where we define the new product $\tilde{P}(n) = K(n) \cdots K(2) K(1)$ of the projected matrices $K(i) = \Pi_c J(i)$.

Using the above description we obtain the sub-system LS for the homogeneous evolution. The projected Jacobian for the homogeneous evolution at time n is

$$J'(n) = \pi(J(n)) = \mu_n \pi(M) = \mu_n \cdot M' \tag{5.11}$$

where $M' = \pi(M)$ is the $N_s \times N_s$ constant matrix

$$M' = \begin{pmatrix} 1-\varepsilon & \varepsilon/2 & 0 & \cdots & 0 \\ \varepsilon/2 & 1-\varepsilon & \varepsilon/2 & \cdots & 0 \\ 0 & \varepsilon/2 & 1-\varepsilon & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \varepsilon/2 & 1-\varepsilon \end{pmatrix}$$

if $N_s < N$, and $M'_N = M$ if $N_s = N$. From now on we only use the notation M' when $N_s < N$. It is important to notice that by taking a sub-Jacobian matrix the periodicity of the boundary conditions is lost. The dynamics of the sub-system at the boundaries could be thought as being coupled to some external noise coming from the adjacent sites. Thus, in contrast to M, the matrix M' is not circulant, however its eigenvalues are well known to be (Barnett 1990)

$$\phi'_j(\varepsilon, N_s) = (1-\varepsilon) + \varepsilon \cos\left(\frac{\pi j}{N_s + 1}\right). \tag{5.12}$$

(where $j = 1$ to N_s); and so the eigenvalues $\beta'_j(n)$ of $J'(n)$ are $\beta'_j(n) = \mu_n \, \phi'_j(\varepsilon, N_s)$. The sub-system LS is given by

$$\lambda'_j = \lambda_0 + \ln |\phi'_j(\varepsilon, N_s)|. \tag{5.13}$$

One can immediately infer from this that the Lyapunov exponents for the homogeneous evolution are interleaved for two consecutive sub-system sizes. More precisely, suppose that we take two sub-systems, one of size N_s and the other of size $N_s + 1$. It is then trivial to see that their respective Lyapunov exponents $\lambda'_i(N_s)$ and $\lambda'_i(N_s + 1)$ satisfy:

$$\lambda'_i(N_s + 1) \leq \lambda'_i(N_s) \leq \lambda'_{i+1}(N_s + 1) \quad \forall 1 \leq i \leq N_s, \tag{5.14}$$

see figure 5.1.a. Interleaving of the sub-system LS with respect to the whole LS $\lambda_i(N)$ also occurs:

$$\lambda_i(N) \leq \lambda'_i(N_s) \leq \lambda_{i+N-N_s}(N) \quad \forall 1 \leq i \leq N_s.$$

This interleaving of the eigenvalues is a consequence of Cauchy's interlace theorem (Parlett 1980) that gives bounds on the eigenvalues of a principal sub-matrix given the eigenvalues of the original matrix. It is important to notice that the interleaving property of the Lyapunov exponents for the homogeneous case is a straightforward consequence of the decoupling of the time dependence of the Jacobian matrix leaving us with the constant matrices M and M'. In a typical non-homogeneous evolution the time dependence of the Jacobian cannot be factorized and an equivalent constant matrix for the Jacobian does not exist. Therefore, Cauchy's interlace theorem cannot be applied in this general case and there is no reason a priori for the interleaving property to hold for a generic extended dynamical system. It is true that, at any particular time, there is interleaving between the eigenvalues of the whole Jacobian and those of a sub-system. However, when computing the LS, one has to compute the product of the Jacobian matrices while for the sub-system LS one uses the product of the sub-Jacobian matrices and therefore the interleaving of the matrix product is no longer assured. The only way, a priori, for the interleaving to work would be to take the product of the whole Jacobians *first* and only then extract the sub-Jacobian. The problem with this procedure is that one has to rely again on re-orthonormalisation procedures involving the original matrix size N, making the task impossible for large N. Nevertheless, as we shall see in the following section, the interleaving of the sub-system LS does hold to a great extent in the thermodynamic limit.

Another important point to note from equations (5.12) and (5.13) is that the LS of the sub-systems all have the same shape. The best way to see this is to rescale the indices of the Lyapunov exponents so that they lie in the range [0,1]: so instead of plotting λ against j we plot it against $j/(N_s+1)$. Equations (5.12) and (5.13) then show that the points always lie on the graph of the function

$$\lambda(z) = \lambda_0 + \ln\left[(1-\varepsilon) + \varepsilon\cos(\pi z)\right], \tag{5.15}$$

irrespective of the value of N_s. This observation suggests another way of looking at the interleaving property. For a given N_s, the z values of the sub-system LS are equally spaced in the interval [0,1]; if we increase N_s by 1 the new z values interleave with the old. Since λ is a monotone function the fact that the z values interleave means that the $\lambda(z)$ values interleave also. It is worthwhile mentioning that we are considering the simple case $1-2\varepsilon > 0$ so the absolute value inside the logarithm in equation (5.13) can be omitted. For $1-2\varepsilon \leq 0$ the eigenvalues need further re-indexing in order to maintain their decreasing order and a similar construction as below is possible.

To compare the sub-system LS with that of the full system we should similarly rescale the indices for the latter, so now we plot the full system Lyapunov exponents against $j/(N+1)$ instead of j. The points of this spectrum do not lie on the graph of $\lambda(z)$; however, equation (5.6) shows that they do lie on the graphs of the functions

$$\lambda_{\text{even}}(z) = \lambda\left(z\left(1+1/N\right)\right)$$

(for exponents with even indices) and

$$\lambda_{\text{odd}}(z) = \lambda\left(z\left(1+1/N\right) - 1/N\right)$$

(for exponents with odd indices), where the function $\lambda(z)$ is given by equation (5.15). Since $z(1+1/N) - 1/N < z < z(1+1/N)$ $(0 < z < 1)$ and $\lambda(z)$ is a decreasing function we see that $\lambda_{\text{even}}(z) < \lambda(z) < \lambda_{\text{odd}}(z)$. Thus λ_{even} and λ_{odd} are bounding curves for λ (see thin dashed lines in figure 5.1.b) and converge to it as $N \to \infty$; the differences between λ and the other curves are $\mathcal{O}(1/N)$.

The similarities between the shapes of the Lyapunov spectra of the sub-systems and of the whole system mean we can use the sub-system LS to estimate the

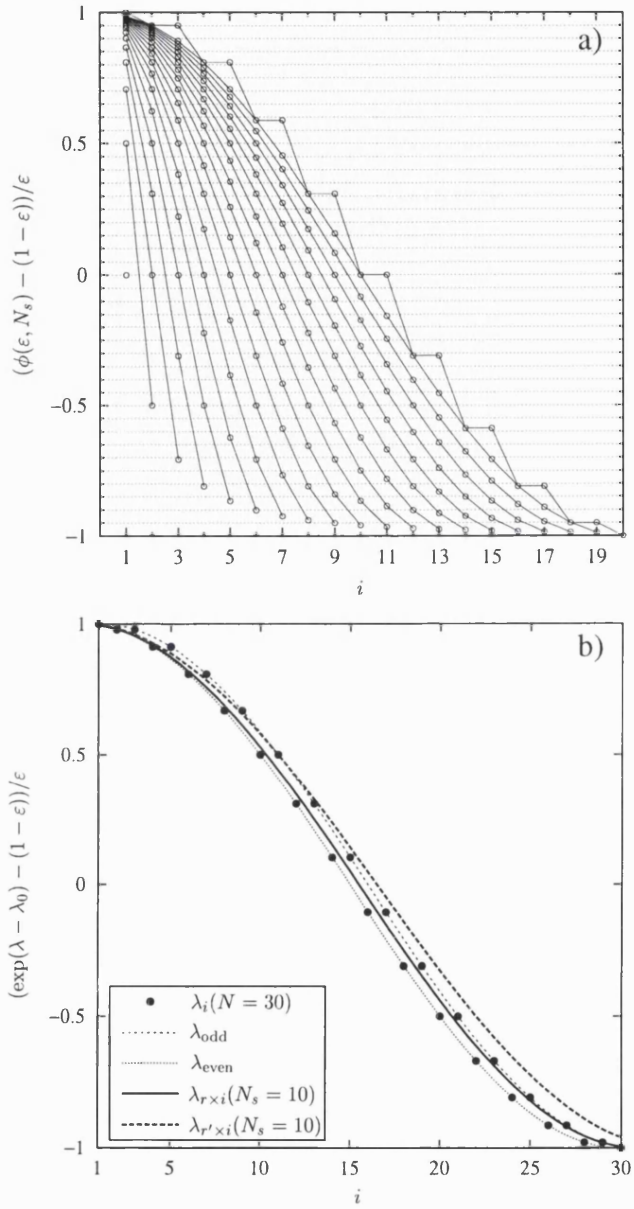


Figure 5.1: Lyapunov spectrum for a homogeneous evolution in a diffusive CML: a) interleaving for sub-system sizes $N_s = 1, \ldots, 20$ ($N = 20$); b) rescaled sub-system LS, the circles represent the whole LS ($N = 30$), the thin dashed lines represent the functions λ_{odd} and λ_{even} passing through the eigenvalues for even and odd indexes respectively, while the thick lines represent the rescaled LS with $N_s = 10$ using the conventional rescaling $r' = N/N_s$ (thick dashed line) and the new rescaling obtained in section 5.1 $r = (N+1)/(N_s+1)$ (thick solid line).

whole LS: to do this we rescale the indices of the sub-system exponents, plotting λ_j against rj where r is a factor chosen so that the rescaled sub-system LS lies as close as possible to the plot of the full system LS. The above discussion shows that if we choose

$$r = \frac{N+1}{N_s+1} \tag{5.16}$$

then the rescaled sub-system LS differs from the full system LS by an amount of $\mathcal{O}(1/N)$.

The scaling given by (5.16) differs from that used conventionally, which is performed by scaling by

$$r' = \frac{N}{N_s}$$

see (Grassberger 1989; Bauer & Martienssen 1991; Kaneko 1989; Kaneko 1986; Mayer-Kress & Kaneko 1989). It is clear however that using r' will give results that differ from those using r by terms of $\mathcal{O}(1/N_s)$, and since this is larger than $\mathcal{O}(1/N)$ the errors in the exponents will also be $\mathcal{O}(1/N_s)$. This suggests that scaling (5.16) should give more accurate results than the conventional scaling; this is certainly true in the homogeneous case. As an example figure 5.1.b shows the original LS for a homogeneous CML with $N = 30$ (circles) along with the rescaled LS with $N_s = 10$ using the conventional rescaling r' (dashed line) and the new rescaling r obtained above (solid line). It is clear that the new rescaling gives a much better approximation to the original LS.

5.2 Interleaving and rescaling for coupled logistic maps

As mentioned in the previous section, the interleaving property for the homogeneous evolution relies on the fact that the Jacobian matrices can be factorized into a time dependent scalar and a *time independent* matrix (see equations (5.3) and (5.11)). For a non-homogeneous evolution the Jacobians cannot be factorized in such a way and thus *a priori* one does not expect interleaving to occur. Surprisingly enough the numerical evidence points towards interleaving of the sub-system

LS for almost every Lyapunov exponent in the fully developed chaotic regime. In this section we shall present such evidence for the familiar logistic CML, and discuss why such behaviour might be expected to occur. More general systems will be considered in the following section.

We thus consider the diffusive CML (5.1) with the fully chaotic logistic map $f(x) = 4x(1-x)$ as we have in the earlier chapters and compute its LS for several values of the coupling parameter ε. As with all numerical work in this chapter we employ a fast HQR algorithm for the computation of Lyapunov exponents (von Bremen, Udwadia & Proskurowski 1997). We then calculate the sub-system LS using principal sub-matrices J' of size $N_s = 1, \ldots, 30$ of the Jacobian. In doing so one is not taking into account the dynamics of the neighbouring sites next to the boundary and their effects are considered as noise. Thus, the algorithm consists in computing the LS of the sub-Jacobian J' by truncating the actual Jacobian J at each time step and then applying the HQR algorithm. The results are shown in figure 5.2 where we plot the sub-system LS for increasing sub-system size ($N_s = 1, \ldots, 30$) for 3 different values of the coupling parameter. In the figure, the filled circles represent the Lyapunov exponents that do not fulfil the interleaving condition. Strikingly, the LS corresponding to $\varepsilon = 0.05$ and $\varepsilon = 0.45$ (figures a and b) are very well interleaved, with the exception of a couple of points. On the other hand, for $\varepsilon = 0.95$ (figure 5.2.c) the LS is not that well interleaved for the smallest Lyapunov exponents, although for the large ones the interleaving is as good as for the previous two figures. The reason for this failure for the smallest Lyapunov exponents is that in the limit $\varepsilon \to 1$ the lattice decouples into two independent sub-lattices: one for odd i and the other for even i. Thus, when successively increasing the sub-system size, one is including in turn contributions from the even and the odd sub-lattice. This is reflected in a variation in the smallest Lyapunov exponents every time we increase the sub-system size by one, hence the bi-periodic nature of the interleaving failure. In fact, by removing the sub-system LS for odd sizes one ends up with almost perfect interleaving. The exact reasons and conditions for the interleaving of the sub-system LS to happen are not yet understood, however we believe that they are connected with the convergence of

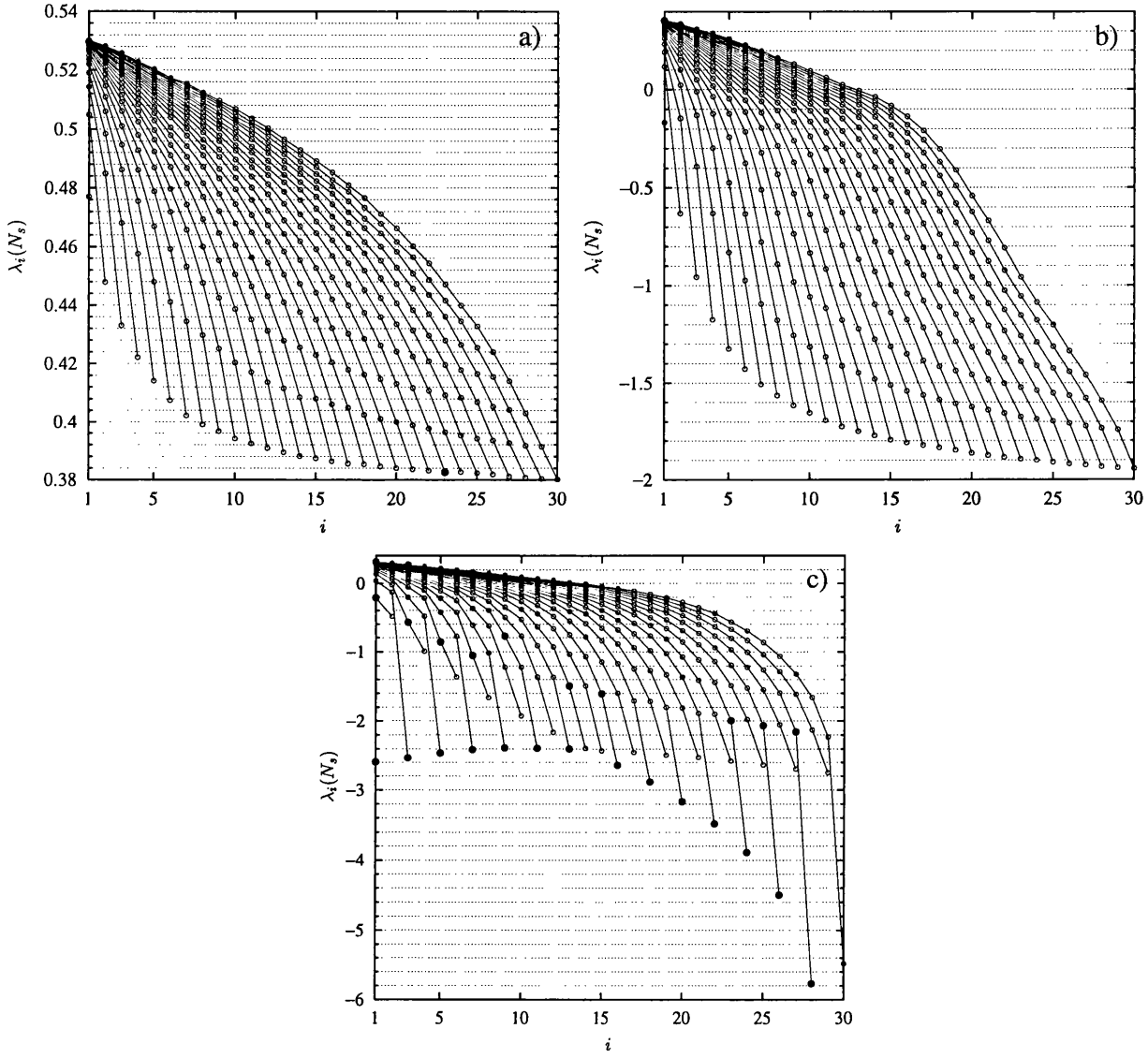


Figure 5.2: Sub-system Lyapunov spectra for the fully chaotic coupled logistic lattice $N = 100$ for sub-system sizes 1 to 30 (left to right) for a) $\varepsilon = 0.05$, b) $\varepsilon = 0.45$ and c) $\varepsilon = 0.95$. The filled circles represent those Lyapunov exponents which fail to interleave.

the sub-system LS to the full system LS—a convergence which may be expected in the thermodynamic limit, see below.

As mentioned in the introduction to this chapter, it has been observed for some time that under appropriate rescaling the sub-system LS approximates the whole LS. The usual argument for this rescaling behaviour makes use of the thermodynamic limit. In the previous section, while studying the interleaving of sub-system LS for the homogeneous case, a new rescaling was suggested (see equation (5.16)). Let us test this for the case of the fully chaotic coupled logistic lattice. In figure 5.3 we compare, for $\varepsilon = 0.05$ and $\varepsilon = 0.45$, the rescaled sub-system LS using the new rescaling $r = (N+1)/(N_s+1)$ (5.16) (circles) and the conventional one $r' = N/N_s$ (crosses) to the whole LS (lines) for different sub-system sizes ($N_s = 15, \ldots, 25$). As is clear from the figures, the new rescaling r gives a much better fit to the original LS than the conventional rescaling.

Let us explore the idea of rescaling the sub-system LS in the thermodynamic limit a bit further. The correspondence between the rescaled LS and the whole LS in figure 5.3 is astonishingly good. The rescaled spectra lie almost perfectly on top of a decreasing curve, therefore, as with the homogeneous case discussed above, it is not surprising that they are interleaved. In general, if the rescaled Lyapunov spectra of the sub-systems converge sufficiently quickly to the whole system LS we expect to have good interleaving of the sub-system Lyapunov spectra. On the other hand, if the rescaled sub-system LS do not approximate the whole system LS well, it is not clear that interleaving will occur. To illustrate this we present the rescaled LS using the new rescaling r for $\varepsilon = 0.95$ in figure 5.4. In this case, the rescaled LS do not give such a good approximation to the whole LS (in particular for the smallest Lyapunov exponents) as seen in the other cases ($\varepsilon = 0.05$ and $\varepsilon = 0.45$). As explained above, this is due to the decoupling of the whole lattice into two sub-lattices when $\varepsilon \to 1$. Therefore it appears that the non-interleaving of the smallest Lyapunov exponents in figure 5.2.c is related to the lack of convergence of the sub-system LS. In general we suppose that failure to interleave is an indication that the sub-system LS have not converged. Clearly however, the presence of interleaving is not a sure indication that convergence has

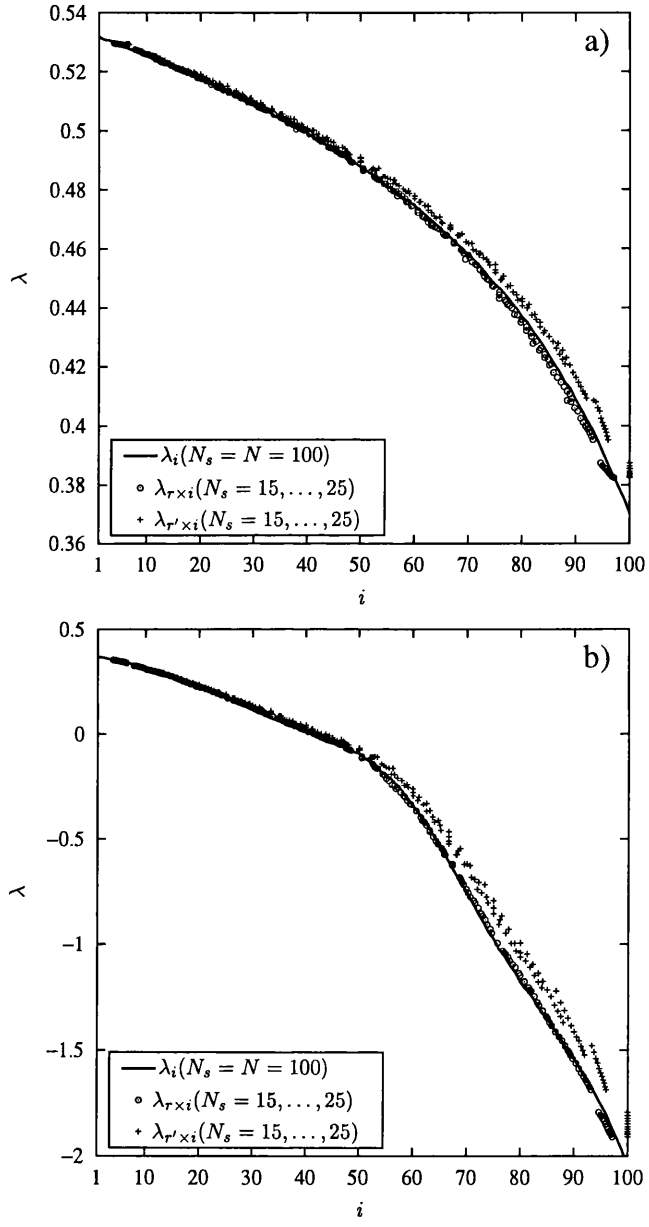


Figure 5.3: Comparison of the whole Lyapunov spectrum (solid line) and the rescaled sub-system Lyapunov spectrum using the new rescaling r (circles) and the conventional rescaling r' (crosses) in the fully chaotic logistic lattice with $N = 100$ for several sub-system sizes ($N_s = 15, \ldots, 25$). a) $\varepsilon = 0.05$ and b) $\varepsilon = 0.45$.

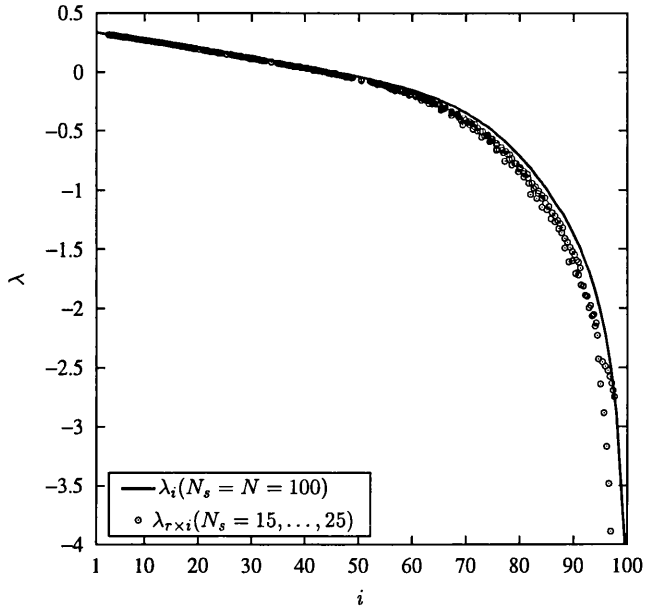


Figure 5.4: Rescaled Lyapunov spectrum (circles) for the coupled logistic lattice with $\varepsilon = 0.95$ for sub-system sizes $N_s = 15, \ldots, 30$. The solid line represents the whole Lyapunov spectrum $N_s = N = 100$.

occurred; this is illustrated by the two-dimensional logistic lattice discussed later.

We believe that the key point in understanding the interleaving behaviour is that although in computing the sub-system LS one is using the product of projected matrices (5.10), one does not modify the original dynamics in any way. Recall that similar matrices share eigenvalues. Thus a feasible explanation for the occurrence of interleaving is to hypothesize that the product of the projected matrices $P'(n)$ is a projection of a $N \times N$ matrix $\mathcal{Q}(\infty)$ which is similar to the limit as $n \to \infty$ of the original product $P(n)$ of the whole Jacobians. In other words, we conjecture that there exists an invertible $N \times N$ matrix S such that

$$\mathcal{Q}(\infty) = \lim_{n\to\infty} S^{-1} P(n) S, \tag{5.17}$$

where the product of the projected matrices $P'(n)$ in the limit is obtained by projecting $\mathcal{Q}(\infty)$:

$$P'(\infty) = \pi\left(\mathcal{Q}(\infty)\right).$$

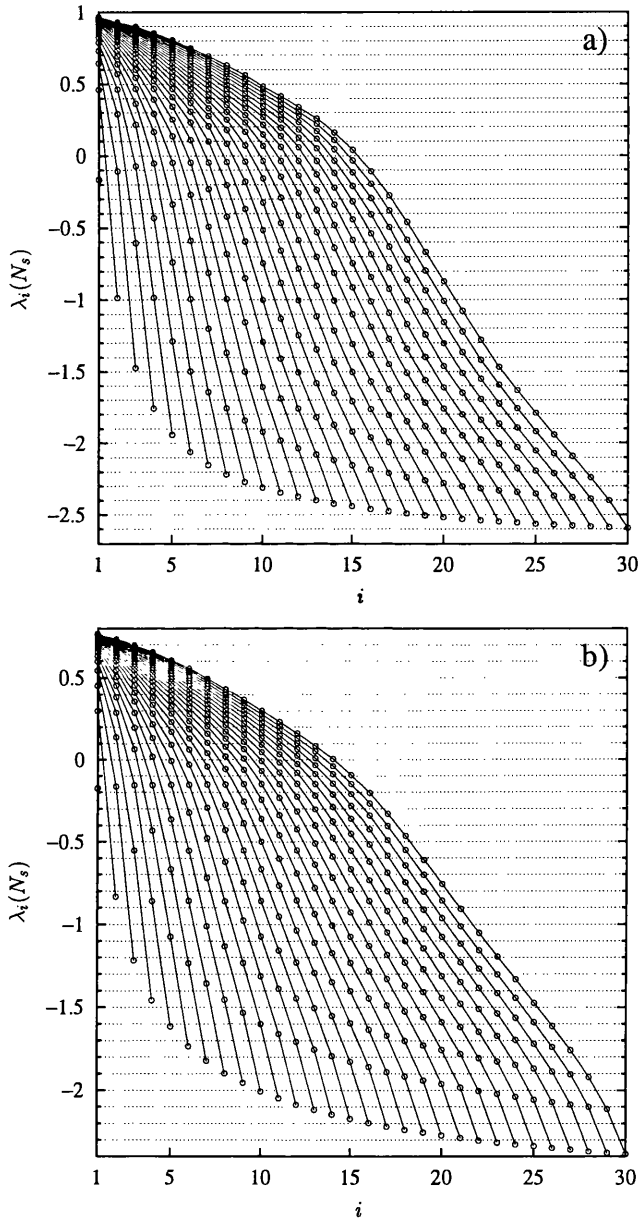


Figure 5.5: Interleaving of the sub-system LS, $N_s = 1, \ldots, 30$, for the fully chaotic logistic lattice with $\varepsilon = 0.45$ using the more general projection matrices (5.18) to extract the sub-system Jacobians: a) Π_1 and b) Π_2.

Informally, this is saying that in some sense in equation (5.10) the projection matrices commute on average with the Jacobians in the $n \to \infty$ limit. We believe that it might be possible to make this statement rigorous by an appropriate generalisation of the multiplicative ergodic theorem.

One might then also ask what is so special about the projection Π_c. Is it possible for interleaving to occur for more general projections? The following two examples suggest that this is indeed the case. Consider the following projection matrices

$$\begin{aligned} \Pi_1 &= \left(\begin{array}{c|c} \Pi_1' & Z \\ \hline Z^{\mathrm{tr}} & 0 \end{array}\right) \\ \Pi_2 &= \left(\begin{array}{c|c} \Pi_2' & Z \\ \hline Z^{\mathrm{tr}} & 0 \end{array}\right) \end{aligned} \tag{5.18}$$

where Z is the $N_s \times (N - N_s)$ null matrix and the $N_s \times N_s$ matrices Π_1' and Π_2' are

$$\Pi_1' = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\Pi_2' = \begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1N_s} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2N_s} \\ 0 & 0 & 1 & \cdots & \alpha_{3N_s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

where the α_{ij} are random numbers chosen from the interval $[0, 1]$ with equal probability. Note that we are still using the term projection matrices for Π_1 and Π_2 which in a strict sense is not correct, since they do not satisfy $\Pi_j = \Pi_j^2$ $(j = 1, 2)$. We use this terminology to stress the fact that they completely remove some of the entries of the original Jacobian. Thus, instead of taking the projection matrix

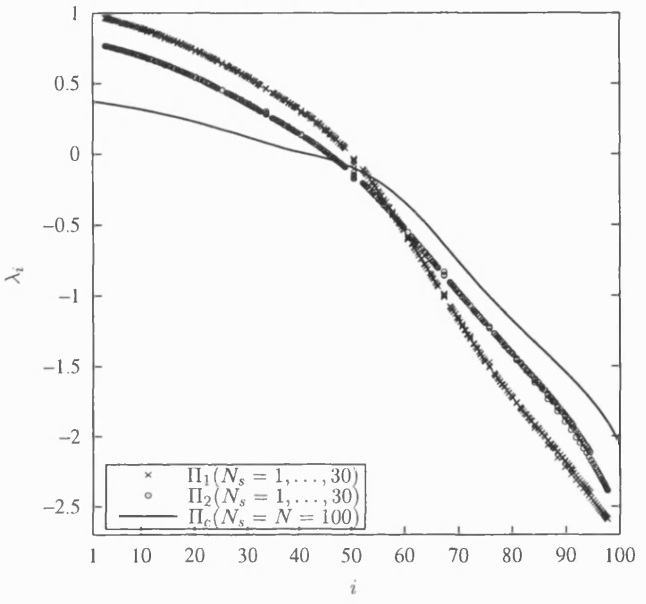


Figure 5.6: Rescaled sub-system LS corresponding to figure 5.2 using the projection matrices Π_1 (crosses) and Π_2 (circles). The continuous line corresponds to the original LS computed with the whole Jacobian.

Π_c let us take Π_1 and Π_2. For the projection Π_2 we randomise its entries *every* time-step; similar results were obtained by randomising only at the beginning and keeping the same projection matrix thereafter. In figure 5.5 we depict the non-rescaled sub-system LS using both projection matrices for the fully chaotic logistic lattice. The figure strongly suggests that in these cases interleaving still occurs. It thus appears that the choice of projection matrix is not a crucial ingredient for interleaving. Nonetheless it is important to say that we do not expect interleaving to hold if one uses a series of projection matrices such that when computing the LS one does not get convergence. In the above examples, Π_1 and Π_2, we do have the required convergence. For the Π_2 case, the convergence of the LS of their product is a well known fact (Furstenberg & Kesten 1960; Johnson, Palmer & Sell 1987).

On the other hand, when we turn to rescaling we find that although for Π_1 and Π_2 we still get convergence of the rescaled sub-system LS to a definite limit, this limit is not the original LS for the full system (figure 5.6). The reason for this discrepancy is easy to understand since the new projections Π_1 and Π_2 combine the

entries of the projected Jacobians and thus one expects the eigenvalues to change.

5.3 Estimation of quantities derived from the Lyapunov spectrum

As illustrated in the previous section, the LS can be well approximated by the rescaled sub-system LS in the thermodynamic limit. We now use the new rescaling in order to approximate the original LS by extrapolating from the sub-system LS. We estimate the largest Lyapunov exponent, Lyapunov dimension and KS entropy and we compare our method to the results obtained with the whole LS and with the conventional rescaling.

The first method to approximate quantities derived from the LS in the thermodynamic limit is by defining intensive quantities from the extensive ones by simply using the corresponding densities (Parekh, Kumar & Kulkarni 1998; Parekh, Kumar & Kulkarni 1997; Bauer & Martienssen 1991; Mayer-Kress & Kaneko 1989; Puccioni, Torcini, Politi & G.D'Alessandro 1991). Let us define the densities of eqs. (2.9) and (2.10):

$$\begin{aligned} \rho_d(N_s) &= \frac{D_L}{N_s} \\ \rho_h(N_s) &= \frac{h}{N_s} \end{aligned} \tag{5.19}$$

corresponding to the Lyapunov dimension density and the KS entropy density respectively. In the thermodynamic limit these densities are intensive quantities (i.e. they do not depend on the sub-system size taken). One then estimates their extensive counterpart when $N_s \to N$ by multiplying the densities (5.19) by N. To estimate the largest Lyapunov exponent for the whole system we directly take the value of the largest Lyapunov exponent of the sub-system (the Lyapunov exponents are not extensive quantities). It is worth mentioning that in order to use these intensive densities to estimate extensive ones we are supposing the size N of the original system to be known.

The second method, which we believe is more accurate, consists of taking the sub-system LS, rescaling it, extrapolating a curve through it to obtain an approx-

imation to the whole LS and only then computing the desired quantities. There are several ways to extrapolate the whole LS from the sub-system LS; here we have chosen a piece-wise linear approximation for simplicity. One could use more accurate methods such as cubic splines but the aim here is to compare both kinds of rescaling and thus a piece-wise linear fit is the most straightforward approach. Therefore, take the rescaled LS $\lambda_i(N_s)$, obtained with either rescaling for a sub-system of size N_s, and consider the polygon $\mathcal{P}$ through all the points $(i, \lambda_i(N_s))$. To estimate a Lyapunov exponent of the whole LS lying between $\lambda_1(N_s)$ and $\lambda_{N_s}(N_s)$ one simply uses the fit given by the polygon $\mathcal{P}$. For Lyapunov exponents lying to the left (right) of the polygon use linear extrapolation from the first (last) two points of the rescaled LS. Here again one could use more sophisticated extrapolation methods but for simplicity we restrict ourselves to the linear one. Once the whole LS is estimated using the above method, or a more complicated one, quantities such as the largest Lyapunov exponent $\lambda_1(N)$, the Lyapunov dimension D_L and the KS entropy h are easily extracted.

In figure 5.7 we compare the estimates of a) the largest Lyapunov exponent $\lambda_1(N)$, b) the Lyapunov dimension D_L and c) the KS entropy h obtained from the intensive densities (diamonds) and the piece-wise linear fitting for both rescalings (conventional rescaling with crosses and the proposed new one with circles) as the sub-system size increases for the coupled logistic lattice. The actual values of these quantities calculated with the whole LS correspond to the horizontal lines. For the largest Lyapunov exponent, figure 5.7.a, we notice that the estimates are almost identical for both rescalings (crosses and circles). This is due to the fact that both rescalings tend to coincide for small i (see figure 5.7.b). The estimate of the largest Lyapunov exponent by just taking the largest Lyapunov exponent of the sub-system (diamonds) shows a slower convergence than the linear fit methods. For the Lyapunov dimension, figure 5.7.b, the method with the slowest convergence corresponds to the conventional rescaling (crosses), while the approximations derived from densities (diamonds) and from a linear fit with the new rescaling (circles) are quite good (note that the new rescaling method does better than the approach using densities). Finally, for the KS entropy, figure

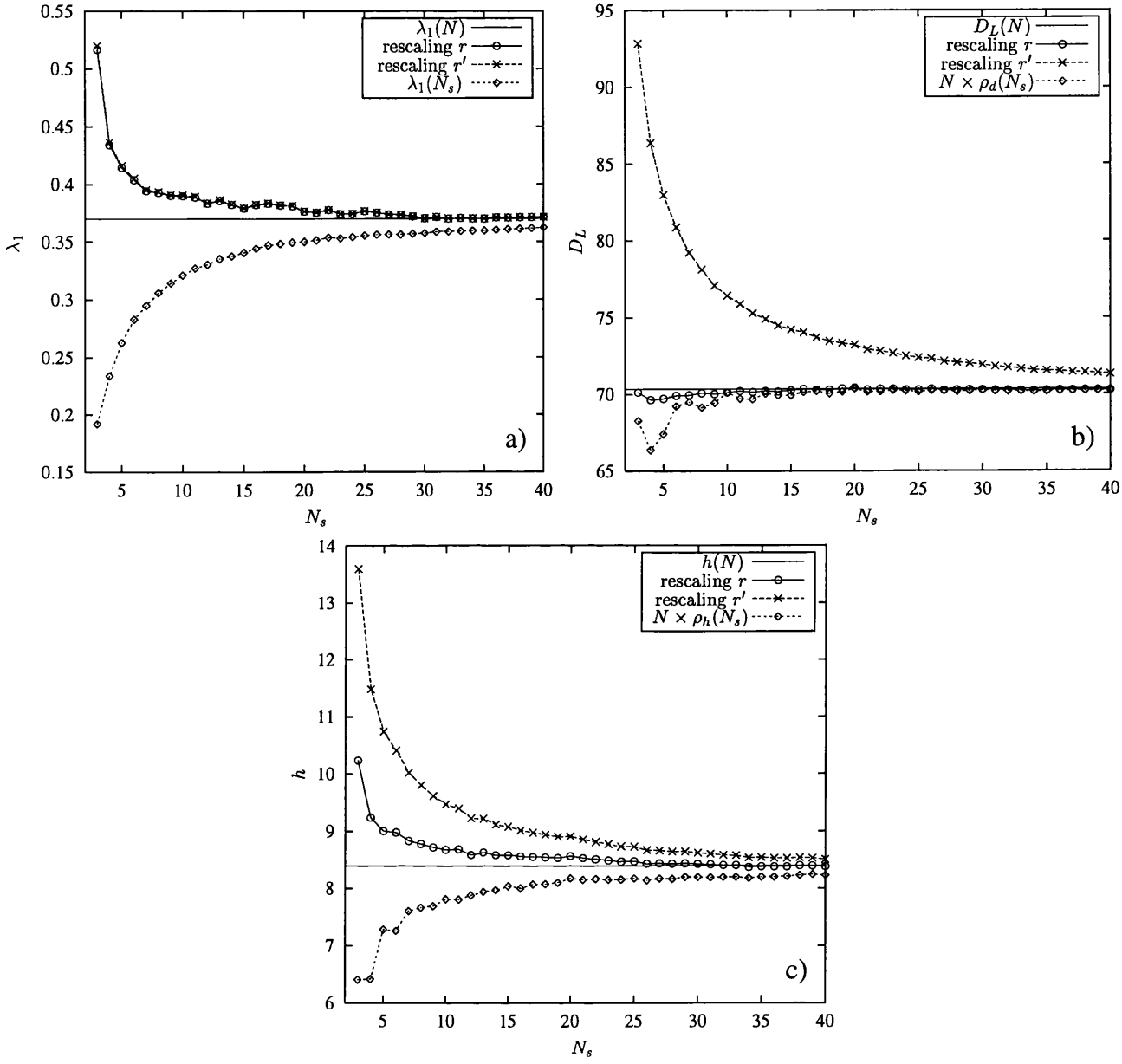


Figure 5.7: Estimation of a) the largest Lyapunov exponent, b) the Lyapunov dimension and c) the KS entropy as a function of the sub-system size N_s in the coupled logistic lattice with $N = 100$ and $\varepsilon = 0.45$. The estimates obtained by using a) the largest Lyapunov exponent of the sub-system and b–c) the associated densities from the sub-system are presented with diamonds, and the estimate obtained from the piece-wise linear fit to the rescaled LS is presented with crosses for the conventional rescaling and circles for the proposed new one. The values obtained with the whole LS are represented by the horizontal line.

5.7.c, the estimates using the density (diamonds) and the conventional rescaling (crosses) have similar convergence rates while the new rescaling method (circles) does considerably better. The evidence given by this set of plots tends to indicate that the new rescaling method gives better convergence to the quantities derived from the sub-system LS.

5.4 More general extended dynamical systems

So far we have only considered interleaving and rescaling in systems in one spatial dimension with nearest neighbour coupling, corresponding to tridiagonal Jacobians. In this section we turn to more general kinds of extended dynamical systems by allowing a larger coupling range (e.g. chaotic neural networks) and by taking a different topology for the lattice (e.g. lattice with two spatial dimensions). The results presented in this section suggest that the interleaving and rescaling properties observed for the simpler one-dimensional CML persist for more general extended dynamical systems.

5.4.1 Chaotic neural networks

We now consider a chaotic neural network (Bauer & Martienssen 1991) of the form

$$x_i^{n+1} = \tanh\left(g \sum_{l=i-k}^{i+k} C_{il} x_l^n\right), \tag{5.20}$$

where g is a real number called the gain parameter, k represents the connectivity (essentially playing the same role as the range of the coupling in a CML) and the weight matrix C_{ij} has entries chosen randomly from $[-1, 1]$ with a uniform probability distribution for $(i-j) \pmod N \leq k$ and $C_{ij} = 0$ otherwise.

Both the CNN and CML dynamics work in two stages—nonlinearity and coupling—but their order is inverted. The CML dynamics applies the nonlinear mapping f first and then the coupling, while the CNN first applies the coupling via a linear weighted combination of neighbouring sites, and then a nonlinear map (the sigmoid). This inversion is reflected in the Jacobian matrix of the transfor-

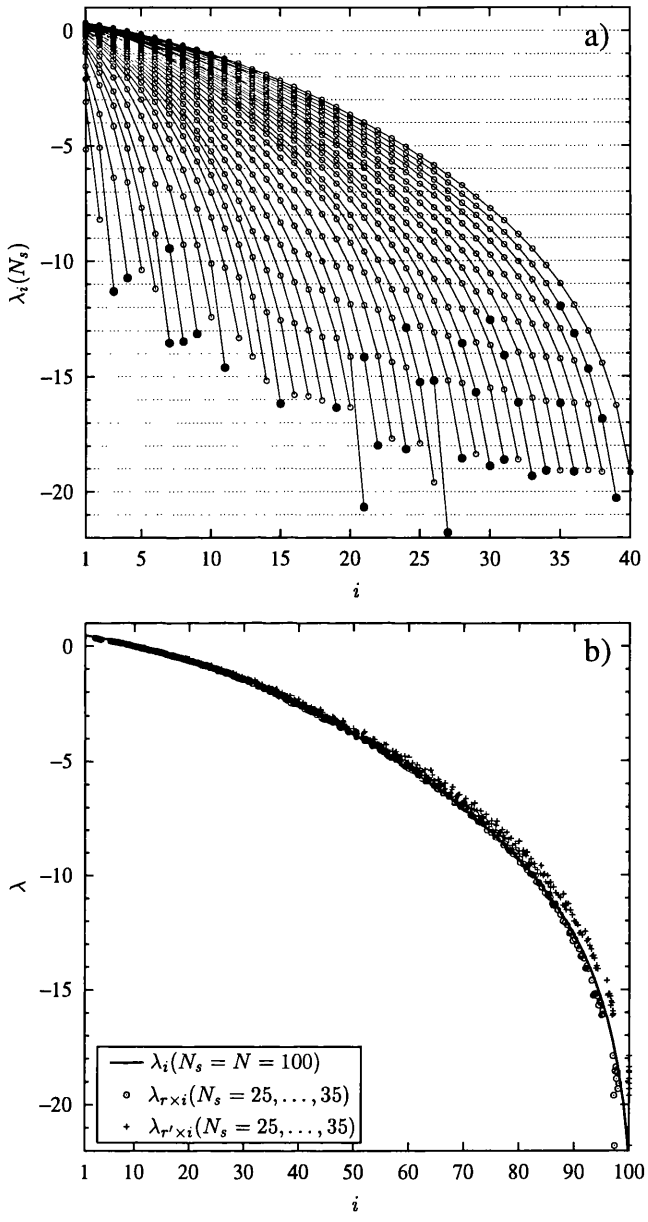


Figure 5.8: a) Interleaving of the sub-system LS in the chaotic neural network (5.20) with $k = 10$ and $g = 2$. b) Comparison between the conventional rescaling of the sub-system Lyapunov spectrum (crosses) and the new rescaling obtained in section 5.1 (circles), the whole LS is depicted by the solid line.

mation: while each entry of the CML Jacobian (5.3) depends on a single site, each entry of the CNN Jacobian depends on a neighbourhood of sites:

$$J_{ij}(n) = \frac{g\, C_{ij}}{\cosh^2\left[\sum_{l=i-k}^{i+k} C_{il} x_l^n\right]}. \tag{5.21}$$

The CNN Jacobian (5.21) inherits the zeros of the coupling matrix C_{ij}, i.e. $J_{ij}(n) = 0$ if $(i-j) \pmod N > k$. Another difference between the CNN that we will consider and the diffusive CML discussed before is that the CNN involves coupling with a larger neighbourhood than just the left and right nearest neighbours.

Let us now analyze the interleaving and rescaling for a CNN with a large k. In figure 5.8 we show the interleaving and rescaling with $k = 10$ and $g = 2$. As we can see, the interleaving is quite good with the exception of a few small Lyapunov exponents. In figure 5.8.b we plot the rescaled sub-system LS for several sub-system sizes using both rescalings (circles: new rescaling and crosses: conventional rescaling) along with the whole LS (solid line). Clearly the new rescaling gives a better estimate of the whole LS. Similar results were obtained for other values of the parameters k and g.

5.4.2 Two-dimensional logistic lattice

The interleaving and rescaling properties of the sub-system LS were obtained in section 5.2 for a one-dimensional array of coupled maps. Here we put to the test the interleaving and rescaling for a two-dimensional CML. Let us take a two-dimensional square lattice of size $L \times L$. The local dynamics x_{ij}^n at each node (i, j) and any time n is governed by the fully chaotic logistic map

$$f_{ij}(x) = f(x) = 4x\,(1 - x).$$

As in the one-dimensional CML the local dynamics is applied first

$$y_{ij}^n(x) = f(x_{ij}^n),$$

and then the coupling dynamics

$$x_{ij}^{n+1} = (1 - \varepsilon)\, y_{ij}^n + \varepsilon\, \overline{y}_{\mathcal{N}_{ij}}^n,$$

where $\overline{y}^n_{\mathcal{N}_{ij}}$ is the average of the y^n_{ij} in the neighbourhood $\mathcal{N}_{ij}$ of site (i,j). The neighbourhood $\mathcal{N}_{ij}$ is taken to be the eight adjacent sites to (i,j) with periodic boundary conditions.

The Jacobian $J(n)$ at time n for this two-dimensional lattice is defined through its elements:

$$J_{kl}(n) = \frac{\partial x^{n+1}_{\sigma_k}}{\partial x^n_{\sigma_l}},$$

where the indices σ_k and σ_l refer to the position in the actual lattice of the chosen kth and lth state variables of the system. If one just wants to compute eigenvalues of the whole Jacobian, the order in which the state variables are taken is not relevant. However we are interested in extracting sub-Jacobian matrices from the whole system and thus the ordering choice of the state variables does matter. There are $L^2!$ different ways to choose the ordering, but the simplest way consists of taking the site (1,1) as the first state variable and then proceeding horizontally to the right until the end of the lattice is reached and then proceeding to the bottom of the lattice by rows:

1	2	3	...	$L-1$	L
$L+1$	$L+2$	$L+3$	...	$2L-1$	$2L$
$2L+1$	$2L+2$	$2L+3$	...	$3L-1$	$3L$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$

,

that is $\sigma_k = (k - \lfloor k/L \rfloor, \lfloor k/L \rfloor)$ where $\lfloor z \rfloor$ denotes the largest integer smaller than or equal to z. From now on this kind of ordering will be called *horizontal wraparound*. There is obviously a vertical counterpart where the order is taken by columns. The problem with this type of ordering is that it does not build up the Jacobian in a natural way. The propagation of a perturbation typically grows equally in both of the two spatial dimensions (in particular for our choice of coupling since all the neighbours contribute with the same weight $\varepsilon/8$). In contrast, with horizontal or vertical wraparound one has to wait until a complete wrap is taken to fall again near the perturbed area. A more natural approach might thus be to attempt to mimic the spatial growth of perturbations by taking

an ordering that fills up a two-dimensional area from the centre outwards. For that purpose, we use the following ordering technique:

1	2	5	10	$\cdots$
4	3	6	11	$\cdots$
9	8	7	12	$\cdots$
16	15	14	13	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

We call this *square wraparound.*

In figure 5.9 we show the non-rescaled sub-system LS for the two wraparound methods, a) square and b) horizontal, and we plot with solid circles the Lyapunov exponents that fail to interleave. Observe that interleaving failure occurs for only a very few Lyapunov exponents. After a careful examination of these Lyapunov exponents one notices that they are very close to interleaving, suggesting that the failure is due to numerical error in the computation of the exponents (and in particular poor convergence). Therefore, we shall consider a Lyapunov exponent to be interleaved if it falls in the interval defined by the inequality (5.14) with an error δ:

$$\lambda_i(N_s+1) - \delta\Lambda \leq \lambda_i(N_s) \leq \lambda_{i+1}(N_s+1) + \delta\Lambda, \tag{5.22}$$

where $\Lambda = \lambda_{i+1}(N_s+1) - \lambda_i(N_s+1)$. From now on we redefine δ such that the errors are given in percentages. Using such a definition, if one allows a small error of 2.5% —$\delta = 0.025$ in (5.22)— for the Lyapunov exponents in figure 5.9, one obtains perfect interleaving for the whole spectrum.

The interleaving seen in figure 5.9 suggests that the ordering choice for the Jacobian entries does not play an important role in this phenomenon. However, as can be seen in figure 5.10, where we plot the rescaled LS for both wraparound methods along with the whole LS, the choice of ordering method is crucial in obtaining good rescaling behaviour. Square wraparound (figure 5.10.a) yields immediate convergence towards the whole LS: even for a very small sub-system size

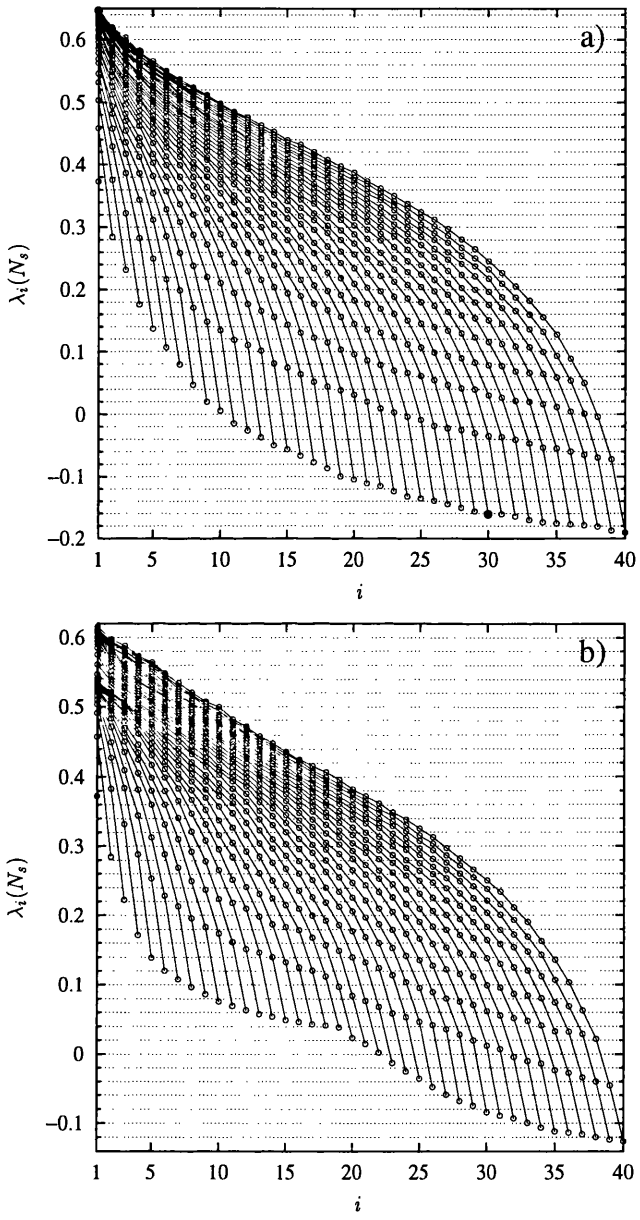


Figure 5.9: Sub-system Lyapunov spectra for the two-dimensional 20×20 coupled logistic lattice for sub-system sizes 1 to 40 (left to right) for $\varepsilon = 0.45$ and for the two wraparound methods for building up the Jacobian: a) square wraparound and b) horizontal wraparound. The filled circles represent the Lyapunov exponents where interleaving fails.

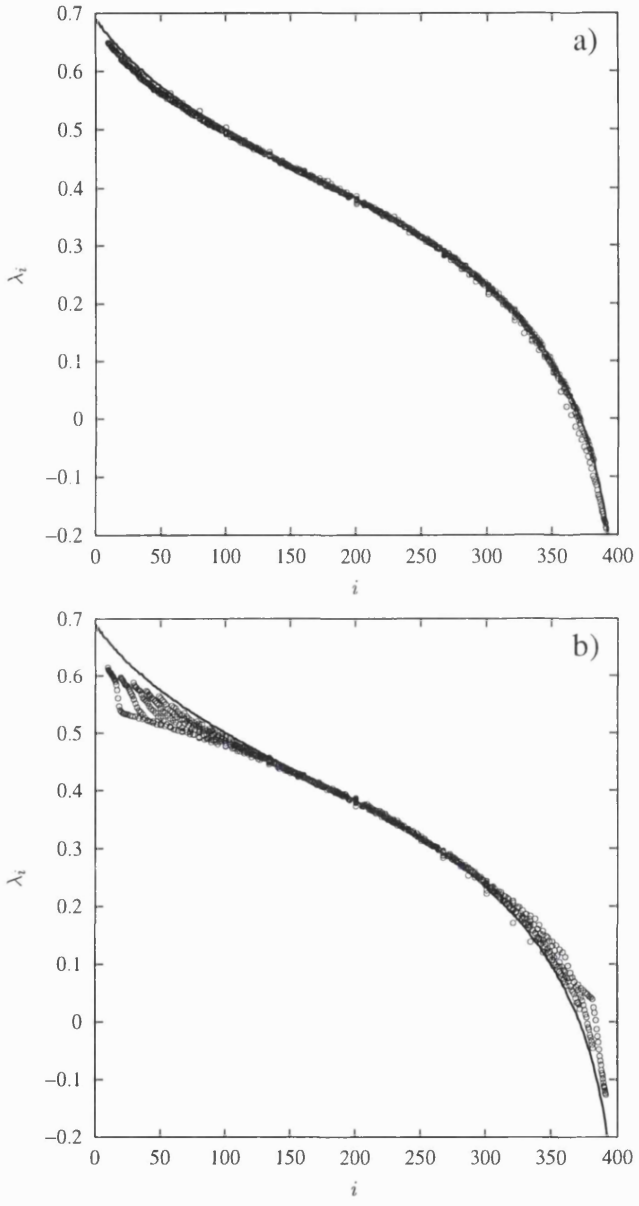


Figure 5.10: Rescaled sub-system LS for the two-dimensional coupled logistic lattice (same parameters as in figure 5.9) using a) square and b) horizontal wraparound methods.

the rescaled LS is almost exactly superimposed on top of the whole LS. On the other hand, horizontal wraparound (figure 5.10.b) gives a rescaled sub-system LS that seems to converge to a different curve for sub-system sizes $N_s = 1, \ldots, 20$ (aligned points in the lower part of the curve for the first Lyapunov exponents). For sub-system sizes larger than 20 the rescaled LS starts a new convergence towards something closer to the whole LS. The explanation for this phenomenon is quite simple. The Jacobian for the horizontal wraparound consists of a main diagonal of non-zero elements coming from the neighbours in the same row of the square lattice, however, the neighbours in the row above and below give rise to two sub-diagonals of non-zero elements. The sub-diagonals start when a whole wraparound has been completed, that is when $N_s = L$ where L is the side length of the square lattice. Thus for sub-system sizes $N_s < L$ the sub-Jacobian only extracts the main diagonal elements and does not capture the two sub-diagonals with vital information about the neighbouring sites in the rows above and below. When $N_s \geq L$ the sub-Jacobian starts capturing these forgotten neighbours and the rescaled sub-system LS now begins to converge to the desired LS. For the example in figure 5.10.b this behaviour starts at $N_s = L = 20$. This effect of horizontal wraparound is reflected when one tries to extract information from the sub-system LS. As an example, we depict in figure 5.11 an estimate of the largest Lyapunov exponent by extrapolating the whole LS from its rescaled version as the sub-system size increases. The results are depicted with circles for square wraparound and with crosses for horizontal wraparound. The vertical solid line corresponds to the largest Lyapunov exponent from the whole LS. The estimate using horizontal wraparound seems to converge to a much smaller value than the desired one for sub-system sizes $N_s < L = 20$. When the sub-system size is increased further, horizontal wraparound performs better but still lacks the desired convergence. On the other hand, the square wraparound converges rapidly in a smooth way: this is because it was designed to build up the Jacobian entries in a more natural way. Therefore, although the interleaving for both wraparound methods is very good it is considerably more reliable to use the square wraparound for rescaling purposes of the sub-system LS.

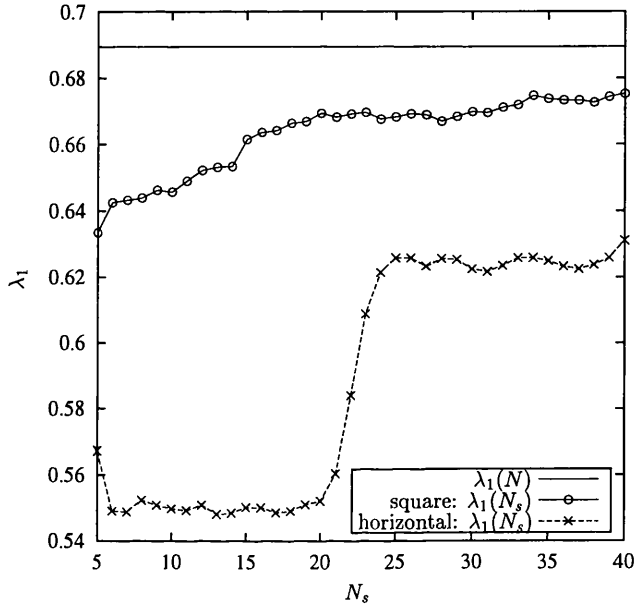


Figure 5.11: Estimation of the largest Lyapunov exponent as a function of the sub-system size N_s in a two-dimensional logistic lattice of size 20×20 and with $\varepsilon = 0.45$ using a linear fit for the rescaled sub-system LS. The circles correspond to building up the Jacobian by square wraparound whilst the crosses correspond to horizontal wraparound. The value of the largest Lyapunov exponent for the whole lattice is represented by the horizontal solid line.

5.4.3 Host-parasitoid system

We now consider a more general type of two dimensional lattice, namely the Host-Parasitoid lattice model (Hassell, Comins & May 1991; Comins, Hassell & May 1992; Wilson & Rand 1997). For this system the local dynamics is no longer one-dimensional but two-dimensional: hosts and parasites. The model evolves again in two phases. First there is at each site (i,j) a local dynamics given by

$$\begin{aligned} \mathcal{H}_{ij}^n &= bH_{ij}^n \mathrm{e}^{-aP_{ij}^n} \\ \mathcal{P}_{ij}^n &= cH_{ij}^n \left(1 - \mathrm{e}^{-aP_{ij}^n}\right) \end{aligned} \tag{5.23}$$

where H^n and P^n are respectively the population size of hosts and parasitoids at time n, a is the per capita parasitoid attack rate, b is the host reproductive rate and c is the conversion efficiency of parasitised hosts into female parasitoids in the next generation. The second phase involves dispersal into a neighbourhood $\mathcal{N}_{ij}$ of site (i,j), i.e. a fraction μ_h of hosts and μ_p of parasitoids disperse equally into the eight neighbouring sites:

$$\begin{aligned} H_{ij}^{n+1} &= (1-\mu_h)\,\mathcal{H}_{ij}^n + \mu_h \overline{\mathcal{H}}_{\mathcal{N}_{ij}}^n \\ P_{ij}^{n+1} &= (1-\mu_p)\,\mathcal{H}_{ij}^n + \mu_p \overline{\mathcal{P}}_{\mathcal{N}_{ij}}^n \end{aligned} \tag{5.24}$$

where $\overline{\mathcal{H}}_{\mathcal{N}_{ij}}^n$ and $\overline{\mathcal{P}}_{\mathcal{N}_{ij}}^n$ are, respectively, the average of the hosts and the parasitoids (after local dynamics (5.23)) in the neighbourhood $\mathcal{N}_{ij}$ of site (i,j). We take a square lattice $(i,j) \in [1,L]^2$ and periodic boundary conditions. The total size of the system is then $N = 2L^2$. Let us build up the whole Jacobian with host-parasite blocks of size 2×2:

$$J = \begin{pmatrix} \mathcal{J}_{\sigma_1}^{\sigma_1} & \mathcal{J}_{\sigma_2}^{\sigma_1} & \cdots & \mathcal{J}_{\sigma_{L^2}}^{\sigma_1} \\ \mathcal{J}_{\sigma_1}^{\sigma_2} & \mathcal{J}_{\sigma_2}^{\sigma_2} & \cdots & \mathcal{J}_{\sigma_{L^2}}^{\sigma_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{J}_{\sigma_1}^{\sigma_{L^2}} & \mathcal{J}_{\sigma_2}^{\sigma_{L^2}} & \cdots & \mathcal{J}_{\sigma_{L^2}}^{\sigma_{L^2}} \end{pmatrix},$$

where the host-parasite blocks $\mathcal{J}_{\sigma_j}^{\sigma_i}$ are given by

$$\mathcal{J}_{\sigma_j}^{\sigma_i} = \begin{pmatrix} \frac{\partial H_{\sigma_i}^{n+1}}{\partial H_{\sigma_j}^{n}} & \frac{\partial H_{\sigma_i}^{n+1}}{\partial P_{\sigma_j}^{n}} \\ \frac{\partial P_{\sigma_i}^{n+1}}{\partial H_{\sigma_j}^{n}} & \frac{\partial P_{\sigma_i}^{n+1}}{\partial P_{\sigma_j}^{n}} \end{pmatrix}.$$

The indices $\sigma_{1,\ldots,L^2}$ refer to the actual position in the two-dimensional lattice of a particular local population. As for the two-dimensional CML the ordering choice of the Jacobian entries plays an important role for the rescaling.

Given a reasonable lattice size ($L > 15$) and depending on the dispersal parameters μ_h and μ_p the evolution of model (5.24) is spatio-temporally chaotic (Rohani & Miramontes 1995). Here we choose $L = 20$, $a = 1$, $b = 2$, $c = 1$, $\mu_h = 0.2$ and $\mu_p = 0.6$. The full system is thus $N = 2L^2 = 800$ dimensional. We start the system with random initial conditions and discard a transient of 10^5 iterations before computing the sub-system LS. In figure 5.12we depict the interleaving of the sub-system LS for sub-system sizes $N_s = 1, \ldots, 40$ where we allow a 5% error in the interleaving —$\delta = 0.05$ in (5.22). Figures 5.12.a and 5.12.b correspond to square wraparound whilst figures 5.12.c and 5.12.d correspond to horizontal wraparound. As the figure shows, interleaving is quite good even for the upper region (see amplifications in figures 5.12.b and 5.12.d) where the density of Lyapunov exponents is very high and the intervals for interleaving are small and thus the margin for error in the inequality (5.22) is reduced. Square wraparound does better for large Lyapunov exponents (figure 5.12.b) whilst horizontal wraparound does better for small ones. However, overall both methods have approximately similar performance.

While the choice of wraparound method is not crucial for interleaving, figure 5.13.a shows that it leads to significant differences in rescaling behaviour. Note that the LS for the whole system $N_s = N = 800$ is not depicted since it would take an enormous amount of time to compute. In figure 5.13.a we depict the rescaled LS for sub-system sizes $N_s = 1, \ldots, 40$ for both wraparound methods (square wraparound with circles and horizontal wraparound with crosses). As for the two-dimensional lattice of coupled logistic maps, horizontal wraparound converges to

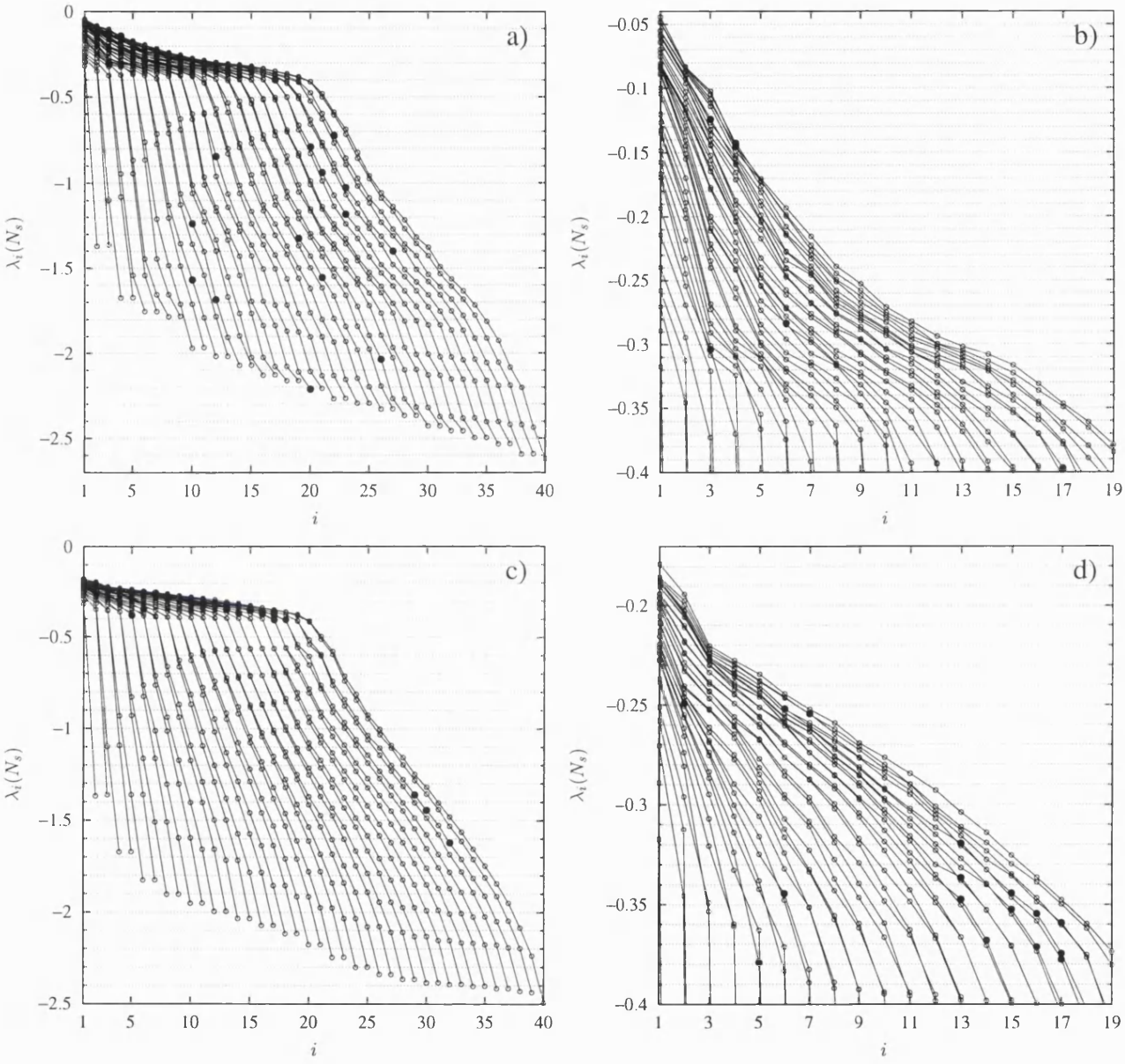


Figure 5.12: Interleaving of the sub-system LS for the host-parasite system in a two-dimensional lattice of size 20×20. The Jacobian was built using a)–b) square wraparound and c)–d) horizontal wraparound. Figures b) and d) correspond, respectively, to amplifications of figures a) and c) for the top half of the spectrum.

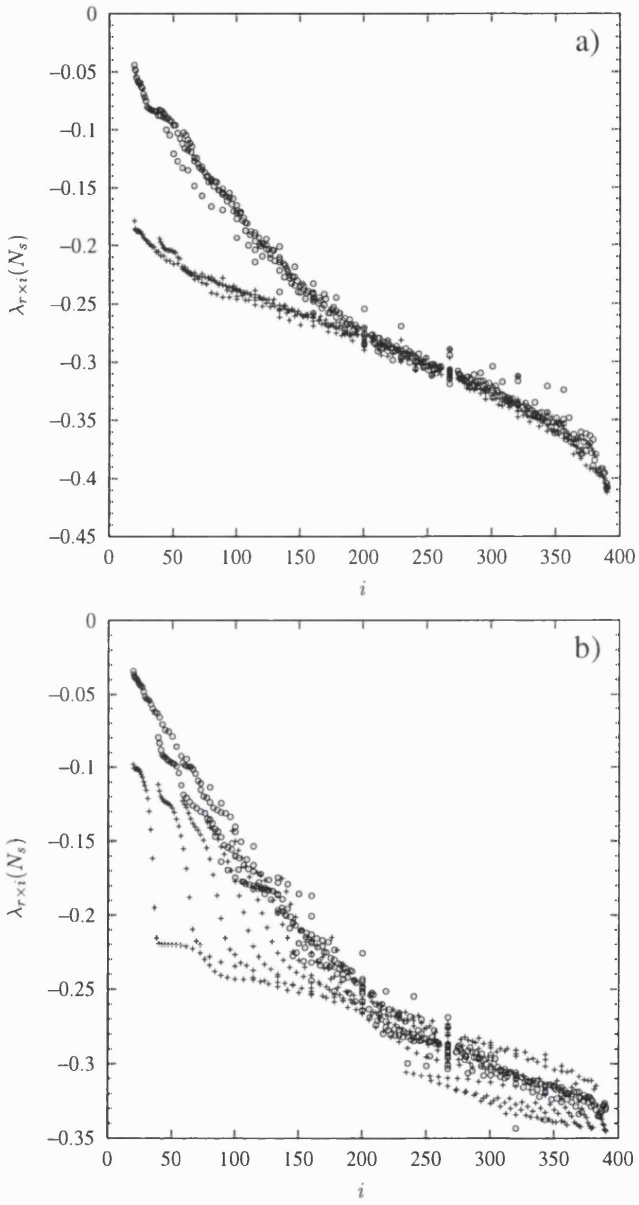


Figure 5.13: First half of the rescaled Lyapunov spectrum for a host-parasitoid system in a two-dimensional lattice for sub-system sizes $N_s = 1, \ldots, 40$. a) Using both the host and parasite variables and b) using only the hosts when building up the Jacobian. The circles (crosses) correspond to the square (horizontal) wraparound.

a different curve than does square wraparound. The reason is again that for the horizontal wraparound one has to wait until a complete wrap is finished until falling again into the neighbouring region. In this case, a horizontal wrap of the Jacobian is achieved when $N_s = 2L = 40$. Moreover, for $N_s = 2L = 40$ one is only including partial derivatives of hosts with respect to hosts and parasitoids. In order to include dependences of parasitoids with respect to hosts and parasitoids one should take a further wrap of the Jacobian, i.e. $N_s = 4L = 80$.

Therefore, the horizontal wraparound technique for sub-system sizes $N_s \leq 40$ does not pick up the dynamics of the neighbours situated in adjacent rows. This problem for horizontal wraparound becomes worse as the dimension of the local dynamics is increased. A partial solution to this problem is to build up the Jacobian by using just one of the local variables of the system. Particularly in the host-parasite system where the parasitoid dynamics is slaved to the host dynamics, one should be able to reproduce the LS from only the host variables. We then build up the Jacobian by taking only host variables using both wraparound methods. The results are shown in figure 5.13.b where again the circles correspond to square wraparound and the crosses to horizontal wraparound. We only plot the first half of the spectrum, the second half of the spectrum differs considerably for both methods (host-parasites variables and only host variables) since the small Lyapunov exponents are more sensitive to the loss of information contained in the parasite variables. On the other hand, the first half of the spectrum is quite similar independently of the choice of host-parasite or only host variables. As we can see in figure 5.13.b, horizontal wraparound seems to converge to a different curve than square wraparound for sub-system sizes $N_s < 20$ (see aligned crosses in the lower part of the spectrum). Since we are only taking the host population (20×20), when $N_s > 20$ the horizontal wraparound has finished a complete wrap and it starts to pick up the neighbours in adjacent rows and thus the rescaled spectrum begins to converge closer to the square wraparound.

5.5 Discussion

When studying high dimensional extended dynamical systems in a spatio-temporal chaotic regime it is possible to rescale the sub-system Lyapunov spectrum to obtain the original Lyapunov spectrum. In this thermodynamic limit, a sub-system of comparatively small size N_s contains a sufficient amount of information to reconstruct the Lyapunov spectrum of the whole system. Usually, when coupling different sub-systems in a lattice one chooses a coupling with a finite neighbourhood (localized coupling) or at least with decreasing effect for further away neighbours. In the context of discrete spatio-temporal systems, this restriction on the choice of coupling causes the Jacobian of the dynamics to be a banded (or quasi-banded) matrix. In the limit of only nearest neighbours interaction in a one-dimensional lattice, the Jacobian is a tridiagonal matrix. If one considers the homogeneous evolution under this dynamical system, the Lyapunov spectrum of sub-Jacobian matrices will inherit the rescaling and interleaving properties described in section 5.1. The evidence presented in this chapter shows that the new rescaling method of the sub-system Lyapunov spectrum gives a much better fit than the conventional rescaling N/N_s for one-dimensional lattices.

We have also observed interleaving of the Lyapunov spectra for consecutive sub-system sizes. We showed that for two-dimensional lattices the rescaling and interleaving are still valid. However, the choice of variables used to build up the sub-Jacobian matrices appears to be crucial to achieve good rescaling properties. In particular one has to choose an ordering method of the system variables that mimics the propagation of information in the particular lattice topology of the system. In two dimensions we showed that choosing the system variables in 'concentric' sub-squares gave a much better rescaled Lyapunov spectrum than by choosing them in a row or column-wise fashion. Generalizing this idea to higher-dimensional lattices one should take the system variables by filling up 'concentric' hyper-cubes.

Another point to take into account when choosing the system variables in high-dimensional lattices is the anisotropy of the coupling. The two-dimensional

systems studied here have an equal relative contribution from all the neighbouring directions (isotropic coupling). It is possible to choose the coupling in order to give more weight to one of the directions (vertical or horizontal) and thus the propagation of information to be faster in that direction. Therefore, instead of building the system variables by 'concentric' squares it should be more natural to take rectangles, the ratio of the rectangle sides being related to the ratio of velocity propagation of disturbances in both directions.

For a continuous spatio-temporal system a similar reconstruction may be used by sampling in a grid of a sub-system at regular time intervals and by reconstructing the Jacobian from time series in the usual manner. The same procedure can be applied for a discrete spatio-temporal system where the dynamics is not explicitly given and the only available dynamic information comes from time series taken at several spatial locations. We expect that rescaling and interleaving should still be observed in these cases. This aspect is investigated in the next chapter.

Chapter 6

Estimation of intensive quantities in spatio-temporal systems from time-series

As we saw in chapter 2, in low-dimensional systems an ample framework has been developed for the characterisation of the corresponding attractor, such as the fractal dimension (by calculating the correlation dimension) and the estimation of the Lyapunov spectrum. Spatio-temporal systems, on the other hand, are typically high-dimensional and moreover the dimension of the chaotic attractor grows with the system size. Since the methods mentioned in chapter 2 require the available data to increase exponentially with the dimension of the attractor (Kantz & Olbrich 1997) it is not clear how they can be utilised in the context of spatially extended systems. One approach, as suggested by (Grassberger 1989; Mayer-Kress & Kaneko 1989), is to use the correlation dimension algorithm to calculate dimension densities and use this to distinguish between spatio-temporal chaos and noise. This approach has been further studied by (Bauer, Heng & Martienssen 1993; Tsimring 1993). However, (Puccioni, Torcini, Politi & G.D'Alessandro 1991) has questioned the ability of this method to distinguish between chaos and noise. Also remember that, as we showed in chapter 4, even a small truncated lattice can model, exponentially close, the dynamics of a potentially ∞-lattice. Thus at least

time-delay embedding techniques are bound to fail in spatio-temporal systems.

In this chapter another approach to characterise a spatially extended system is adopted, namely we use time-series of sub-systems to estimate the Lyapunov spectrum (LS) of the full system. From this we are able to estimate related quantities such as the Lyapunov dimension density and KS entropy density. The approach was first suggested by (Kaneko 1989). There are, however, some shortcomings to this method as it was described there. Firstly the method relies on a local approximation of the tangent map, which involves finding close neighbours in a reconstructed state space. Since we have to work in dimensions at least up to 20 it is not clear how this may be achievable with limited data. Furthermore, Kaneko was not able to estimate the negative part of the LS accurately enough to be able to calculate the Lyapunov dimension density, thus only the KS entropy density (which only relies on the positive part of the LS) was calculated. We address these shortcomings and suggest ways in which they may be overcome.

6.1 Lyapunov spectrum and related densities

As already mentioned in chapter 5, in the case of spatially extended systems, e.g. in the case of a turbulent flow, is has been pointed out that sub-systems should be weakly correlated thus the Lyapunov spectrum of the whole system should be approximately equal to the union of exponents from each non-interacting sub-system (Ruelle 1982). This implies that the LS should be intensive in that $\lambda_i = F(i/V)$ is a function of an intensive index i/V, where V is the volume of the system. To recall, as a consequence the Lyapunov dimension and KS entropy are extensive quantities that grow linearly with increasing system size V. Thus we can define the corresponding Lyapunov dimension density ρ_d as

$$\rho_d(V) = \lim_{V\to\infty} d_L/V, \tag{6.1}$$

and entropy density ρ_h as

$$\rho_h(V) = \lim_{V\to\infty} h/V. \tag{6.2}$$

The densities defined in eq. (6.1) and eq. (6.2) are thus intensive quantities that do not vary with system size. These quantities were first found experimentally in (Manneville 1985).

We saw in chapter 5 that for spatially extended system, and in particular by analysing coupled map lattices, the sub-system LS converges rapidly towards the LS of the whole system for increasing sub-system size (also see (Kaneko 1989; Parekh, Kumar & Kulkarni 1998)). Instead of using all the N variables of the system to build the Jacobian we only took a subset N_s of these variables (from N_s neighbouring sites) and built the Jacobian for this N_s-dimensional sub-system without changing the underlying dynamics for the whole original system. When calculating the sub-system LS for increasing N_s, we found that Lyapunov exponents for two consecutive sizes were interleaved and that the sub-system LS converges to the LS for the whole system when using a suitable rescaling.

The convergence to a limiting LS suggested that the Lyapunov dimension and KS entropy scale with sub-system size and this gave us the opportunity to estimate ρ_d and ρ_h as

$$\rho_d(N) = \lim_{N_s \to N} d_L(N_s)/N_s, \tag{6.3}$$

and

$$\rho_h(N) = \lim_{N_s \to N} h(N_s)/N_s, \tag{6.4}$$

respectively for large enough N_s.

Recall that there are two ways of doing this from sub-system information. Firstly one could use eq. (6.3) and eq. (6.4) and plot $\rho_d(N_s)$ and $\rho_h(N_s)$ for increasing N_s. In the thermodynamic limit (infinite number of lattice sites) these densities are intensive quantities (they do not depend on sub-system size). One can then estimate the Lyapunov dimension and KS entropy by multiplying by N. The second method, as suggested in chapter 5, is to take the sub-system LS, rescale it, and then extrapolate a curve through it to obtain an approximation of the whole LS. A simple choice of extrapolation method is to use piece-wise linear fits.

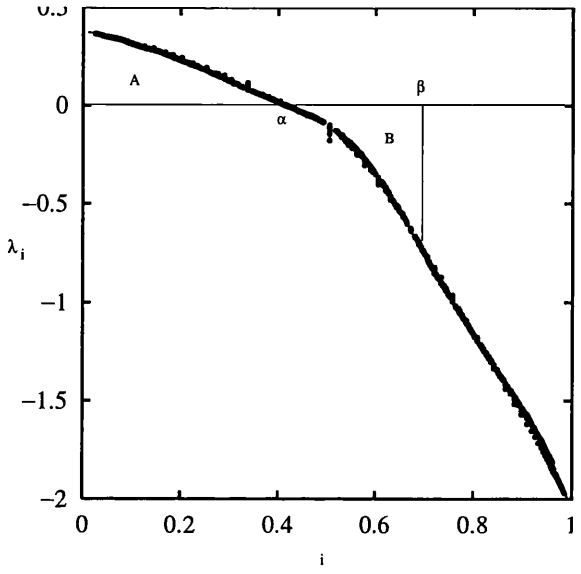


Figure 6.1: Convergence to the limit curve using the new rescaling. The real LS is super-imposed as a continuous curve. The area A equals the KS entropy density and the value of β that makes area A equal to area B is an estimate of the Lypunov dimension density.

Now consider the case where only one or more time-series sampled from the system are available. As an example, consider the logistic coupled map lattice with $N = 100$ and $\epsilon = 0.45$ as before.

In figure 6.1 we have plotted the rescaled sub-system LS for sub-system sizes $N_s = 1, \ldots, 40$ (this corresponds to figure 5.3.b but for a slighly different index axis range. The reason for this change will become apparent below). Is it possible to get the same information as in the examples in chapter 5 without knowing the dynamical equations governing the system? Ideally we would like to be able to calculate Lyapunov exponents for sub-systems directly from time-series with some accuracy and one would hope that they converge to some limit curve for increasing sub-system sizes. It is clear that in this case we would not know the size of the full system, N. Hence we would not be able to use the rescaling in eq. (5.16) to produce the limit curve that would approximate the whole LS. In eq. (5.16) the rescaling confines the index axis to the range $[1, N]$. Instead choose the index axis

to lie in the range $[0, 1]$ by dividing eq. (5.16) by N to get the rescaling

$$r' = \frac{(N+1)}{N(N_s+1)}.$$

However, since typically N is big, we could assume $(N+1)/N \approx 1$ and therefore it seems plausible to use the approximated rescaling

$$r = \frac{1}{N_s+1}, \tag{6.5}$$

when the system size N is unknown. Further, not knowing N would mean that we will only be able to estimate densities (and not the extensive quantities Lyapunov dimension and LS entropy for the whole system). We could choose to plot ρ_d and ρ_h for increasing N_s using eq. (6.3) and eq. (6.4) and hope to observe a convergence for N_s large enough. Again, another method would be to use the polygon method introduced in chapter 5 but modified as follows. The polygon can be described as an approximation of the function $F(i/N)$, i.e. the whole LS. In the thermodynamic limit we can define ρ_h as

$$\rho_h = \lim_{N\to\infty} \int_{i=1}^{\alpha} F(i/N), \tag{6.6}$$

such that $F(\alpha) = 0$. Note that the value of this integral corresponds to the area A in figure 6.1. Thus to get an approximation of ρ_h, $\rho_h(N_s)$, from sub-system LS estimated from time-series, we calculate the area A as given by the polygon. Similarly we can find ρ_d by solving for β the equation

$$\int_{i=0}^{\beta} F(i/N) = 0. \tag{6.7}$$

This is equivalent to finding β such that areas A and B in figure 6.1 satisfy $A = B$. The corresponding β is then an estimate of ρ_d. We will refer to this way of computing ρ_h and ρ_d, from areas, as the polygon method and call the estimation of ρ_d and ρ_h from eqs. (6.3) and (6.4) as the direct method.

6.2 Sub-system Lyapunov spectrum from time-series

In the previous section we discussed a possible way of how to utilise sub-system LS from time-series. Before presenting the results we need to discuss how to find the sub-system Lyapunov exponents themselves.

Usually the LS is calculated from an observed time series using the method of time delays as described in chapter 2 and 3. However, using the method of delays to obtain good estimates of the LS is known to be difficult in practice, even for low-dimensional systems. Spatially extended systems, which we are interested in here, are typically high-dimensional and thus the extraction of the LS is even more difficult. Takens theorem (Takens 1981) require us to work in dimensions larger than $2M$ where M is the dimension of the attractor of the system. Since the method involves finding neighbouring points in high dimensions this means that we require an enormous amount of data to have any hope of decently reconstructing the dynamics.

Here we will consider the case where N_s time series are observed at N_s different neighbouring spatial locations simultaneously. Spatial delay vectors are built up using data from neighbouring sites as entries. Intuitively one would expect that the added information included by utilising several sites at once would increase our ability to estimate the LS.

Consider spatial delay vectors of size d. We can treat these as sub-systems of size $N_s = d$ and from its corresponding reconstruction we hope to estimate the densities. It should be emphasized that when using time-series we assume the size of the original system to be unknown. Therefore, only densities may be approximated since they do not depend on the system size. Their extensive counterparts d_L and h can obviously not be estimated without knowing N.

In chapter 2 we described how the Jacobian may be estimated by a local linear fit. It is also possible to include higher order terms in the neighbourhood matrix e.g. by considering second order terms. The hope is that by including these, the estimate of the Jacobian will improve, hence giving us a better approximation of

the Lyapunov exponents. As a special case let us choose the spatial delay vectors as

$$\mathbf{x}^n = (x_1, x_2, \ldots, x_d, x_1^2, x_2^2, \ldots, x_d^2), \tag{6.8}$$

where all entries are evaluated at time n and where we deliberately excluded cross terms for simplicity. The neighbours and neighbourhood matrices are now found using these vectors and the Jacobian is given by the upper-left $d \times d$ sub-matrix of J_n (see chapter 2.1.4). The problem with including higher order terms is that the number of neighbours needed to fit the Jacobian grows rapidly for increasing embedding dimension d and the data requirements soon get too large. In fact, the minimal number of neighbours required to fit the Jacobian (including cross-terms) with an embedding dimension d is $\prod_{k=1}^{d}[\frac{d+k}{k}] - 1$.

A further improvement in the calculation of the sub-system LS can be made by careful consideration of the nature of the Jacobian and of what happens at the edges of the sub-system. In a lattice with nearest neighbour coupling, a perturbation will travel with only finite velocity in the lattice (this is not the case for globally coupled maps). Thus the entries in the Jacobian should be non-zero only along the diagonal and in its vicinity. (Note that this is in the case of a system with nearest neighbour coupling. If the coupling is over more than one site, more near-diagonals will be non-zero. In practice we treat the input at the edges of the sub-system as noise affecting the estimation of the Jacobian in that the whole of the first and last row will be non-zero. This 'noisy' Jacobian leads to a less accurate estimation of the sub-system LS. We thus discard the outer layer of the Jacobian to eliminate this noise and get better estimates for the sub-system LS, i.e. disregard the information from the nodes at the edges. This leads to an estimate of $(N_s - 2)$ Lyapunov exponents when considering a sub-system of size N_s for nearest neighbour coupling.

6.3 Numerical results for the estimation of the LS from time-series

In this section we present the numerical results when the LS is determined from time-series. Let us first compare the results obtained before for the known dynamics of the coupled logistic lattice with their reconstructed counterpart. Thuse we are again using the logistic map lattice with local dynamics $f(x) = 4x(1-x)$, coupling $\epsilon = 0.45$ and lattice size $N = 100$. We sample 20 neighbouring sites after discarding transient behaviour and keep 11000 time steps. Our sample set is thus 20×11000 long.

Using spatial delay reconstructions $d = 1, \ldots, 20$ we calculated sub-system LS using either linear fits (L-fit) or linear-quadratic fit (LQ-fit) as in eq. (6.8) and in each case with and without discarding the outer layer of the Jacobians. In all these cases we used data from 10000 time steps in the fitting, 1000 time steps to assure convergence of the sub-system LS and $d + 20$ nearest neighbours in the L-fits and $2d + 20$ nearest neighbours in the LQ-fits.

In figure 6.2 we depict the interleaving properties for the different ways of calculating the Jacobians. The interleaving works better for the LQ-fit (Figs. 6.2.c and 6.2.d) than for the L-fit (Figs. 6.2.a and 6.2.b) and even better when using discarded outer layers in the Jacobians (figure 6.2.d). This is especially clear for the most negative exponents.

In figure 6.3 we have plotted the corresponding convergence to the limit curve for the 4 possible methods using the rescaling in eq. (6.5). In all the cases the upper part of the spectrum is well approximated when d is large enough. It is clear though, that using discarded Jacobians give a much faster convergence for the upper part of the sub-system LS (compare Figs. 6.3.b and 6.3.d to Figs. 6.3.a and 6.3.c). Using the LQ-fit with discarded Jacobians leads to a very good approximation to the whole LS even for small d (figure 6.3.d).

Now we use these approximations of the limit curve to extract estimates of the largest Lyapunov exponent, Lyapunov dimension density and KS entropy density using both the direct method and the polygon method discussed earlier.

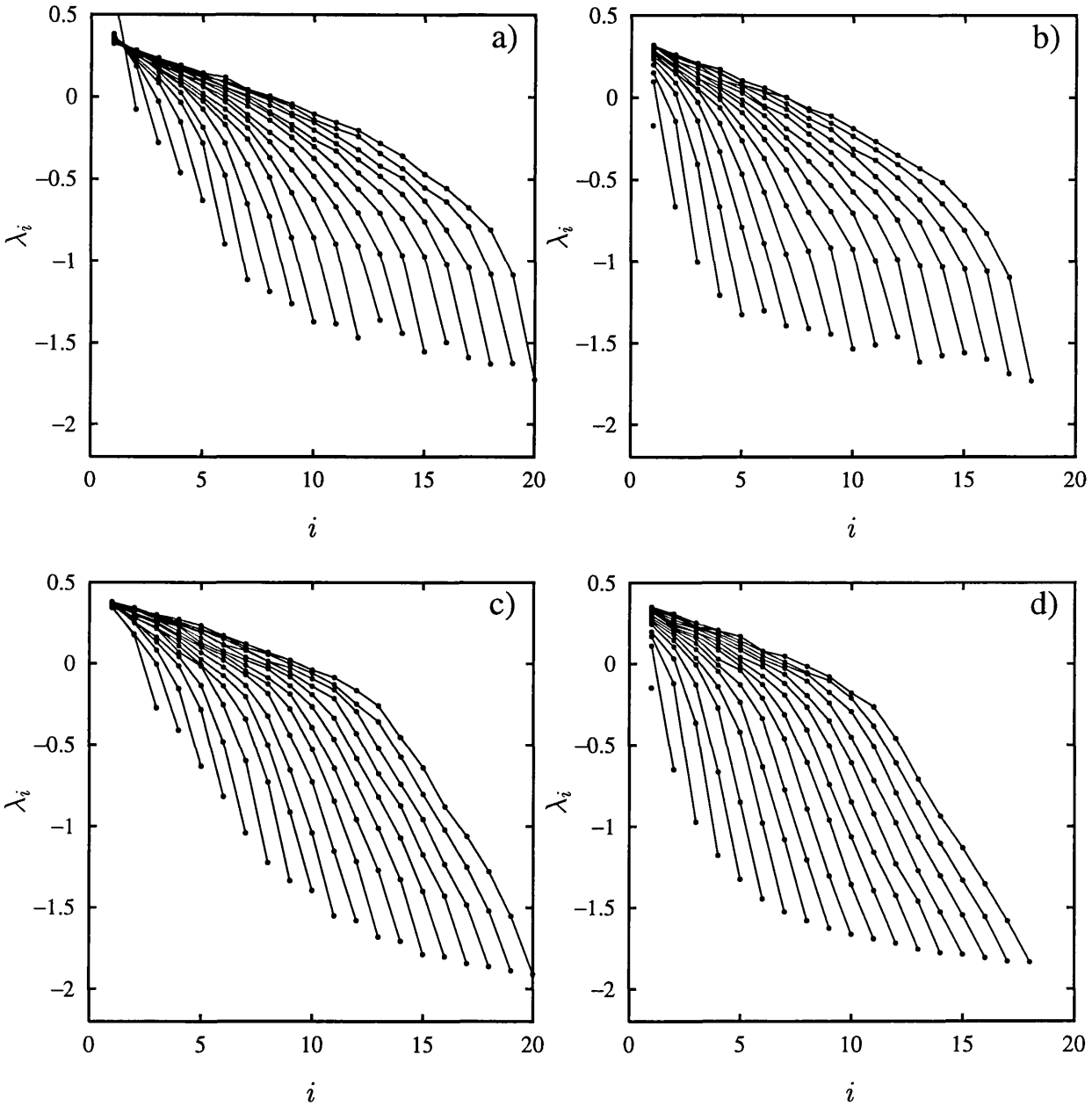


Figure 6.2: Interleaving properties using spatial delay reconstructions with dimensions $d = 1, \ldots, 20$ and using different ways of estimating the Jacobians. a) L-fit without discarding the edges of the Jacobians; b) LQ-fit without discarding; c) L-fit with discarding; d) LQ-fit with discarding.

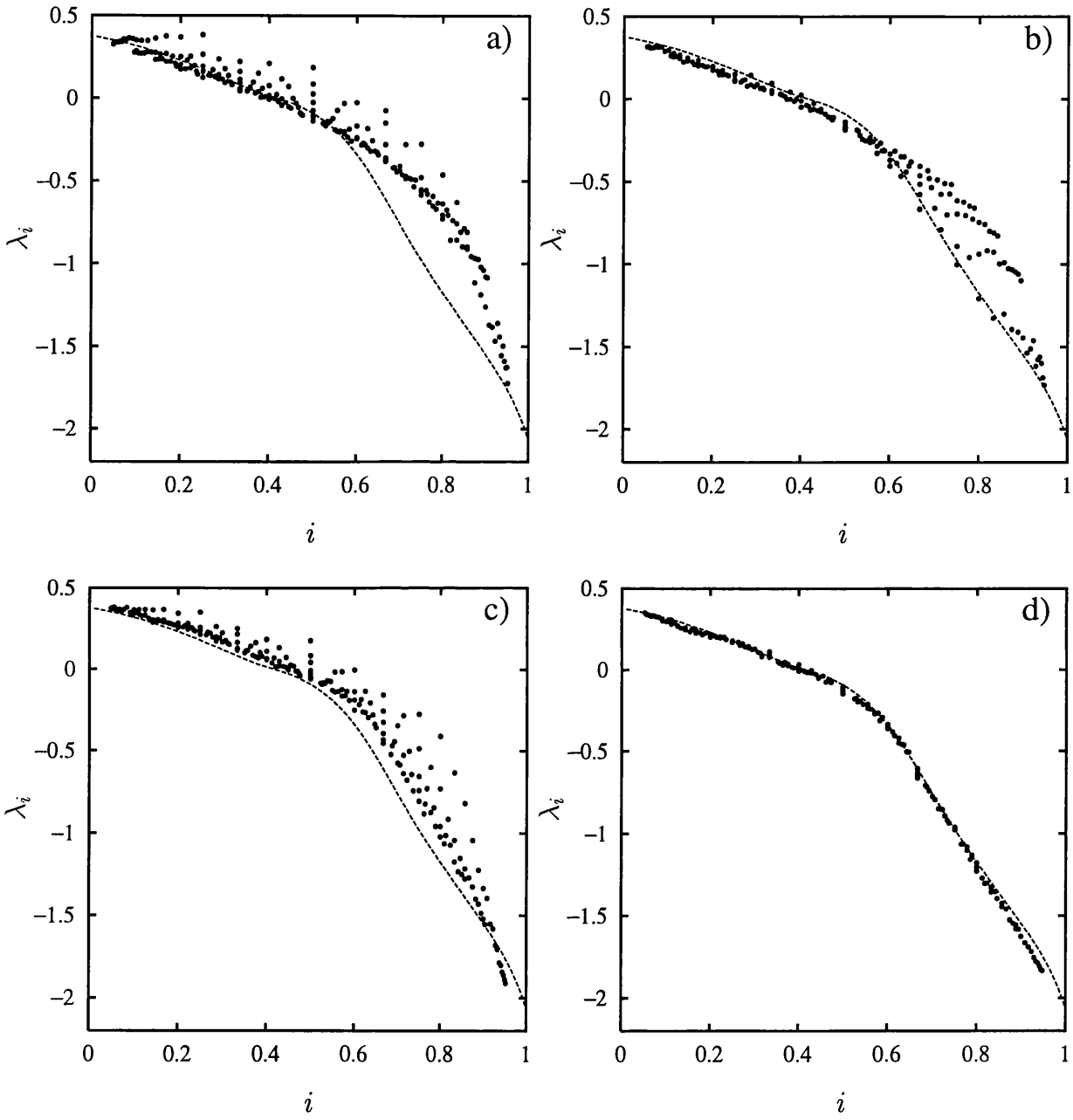


Figure 6.3: Estimated sub-system LS using spatial delay reconstructions with dimensions $d = 1, \ldots, 20$ and using the 4 different ways of estimating the Jacobians; a) L-fit without discarding the edges of the Jacobians; b) LQ-fit without discarding; c) L-fit with discarding; d) LQ-fit with discarding.

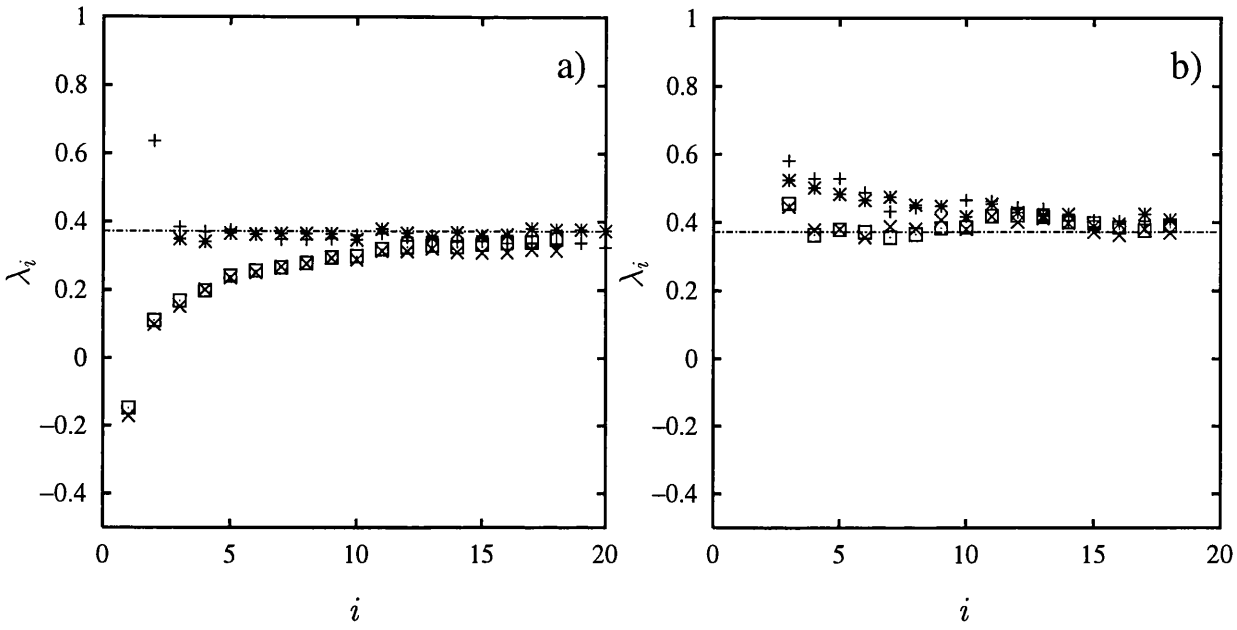


Figure 6.4: Estimated largest Lyapunov exponent using a) the direct method; b) the polygon method. ((+) L-fit without discarded edges, (×) L-fit with discarded edges, (∗) LQ-fit without discarded edges, (□) LQ-fit with discarded edges).

In figure 6.4 we have plotted estimates of the largest Lyapunov exponent using a) the direct method and b) the polygon method. We see that the largest Lyapunov exponent is well approximated using both approaches. Either by using the direct method with non-discarded Jacobians (+ and ∗ in figure 6.4.a) or by using the polygon method and discarded Jacobians (× and □ in figure 6.4.b). Both L-fits and LQ-fits work well in these cases. The success of the direct method in low dimensions can be explained by considering the fits that we are actually performing when approximating the Jacobians. When d is low we are able to find true nearest neighbours, and are indeed doing a local approximation of the dynamics, hence finding Jacobians in some reconstructed space. When d is higher this is not longer possible. However, if the map are approximated *globally* by the chosen basis, the Jacobians are still well approximated in high dimensions.

The fact that the largest Lyapunov exponent can be approximated well in low dimensions by local fits suggest that we can use these estimates as an upper bound for the estimates of the whole LS from the global fit.

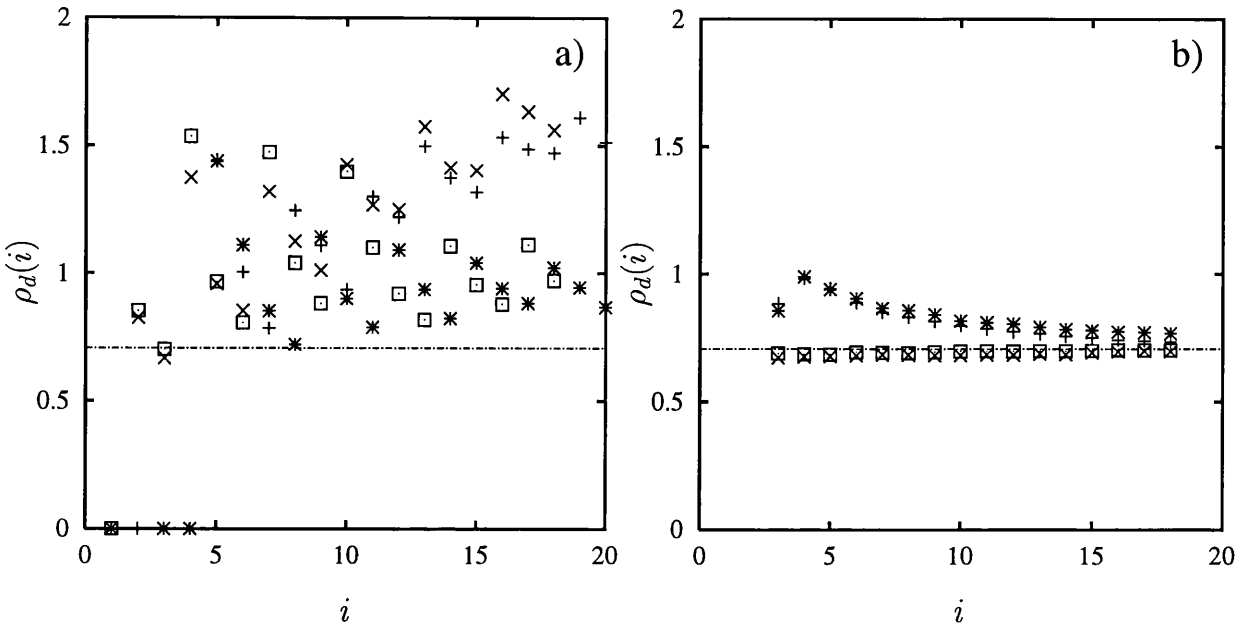


Figure 6.5: Estimated Lyapunov dimension density using a) the direct method; b) the polygon method. ((+) L-fit without discarded edges, (×) L-fit with discarded edges, (∗) LQ-fit without discarded edges, (□) LQ-fit with discarded edges).

Similarly we compare the efficiency of different approaches in the estimation of the Lyapunov dimension density as depicted in figure 6.5. We see that only when using the polygon method as in figure 6.5.b are we able to get a good estimate of ρ_d and the estimate is improved further by discarding the edges of the Jacobians. The failure of the direct method to estimate ρ_d comes from the second term in eq. (2.9) which are used in eq. (6.3). A small error in the denominator in eq. (2.9) amplify the error in the estimate of ρ_d. This implies that the polygon method is preferable to the direct method when estimating ρ_d from time-series.

In the case of KS entropy density (see figure 6.6), the polygon method works best, but in this case the direct method also gives good estimates. The discarded edges in the Jacobians improve the estimate for both methods.

The results suggest that when trying to approximate the whole LS we should consider the fitting as global, and therefore it becomes important to disregard the outer layer of the approximated Jacobian and use the polygon method.

Typically, the underlying system will not be contained in the chosen basis. To

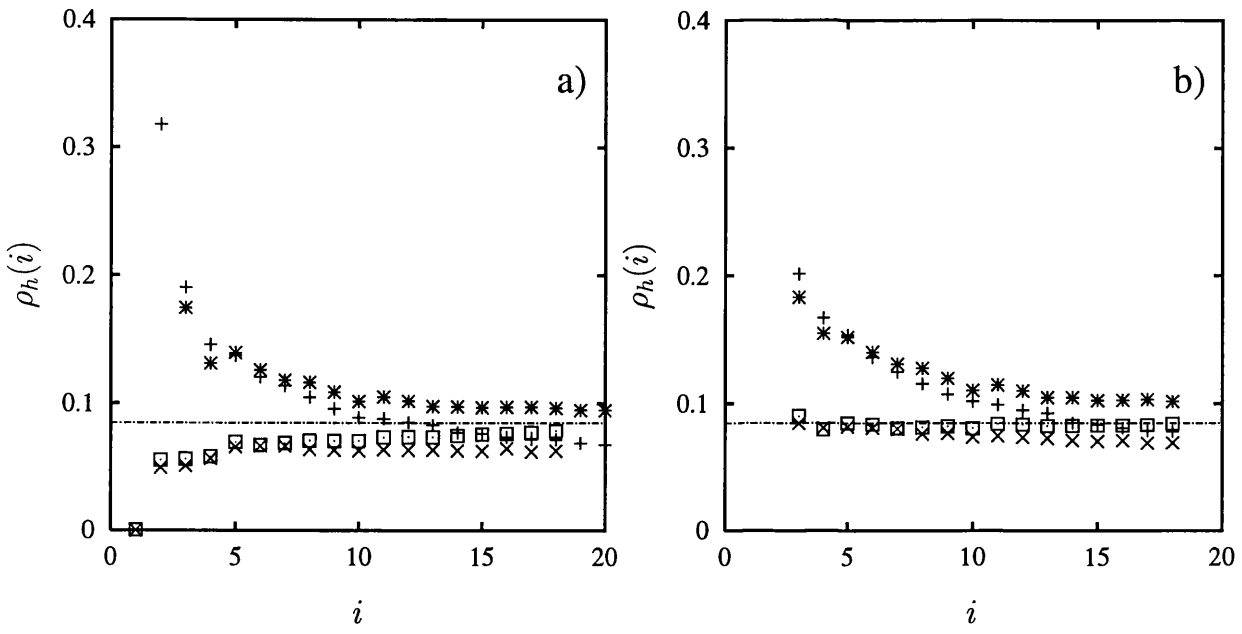


Figure 6.6: Estimated KS entropy density using a) the direct method; b) the polygon method. ((+) L-fit without discarded edges, (×) L-fit with discarded edges, (∗) LQ-fit without discarded edges, (□) LQ-fit with discarded edges).

investigate this further we looked at time-series taken from two other spatially extended systems where the local map is not contained in the fitting basis. First, we used the skewed logistic map as the local dynamics $f(x)$, i.e.

$$f(x) = \frac{-1 + ab(2x-1) + \sqrt{(1+ab)^2 - 4abx}}{2ab^2}, \tag{6.9}$$

where a is a nonlinearity parameter (it comes from the a in $ax(1-x)$) and b a skewness parameter. When $b = 0$ we recover the familiar logistic map studied above and as b is moved away from this value the map is then skewed to one side. The parameter is bounded by $-1/a < b < 1/a$ as in these limits the derivative becomes infinite at 0 and 1. We study here the case $b = 0.2$ corresponding to a high degree of skewness.

As a further example we also study the following local map

$$f(x) = xe^{r(1-x)}, \tag{6.10}$$

using $r = 4.5$. In both cases we used coupling parameter $\epsilon = 0.40$.

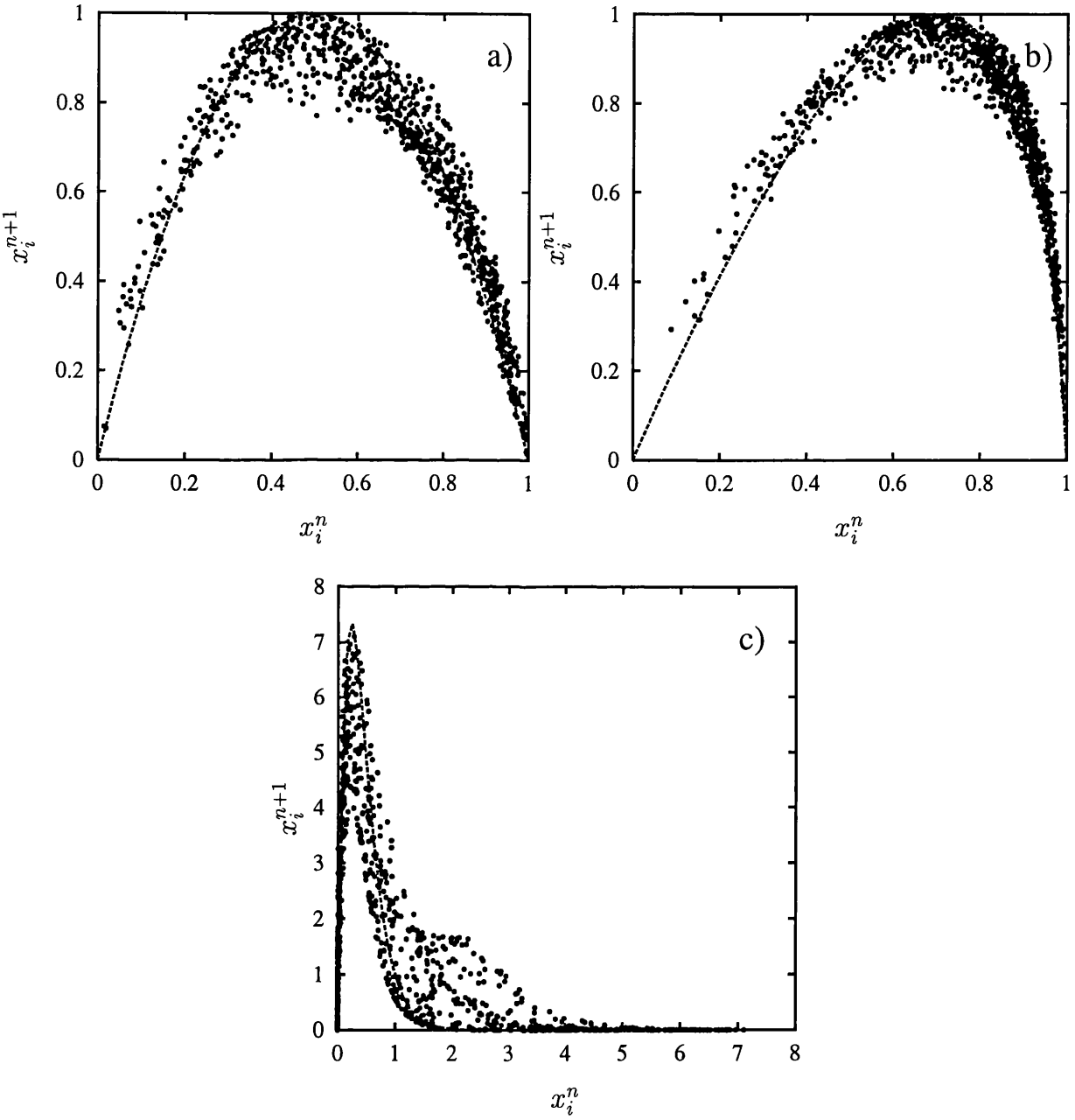


Figure 6.7: Time-delay plot of system with local dynamics: a) Logistic map ; b) Skewed logistic map with $b = 0.2$; c) Exponential map. The local maps are super-imposed as dashed lines.

In figure 6.7 we show time-delay plots of one site for the three systems we are looking at (i.e. plots of x_i^{n+1} vs. x_i^n). The local maps are super-imposed as continuous lines. We see that the time-delay plots retain some of the shape of the respective local maps.

We sampled 20 neighbouring sites as before and repeated the calculations to get approximations of sub-system LS and from there extract estimates of λ_1, ρ_d and ρ_h. We used the same choice of basis as before (L-fit and LQ-fit). As previously noted, using the polygon method and Jacobians with discarded edges gave consistently better results, thus we restrict our next analysis to this method.

In figure 6.8 we depict the interleaving and convergence to the limiting LS curve for the CML with local skewed logistic maps. As it can be noticed from a careful examination of Figs. 6.8.a and 6.8.b the interleaving in this case is not fully achieved. Moreover, from Figs. 6.8.c and 6.8.d we note that the convergence towards the real limit curve of the LS fails, especially for the most negative Lyapunov exponents.

The estimates for λ_1, ρ_d and ρ_h for the skewed logistic lattice are given in figure 6.9. The estimate for λ_1 is a bit higher than the real value (0.427) and the LQ-fit does not perform significantly better than the L-fit (figure 6.9.b). The estimates for ρ_d and ρ_h are also worse than when using the (un-skewed) logistic map as the underlying local dynamics. However, the real value of ρ_d (0.844) is not too badly approximated (figure 6.9.b). In the case of ρ_h the L-fit is better that the LQ-fit. In general we see that as the used basis (L-fit or LQ-fit) does not contain the underlying dynamics, the results deteriorate. In this particular case, the results are not too bad since the skewed logistic map is not far from being reproduced by the LQ basis.

To enhance the effects of the choice of the underlying system we now consider an exponential map as local dynamics. We continue using L-fit and LQ-fit to estimate Jacobians (discarding the outer layer as before). In this case it is clear that the basis is far from being a reasonable representation for the dynamics.

Again we calculated sub-system LS for increasing spatial delay reconstructions. The resulting interleaving and convergence is shown in figure 6.10. The interleaving

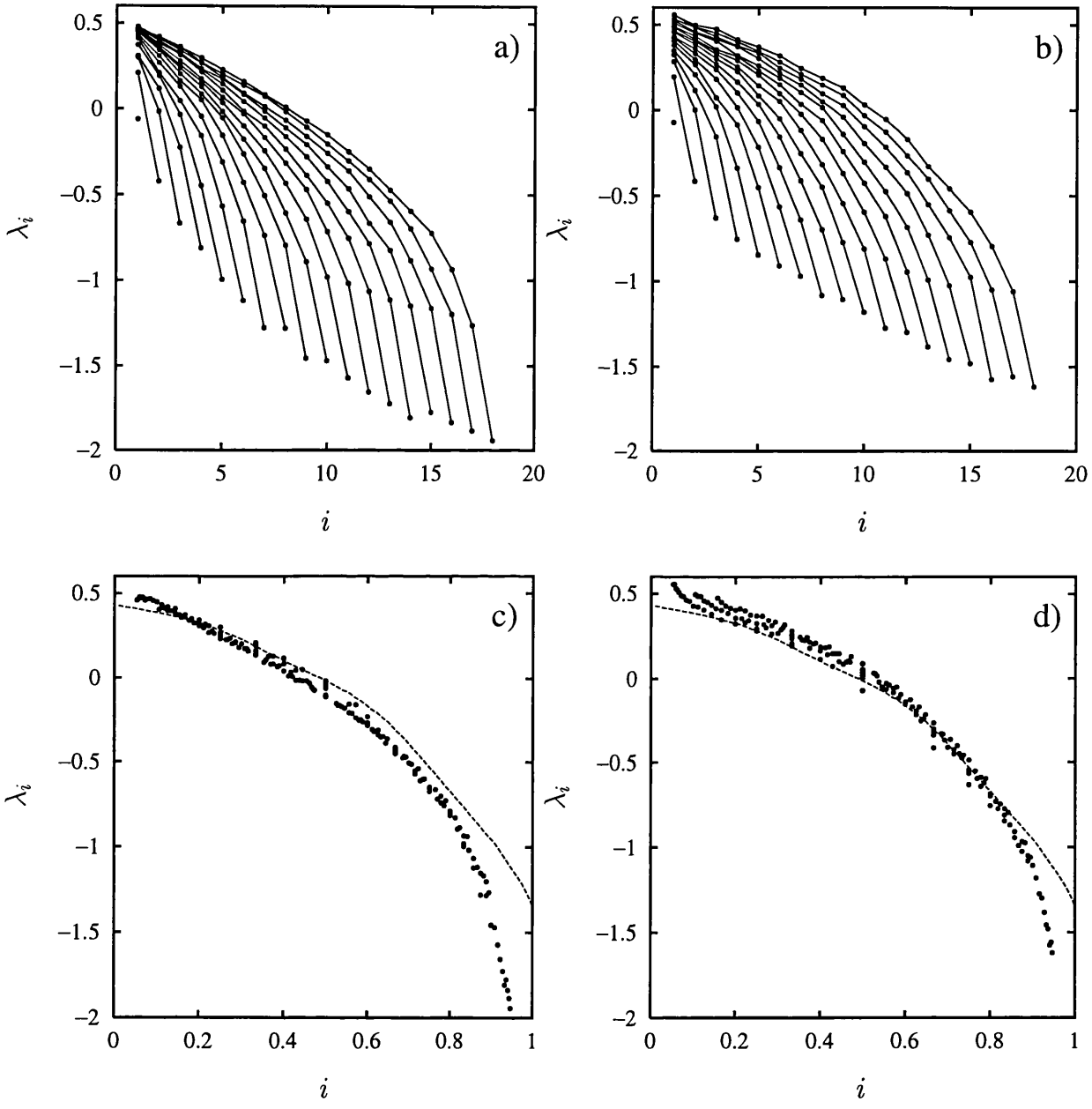


Figure 6.8: Estimated sub-system LS from time-series for increasing spatial delay reconstructions for the system using the skewed logistic map as local dynamics. a) Interleaving when using L-fit; b) interleaving using LQ-fit; c) convergence for L-fit; d) convergence for LQ-fit. In c) and d) the exact LS for the whole system is super-imposed as a continuous curve.

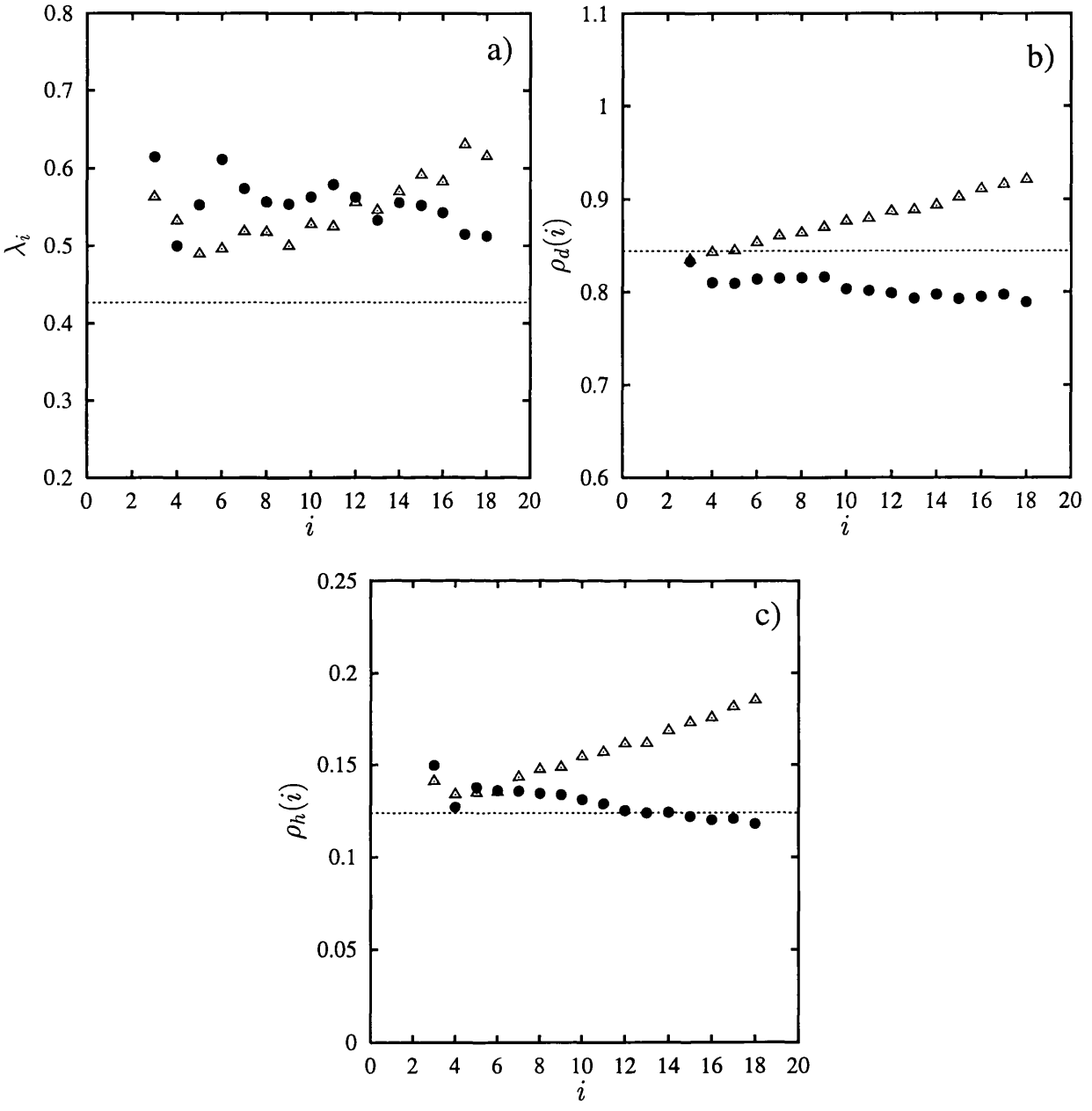


Figure 6.9: Estimation of a) largest Lyapunov exponent; b) Lyapunov dimension density; c) KS entropy density for the skewed system. (△) L-fit with discarded edges, (•) LQ-fit with discarded edges.

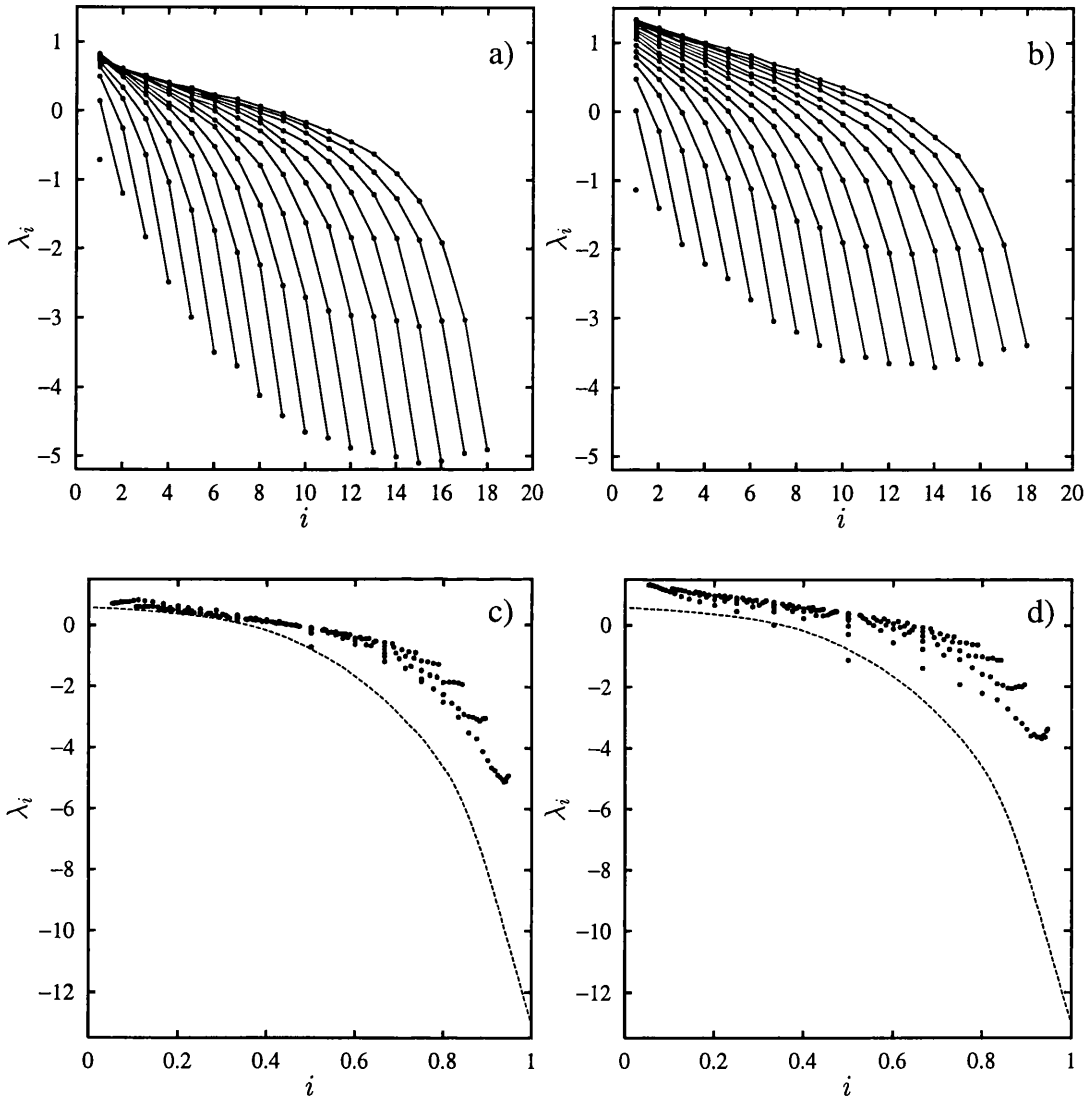


Figure 6.10: Estimated sub-system LS from time-series for increasing spatial delay reconstructions for the system using the exponential map as local dynamics. a) Interleaving when using L-fit; b) interleaving using LQ-fit; c) convergence for L-fit; d) convergence for LQ-fit. In c) and d) the exact LS for the whole system is super-imposed as a continuous curve.

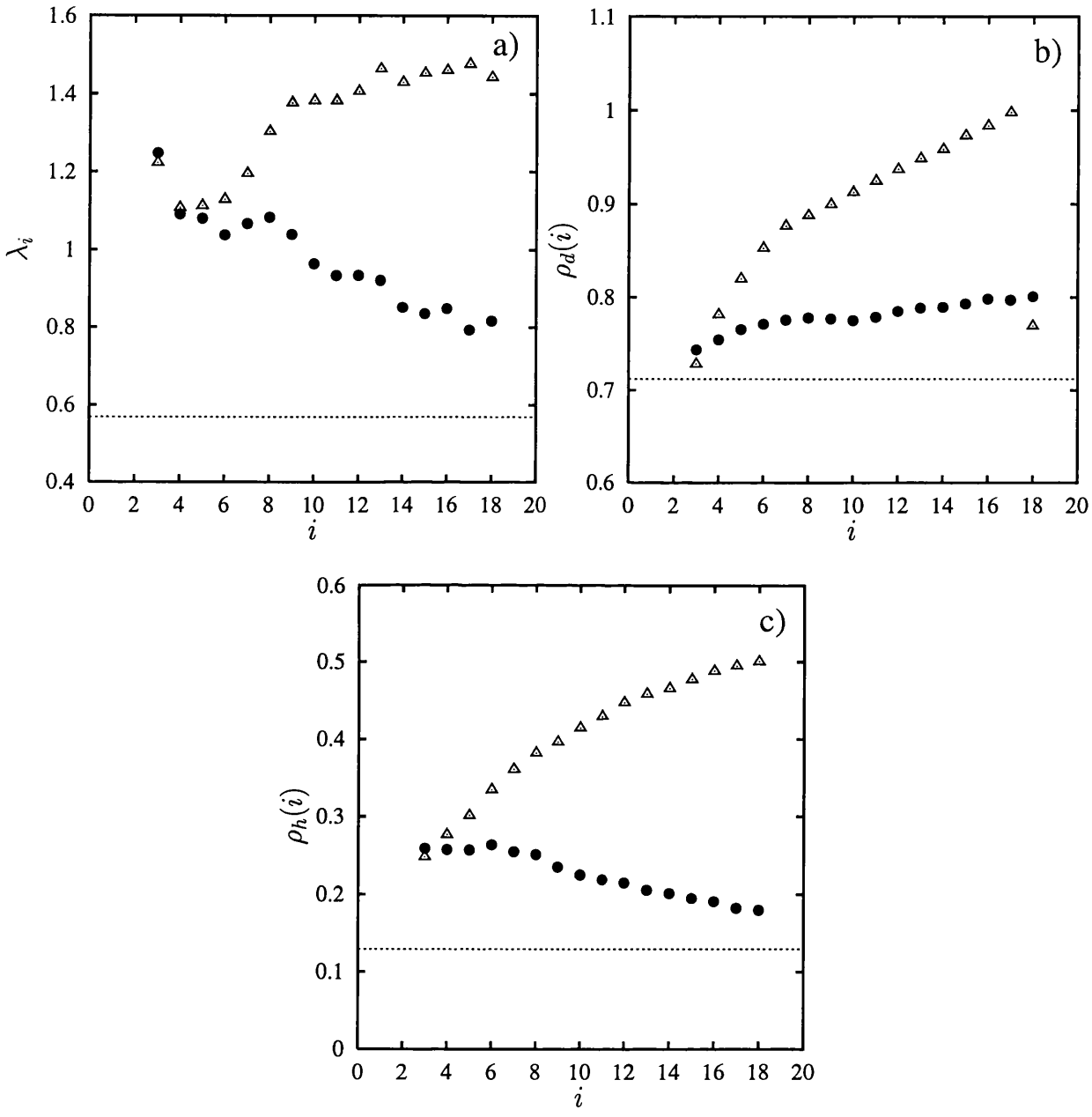


Figure 6.11: Estimation of a) largest Lyapunov exponent; b) Lyapunov dimension density; c) KS entropy density for the exponential system. (△) L-fit with discarded edges, (•) LQ-fit with discarded edges.

(figures 6.10.a and b), is still good for small d but starts to fail as $d \approx 10$ or greater. This result reflect that in small dimensions the approximation is local (so it is good). For large d the approximation is global and obviously bad for this exponential map. This is reflected in the convergence plots (figures 6.10.c and d). There is no convergence for the negative part of the LS. The L-fit perform much better than the LQ-fit and seems to converge to the whole LS in the upper part.

As before, we use the convergence plots with the polygon method to estimate λ_1, ρ_d and ρ_h as shown in figure 6.11. Here we clearly see that in all three cases the LQ-fit does not give any sensible results.

6.4 Discussion

The results given here can be summarized as follows.

We find that in certain circumstances it is possible to estimate intensive measures such as the largest Lyapunov exponent, Lyapunov dimension density and KS entropy density in spatially extended systems from time-series sampled from sufficiently large sub-systems. The success is governed by several factors. Firstly, the fitting of the Jacobians is done by a global fit, and care must go into choosing the right basis. The time-delay plots might be one way to get an idea of the kind of basis that should work better. Secondly, we find that discarding the outer layer of the Jacobian at each time step allow us to eliminate what would otherwise be noise input. We suggest to use a specific rescaling to form a convergence to the right limit curve depicting an approximation of the LS of the whole system and to use the described polygon method or a variant of this to extract the intensive quantities.

It seems clear that the usual method of local fitting of Jacobians does not let us estimate the Lyapunov dimension density and KS entropy density in that we cannot find close neighbours when we work in higher dimensions. The special shape of the Jacobians for the type of systems we have considered here, with non-zero entries only along the tri-diagonal, suggest that we might be able to estimate these using low-dimensional local approximations.

Chapter 7

Conclusions

Spatio-temporal dynamical systems are inherently very complex, not least because of their high dimensionality. Since these types of systems are so widespread and incorporate so many real-life systems and phenomena, understanding them is of great importance.

In this thesis we started out by describing the theory for the analysis of time-series from low-dimensional chaotic systems. It was our hope that we would be able to utilise these methods in the context of high-dimensional spatio-temporal dynamical systems. We feel that we have succeeded in this.

By using coupled map lattices as our model system we have been able to show that predictability is best achieved by using combination of time- and spatial delays locally. Remarkably, the predictability is best achieved in small dimensions, even if the dimension of the chaotic attractor is high-dimensional. We therefore suggest that one should think of such a system as a low dimensional local system, weakly coupled to a high dimensional system (i.e. 'the rest of the system'). A framework for doing this has been introduced.

Building on these results we compared a truncated lattice with noise inputs at the edges with a large deterministic lattice (essentially at the thermodynamic limit). As we increased the truncated lattice size we found an exponential convergence for measures such as the probability density, predictability, power spectrum and two-point correlation. This suggested that spatio-temporal embedding tech-

niques cannot detect the presence of spatial extent, suggesting the impossibility of reconstructing the whole system from localised information.

The most common tool to characterise chaos in a dynamical system is theLyapunov exponents which together give the Lyapunov spectrum (LS). Since spatio-temporal systems are high-dimensional the calculation of the LS becomes prohibitive. While studying sub-systems of coupled map lattices we noticed that the sub-system Lyapunov spectrum soon converged to the limit curve for the LS of the full system. We found a new rescaling that leads to better estimates of the real LS from sub-system information. It was inspired by the stability analysis of the homogeneous evolution in a one-dimensional coupled map lattice but appears to be equally valid in a much wider range of cases. We evaluated the performance of our rescaling method by comparing it to the conventional rescaling (dividing by the relative sub-system volume) for one and two-dimensional lattices in spatio-temporal chaotic regimes. In doing so we noticed that the Lyapunov spectra for consecutive sub-system sizes are interleaved and we discussed the possible ways in which this may arise. A full theoretical understanding of this phenomen is not given and should be subject to further research. We used the new rescaling to approximate quantities derived from the Lyapunov spectrum (largest Lyapunov exponent, Lyapunov dimension and Kolmogorov-Sinai entropy) finding better convergence as the sub-system size is increased than with conventional rescaling. We also proposed a natural method for constructing the Jacobian of systems on high-dimensional lattices.

The success of estimating the full LS from sub-system information led us to believe that we could achieve the same results when only time-series of the system was available. We found that by using increasing spatial embeddings, and given that the local map was suitable reflected in the basis when approximating the Jacobian, the LS was estimated remarkably well from time-series. We discussed how the new rescaling that we introduced for systems where the dynamics and system size is known *a priori*, could be modified in the case when only time-series is available. Using this modified version we suggested a geometric way of calculating the Lyapunov dimension density and KS entropy density. We found

that local linear approximations of the Jacobians along a trajectory may not work when calculating sub-system LS from time-series, but that by using an appropriate global basis the Jacobians might be extracted. We also considered a way of dealing with boundary effects in the sub-systems.

The special shape of the Jacobians for the type of systems we have considered, with non-zero entries only along the tri-diagonal, suggests that we might be able to estimate these using low-dimensional local approximations. This should be considered in future research.

All the results in this thesis were based on the analysis of coupled map lattices (and in most cases by using the one dimensional logistic map lattice). It would be interesting to see if the results are as good for other model systems, e.g. for coupled map lattices with higher-dimensional local dynamics, longer range coupling or in a partial differential equation framework.

Even more interesting would be to test the methods proposed in this thesis on data taken from a real system. Suitable data could e.g. be taken from a fluid flows, EEG recordings or remote satellite data. Only if the methods could show consistent results in these cases would they be entirely successful.

References

Abarbanel, H.D.I., Brown, R., Sidorowich, J.J. & Tsimring, L.Sh. (1993). The analysis of observed chaotic data in physical systems. *Rev.Mod.Phys.* **65**, 1331–1392.

Afraimovich, V.S., Ezersky, A.B., Rabinovich, M.I., Shereshevsky, M.A. & Zheleznyak, A.L. (1992). Dynamical Descriptions of Spatial Disorder. *Physica D* **58**, 331–338.

Barnett, S. (1990). *Matrices: methods and applications.* Oxford University Press. p. 349.

Bauer, M., Heng, H. & Martienssen, W. (1993). Characterization of spatiotemporal chaos from time series. *Phys. Rev. Lett.* **71**, 521–524.

Bauer, M. & Martienssen, W. (1991). Lyapunov exponents and dimensions of chaotic neural networks. *J. Phys. A* **24**, 4557–4566.

Beck, C. (1994). Chaotic cascade model for turbulent velocity distribution. *Phys. Rev. E* **49**, 3641–3652.

Bellman, R. (1960). *Introduction to matrix analysis.* McGraw-Hill.

Boffetta, G., Giuliani, P., Paladin, G. & Vulpiani, A. (1998). An Extension of the Lyapunov Analysis for the Predictability Problem. *J. Atmos. Sci.* **55**, 3409–3416.

Carretero-González, R. (1997). *Front propagation and mode-locking in coupled map lattices.* Ph. D. thesis, Queen Mary and Westfield College, London, U.K.

Carretero-González, R., Arrowsmith, D.K. & Vivaldi, F. (1997). Mode-locking in coupled map lattices. *Physica D* **103**, 381–403.

Carretero-González, R., Ørstavik S., Huke J., Broomhead D.S. & Stark J. (1999a). Scaling and interleaving of subsystem Lyapunov exponents for spatio-temporal systems. *Chaos* **9**, 466–482.

Carretero-González, R., Ørstavik S., Huke J., Broomhead D.S. & Stark J. (1999b). Thermodynamic limit from small lattices of coupled maps. Accepted for publication in *Phys. Rev. Lett.*.

Casdagli, M. (1991). Chaos and Deterministic *versus* Stochastic Non-linear Modelling. *J. R. Statist. Soc. B* **54**, 303–328.

Comins, H.N., Hassell, M.P. & May, R.M. (1992). The spatial dynamics of host-parasitoid systems. *J. Anim. Ecology* **61**, 735–748.

Davis, P.J. (1979). *Circulant Matrices.* John Wiley & Sons.

Eckmann, J.-P. & Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Modern Phys.* **57**, 617–656.

Essex, C., Lookman, T. & Nerenberg, M.A.H. (1987). The climate attractor over short timescales. *Nature* **326**, 64–66.

Farmer, J.D. & Sidorowich, J.J. (1987). Predicting Chaotic Time Series. *Phys. Rev. Lett.* **59**, 845–848.

Furstenberg, H. & Kesten, H. (1960). Products of Random Matrices. *Ann. Math. Stat.* **31**, 479–469.

Gade, P.M. & Amritkar, R.E. (1993). Spatially periodic orbits in coupled map lattices. *Phys. Rev. E* **47**, 143–153.

Geist, K., Parlitz, U. & Lauterborn, W. (1990). Comparison of different methods for computing Lyapunov exponents. *Prog. Theor. Phys.* **83**, 875–893.

Grassberger, P. (1989). Information content and predictability of lumped and distributed dynamical systems. *Phys. Scr.* **40**, 346–353.

Grassberger, P. & Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D* **9**, 189–208.

Hassell, M.P., Comins, H.N. & May, R.M. (1991). Spatial structure and chaos in insect population dynamics. *Nature* **353**, 255–258.

Johnson, R.A., Palmer, K.J. & Sell, G.R. (1987). Ergodic properties of linear dynamical systems. *SIAM J. Math. Anal.* **18**, 1–33.

Kaneko, K. (1983). Transition from torus to chaos accompanied by frequency lockings with symmetry breaking. *Prog. Theor. Phys.* **69**, 1427–1442.

Kaneko, K. (1984). Period-doubling of kink-antikink patterns, quasiperiodicity in anti-ferro-like structures and spatial intermittency in coupled logistic lattice. *Prog. Theor. Phys.* **72**, 480–486.

Kaneko, K. (1986). Lyapunov analysis and information flow in coupled map lattices. *Physica D* **23**, 436–447.

Kaneko, K. (1989). Towards thermodynamics of spatiotemporal chaos. *Prog. Theor. Phys. Suppl.* **99**, 263–287.

Kaneko, K. (1992). Mean field fluctuation of a network of chaotic elements - remaining fluctuation and correlation in the large size limit. *Physica D* **55**, 368–384.

Kaneko, K. (1993). *Theory and applications of coupled map lattices.* John Wiley & Sons, Chichester.

Kantz, H. & Olbrich, E. (1997). Scalar Observations From a Class of High Dimensional Chaotic Systems: Limitations of the Time Delay Embedding. *Chaos* **7**, 423–429.

Kaplan, J.L. & Yorke, J.A. (1979). Chaotic behaviour of multidimensional difference equations. *Lecture Notes in Mathematics, Springer, New York* **730**, 204–207.

Kapral, R., Livi, R., Oppo, G.-L. & Politi, A. (1994). Dynamics of complex interfaces. *Phys. Rev. E* **49**, 2009–2022.

Keeler, D. & Farmer, J.D. (1986). Robust space-time intermittency and $1/f$ noise. *Physica D* **23**, 413–435.

Kennel, M.B., Brown, R. & Abarbanel, H.D.I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **45**, 3403–3410.

Lemaître, A., Chaté, H. & Manneville, P. (1997). Conditional mean-field for chaotic coupled map lattices. *Europhys. Lett.* **39**, 377–382.

Little, S., Ellner, S., Pascual, M., Neubert, M., Kaplan, D., Sauer, T., Caswell, H. & Solow, A. (1996). Detecting nonlinear dynamics in spatio-temporal systems, examples from ecological models. *Physica D* **96**, 321–333.

Lorenz, E.N. (1963). Deterministic Nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

Lorenz, E.N. (1969). Atmospheric predictability as revealed by naturally occurring analogoues. *J. Atmos. Sci.* **26**, 636–646.

Lorenz, E.N. (1991). Dimension of Weather and Climate Attractors. *Nature* **353**, 241–242.

Manneville, P. (1985). Liapounov exponents for the Kuramoto-Sivashinsky model, in "Macroscopic modeling of turbulent flows", Springer-Verlag, Berlin. *Lecture Notes in Physics* **230**, 319–326.

Mayer-Kress, G. & Kaneko, K. (1989). Spatiotemporal chaos and noise. *J. Stat. Phys.* **54**, 1489–1508.

Muldoon, M.R., Broomhead, D.S. & Huke, J.P. (1994). Delay Reconstruction for multiprobe signals. *IEE Digest* **143**, 3/1–3/5.

Ørstavik, S., Carretero-González R., Huke J., Broomhead D.S. & Stark, J. (1999). Estimation of invariant quantities in spatio-temporal systems from time-series. Preprint, to be submitted to *Physica D.*

Ørstavik, S. & Stark, J. (1998). Reconstruction and cross-prediction in coupled map lattices using spatio-temporal embedding techniques. *Phys. Lett. A* **247**, 145–160.

Osborne, A.R. & Provenzale, A. (1989). Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* **35**, 357–381.

Oseledec, V.I. (1968). A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19**, 197–231.

Ott, E., Sauer, T. & Yorke, J.A. (1994). *Coping with Chaos: Analysis of chaotic data and the exploitation of chaotic systems.* John Wiley & Sons, New York.

Packard, N.H., Crutchfield, J.P., Farmer, J.D. & Shaw, R.S. (1980). Geometry from a time series. *Phys. Rev. Lett.* **45**, 712–716.

Parekh, N., Kumar, V.R. & Kulkarni, B.D. (1996). Analysis and characterization of complex spatio-temporal patterns in nonlinear reaction-diffusion systems. *Physica A* **224**, 369–381.

Parekh, N., Kumar, V.R. & Kulkarni, B.D. (1997). Control of spatiotemporal chaos: a study with an autocatalytic reaction-diffusion system. *Pramana J. of Physics* **48**, 303–323.

Parekh, N., Kumar, V.R. & Kulkarni, B.D. (1998). Synchronization and control of spatiotemporal chaos using time-series data from local regions. *Chaos* **8**, 300–306.

Parlett, B.R. (1980). *The symmetric eigenvalue problem.* Prentice-Hall.

Pikovsky, A. & Kurths, J. (1994). Collective behaviour in ensembles of globally coupled maps. *Physica D* **76**, 411–419.

Pritchard, D. & Theiler, J. (1994). Generating Surrogate Data for Time Series with Several Simultanously Measured Variables. *Phys. Rev. Lett.* **73**, 951–954.

Puccioni, G.P., Torcini, A., Politi, A. & G.D'Alessandro (1991). Fractal dimension of spatially extended systems. *Physica D* **53**, 85–101.

Rand, D.A. & Wilson, H.B. (1995). Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. *Proc. R. Soc. Lond. B* **259**, 111–117.

Rohani, P. & Miramontes, O. (1995). Host-parasitoid metapopulations: the consequences of parasitoid aggregation on spatial dynamics and searching efficiency. *Proc. R. Soc. Lond. B.* **260**, 335–342.

Rombouts, S.A.R.B, Keunen, R.W.M. & Stam, C.J. (1995). Investigation of nonlinear structure in multichannel EEG. *Phys. Lett. A* **202**, 352–358.

Ruelle, D. (1982). Large volume limit distribution of characteristic exponents in turbulence. *Commun. Math. Phys.* **87**, 287–302.

Ruelle, D. (1983). Five turbulent problems. *Physica D* **7**, 40–44.

Sauer, T., Yorke, J.A. & Casdagli, M. (1991). Embedology. *J. Stat. Phys* **65**, 579–616.

Schiff, S.J., So, P., Chang, T., Burke, R.E. & Sauer, T. (1996). Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble. *Phys. Rev. E* **54**, 6708–6724.

Schreiber, T. (1995). Efficient neighbor searching in nonlinear time series analysis. *Int. J. Bif. Chaos* **5**, 349–358.

Schreiber, T. & Schmitz, A. (1997). Discrimination power for nonlinearity in a time series. *Phys. Rev. E* **55**, 5443–5447.

Schuster, H.G. (1988). *Deterministic chaos* (second ed.). VCH Verlagsgesellschaft, Weinheim.

Sinai, Ya.G. (1996). A remark concerning the thermodynamic limit of the Lyapunov spectrum. *Int. J. Bifurcation and Chaos* **6**, 1137–1142.

Sinai, Ya.G. & Chernov, N.I. (1982). Entropy of a gas of hard spheres with respect to the group of space-time shifts. *Proc. of Petrovski Seminar* **8**, 218–238.

Sinha, S., Biswas, D., Azam, M. & Lawande, S.V. (1992). Local-to-global coupling in chaotic maps. *Phys. Rev. A* **46**, 6242–6246.

Stark, J. (1999). Delay Embeddings for Forced Systems: I. Deterministic Forcing. *J. Nonlinear Sci.* **9**, 255–332.

Stark, J., Broomhead, D.S., Davies, M.E. & Huke, J.P. (1997). Takens Embedding Theorems for Forced and Stochastic Systems. *Nonlinear Analysis* **30**, 5303–5314.

Sugihara, G. & May, R.M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344**, 734–741.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L.-S. Young (Eds.), *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, pp. 366–381. Springer-Verlag.

Tsimring, L.S. (1993). Nested strange attractors in spatiotemporal chaotic systems. *Phys. Rev. E* **48**, 3421–3426.

Tsonis, A.A. & Elsner, J.B. (1992). Nonlinear prediction as a way of distinguishing chaos from random fractal sequences. *Nature* **358**, 217–220.

Vannitsem, S. & Nicolis, C. (1996). Error growth dynamics in spatially extended systems. *Int. J. Bifurcation and Chaos* **6**, 2223–2235.

von Bremen, H.F., Udwadia, F.E. & Proskurowski, W. (1997). An efficient QR based method for the computation of Lyapunov exponents. *Physica D* **101**, 1–16.

Wales, D.J. (1991). Calculating the rate of loss of information from chaotic time series by forecasting. *Nature* **350**, 485–488.

Willeboordse, F.H. & Kaneko, K. (1995). Pattern dynamics of a coupled map lattice for open flow. *Physica D* **86**, 428–455.

Wilson, H.B. & Rand, D.A. (1997). Reconstructing the dynamics of unobserved variables in spatially-extended systems. *Proc. R. Soc. Lond. B.* **264**, 625–630.

Zheleznyak, A.L. & Chua, L.O. (1994). Coexistence of Low- and High-Dimensional Spatiotemporal Chaos in a Chain of Dissipatively Coupled Chua's Circuits. *Int. J. Bifur. Chaos* **4**, 639–674.

Zhilin, Q. & Gang, H. (1994). Spatiotemporally periodic states, periodic windows, and intermittency in coupled-map lattices. *Phys. Rev. E* **49**, 1099–1107.

Zhilin, Q., Gang, H., Benkun, M. & Gang, T. (1994). Spatiotemporally periodic patterns in symmetrically coupled map lattices. *Phys. Rev. E* **50**, 163–170.