UNIVERSITY OF TRENTO
DEPARTMENT OF PHYSICS

THESIS
SUBMITTED TO THE
DOCTORAL SCHOOL IN PHYSICS
BY

MATTEO FRANCHI

IN CANDIDATURE FOR THE DEGREE OF
PHILOSOPHIAE DOCTOR - DOTTORE DI RICERCA

DEVELOPMENT OF NEW ANALYTICAL TECHNIQUES FOR CHAOTIC TIME SERIES

TUTOR: LEONARDO RICCI

27°PhD cycle

Contents

Chapter 1

Introduction

Noise, referred to as a random fluctuation of a physical quantity, is an intrinsic and unavoidable aspect in physical phenomena. The reason lays first of all in the very core of quantum mechanics, namely in theory of measurement and indetermination principle. A second main source, within the framework of statistical mechanics, is the fluctuation/dissipation theorem: every measurement or detection process requires a transfer and a dissipation of energy; this dissipation leads to an energy fluctuation and thus to a noise contribution to the measurement. Typically, time-varying quantities are of great interest. We refer to these quantities as "signals".

The term "noise" is often given a negative connotation. However, the presence of noise is not necessarily a detrimental aspect. For example, the physical phenomenon occurring in nonlinear systems and known as "stochastic resonance" [1, 2] is linked to the amplification of small signals by means of an optimal level of noise.

Noise is also unavoidably present within neural systems. A main question is whether it is simply a by-product of the communication processes within these systems or, rather, whether it has a role in their operation. The work described in this thesis is part of a project whose aim is the investigation of possible correlations between the level of stochastic noise in the brain and neurological pathologies and disorders like epilepsy, autism and obsessive-compulsive disorder. With this goal in mind, the first main issue

to tackle is the determination of the amount of stochastic noise within a signal stemming from the brain.

The signals we consider in our project are electroencephalographic recordings (EEG). Electroencephalography is a noninvasive technique invented in the second part of the nineteenth century. Since then, many techniques have been developed to analyse EEG recordings. Typical sampling bandwidth are of order 1 kHz. In a seminal paper in 1985, Babloyantz *et al.* [3] demonstrated the presence of a deterministic chaotic component in EEG recordings. Presently, nonlinear analysis is carried out on EEG recordings to study the dynamics of the brain activity and its relation with to the cognitive state and also with pathologies and disorders like epilepsy, autism, Alzheimer's and Parkinson's disease [4].

The chaotic nature of the EEG signal and its apparent similarity to stochastic noise prevent us from measuring the level of stochastic noise by using traditional techniques. Thus, the second main issue concerns the separation of the noisy component from the chaotic one within an EEG recording. In 1997, Gao [5] proposed a method to separate these two components within a sampled time series. This method relies on the possibility to reconstruct the deterministic dynamics by means of the Takens embedding theorem [6, 7], by using the so called embedding technique.

Takens embedding theorem is indeed of crucial importance to analyse experimental signals: it proves that is not necessary to know all the components of a multi-dimensional signal generated by a system in order to determine the characteristic parameters of the system's dynamics. Unfortunately, Takens embedding theorem does not give any clue on how to identify the optimal parameters that define the embedding. As a consequence, finding a method to determine the optimal embedding becomes a crucial point in the quest for a reliable method to assess the level of deterministic chaos and of stochastic noise within a time series.

The scientific literature contains several proposals that tackle the issue of optimal embedding; as a matter of fact, none of these methods relies on an objective algorithm but, rather, require a "subjective" (i.e. made by an observer) trimming and evaluation of key quantities. These methods

are: Gao–Zheng method [8, 9, 10], the method of Schuster [11], the method of characteristic length, the method of global false-nearest-neighbours and autocorrelation function, the method of global false-nearest-neighbours and mutual information [12, 13]. Another major issue with these optimal embedding methods regards which one is the most reliable one. Cellucci *et al.* [14] proposed a technique to assess the best optimal embedding method. Given a method and a known dynamical system $\mathcal{S}$, the idea consists in evaluating the maximum Lyapunov exponent $\mathrm{MLE}_\mathcal{S}$ of the system by analysing a time series generated by it and embedded according to the method. The maximum Lyapunov exponent $\mathrm{MLE}_\mathcal{S}$ is then compared with the reference value $\chi_\mathcal{S}$ calculated by means of the so-called *standard method* [15, 16], a procedure that is independent from any embedding and relies on the integration of the system's tangent map. The calculation of maximum Lyapunov exponent $\mathrm{MLE}_\mathcal{S}$ out of the time series is carried out by means of the *divergent rate method*, i.e. a method that evaluates the local divergence rate of nearby trajectories. According to Cellucci *et al.*, the best optimal embedding method is that one that, given a set of known dynamical systems, provides the smallest differences between $\mathrm{MLE}_\mathcal{S}$ and $\chi_\mathcal{S}$. This comparison method can be used only with systems characterized by equations that allow for the implementation of the *standard method*, like, for example, the Lorenz attractor.

To sum up, the goal of assessing the amount of stochastic noise within an EEG recording or, more in general, an experimental time series generated by a chaotic dynamical system, relies on the solution of several issues: above all, a reliable method to determine the optimal embedding, and a reliable method to distinguish the entity of the stochastic noise component from the deterministic chaos component in a time series.

In this work, two main results are presented. The first is a study of the statistical properties of the finite-time maximum Lyapunov exponent determined out of a time series by using the *divergent rate method*. To reach this goal, we developed a new, completely automatic algorithm based on the method developed by Gao and Zheng [8]. A main achievement of this part of the work is the interpretation [17] of the uncertainty in the light

of the work by Grassberger, Badii e Politi of 1988 [18] on the theoretical distribution of maximum Lyapunov exponents. We showed that the analysis and identification of clusters in diagrams representing uncertainty vs. maximum Lyapunov exponent can provide useful information about the optimal choice of the embedding parameters. In addition, our results allow for the identification of systems that can provide suitable benchmarks for the comparison and ranking of different embedding methods.

The second main result concerns the development of a new method for the assessment of the optimal embedding parameters. Our method is based on two assumptions: a potential-like quantity is defined on the lattice of points that characterize the embedding; the optimal embedding choice coincides with local extrema (maxima or minima) of this potential.

Throughout the work, we used "synthetic" time series generated by numerically integrating the difference and differential equations that describe the following dynamical systems: the Hénon map, the Lorenz attractor, the Rössler attractor and the Mackey–Glass attractor. These four systems are widely used as references in the scientific literature.

In the last part of the work, we have started to examine EEG recordings by using the techniques developed in the main part of the work. The EEG recordings are sampled on healthy subjects in resting-state. These investigations are still at a starting phase.

The present work is organized as follows.

Chapter 2 is devoted to a general description of dynamical systems and chaotic dynamics. The most recurrent definitions of chaos in the scientific literature are presented; in particular, we discuss the definition of chaos proposed by Devaney. The second part of this chapter is devoted to examples of chaotic systems and the computational methods to simulate them.

Chapter 3 deals with the embedding of a time series and presents the Takens

embedding theorem. Besides a discussion on the main characteristics and issues linked to this theorem, we introduce the concept of optimal embedding.

Chapter 4 presents the main characteristics of a chaotic dynamical system. A special attention is devoted to the Lyapunov characteristic exponents, evaluated by using the *standard method*, and the maximum Lyapunov exponent, evaluated by means of the *divergent rate method*. The chapter contains also a review on other important tools for the analysis of a dynamical system: the dimension of an attractor, the Kolmogorov entropy and the mutual information. In this chapter we also present our implementation of the *divergent rate method* to estimate the maximum Lyapunov exponent and its uncertainty.

Chapter 5 discusses the statistical properties of the maximum Lyapunov exponent and its uncertainty as a function of the embedding parameters. A main achievement is the interpretation of the observed uncertainty of the maximum Lyapunov exponent calculated via the *divergent rate method* [17] in the light of the work by Grassberger, Badii e Politi of 1988 [18] on the theoretical distribution of the maximum Lyapunov exponent.

Chapter 6 , upon a detailed presentation of two optimal embedding methods used in the scientific literature, discusses a new optimal embedding method based on a Kolmogorov–Smirnov approach. The new approach is "automatic", in the sense that it does not require any subjective interpretation (as it happens for the others).

Chapter 7 contains a review on electroencephalography and a discussion on the chaotic behaviour of brain function.

Chapter 8 deals with final considerations and future perspectives of the research project.

Chapter 2

Dynamical systems

> Behold the rule we follow, and the only one we
> can follow: when a phenomenon appears to us as
> the cause of another, we regard it as anterior. It is
> therefore by cause that we define time ...
>
> Henri Poincaré

This chapter deals with dynamical systems, the definition of chaotic dynamics and the integration method to simulate a continuous dynamical system. The chapter is organized as follows. Sec. 2.1 is an introduction to the concept of dynamical system. In Sec. 2.2 the definition of a chaotic system is presented. In Sec. 2.3 some examples of chaotic attractors are shown. Sec. 2.4 finally concerns the simulation of a chaotic attractor by using an integration computer algorithm.

2.1   Introduction

The concept of dynamical system was born with Newton in the 17th century. However, the father of the modern theory of dynamical systems is Henri Poincaré. Poincaré was also the pioneer of simulating a differential equation by means of a discrete-time system. Indeed, any computer simulation as well as any experimental recording is based on this assumption.

The concept of dynamical system is based on the fact that the time-evolution of a system follows "deterministic" rules: the state of a system at a given moment of time determines the state of this system for any future time.

A dynamical system is a tuple $\mathcal{S} \equiv (\mathcal{T}, M, \Phi)$ where $\mathcal{T}$ is a monoid that represents time, M is a manifold (i.e. the phase space) of the system, and Φ is the evolution function:

$$\Phi : \mathcal{T} \times M \to M. \tag{2.1}$$

Dynamical systems can be classified in real dynamical systems, discrete dynamical systems and cellular automata.

A real dynamical system, also known as continuous time dynamical system, is a tuple $(\mathcal{T}, M, \Phi)$, where $\mathcal{T} \subset \mathbb{R}$, M is a manifold locally diffeomorphic to a Banach space, and Φ is a continuous function. Often, for this kind of systems, Φ corresponds to the integration of a differential equation:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \quad t \in \mathbb{R}, \, \mathbf{x} \in \mathbb{R}^n, \tag{2.2}$$

A discrete dynamical system is a tuple $(\mathcal{T}, M, \Phi)$, with $\mathcal{T} \subset \mathbb{Z}$. This kind of systems is often described by means of a set of difference – rather than differential – equations:

$$\mathbf{x}(t + 1) = \mathbf{f}(\mathbf{x}(t)) \quad t \in \mathbb{Z}, \, \mathbf{x} \in \mathbb{R}^n.$$

A cellular automaton is a dynamical system made of "cells" and characterized by a discrete time and a discrete state space. The evolution rule assigns a new state of each cell as a function of the old state of this cell and its neighbouring cells. The rule is the same for each cell. An example of a cellular automaton is the *game of life* discussed by Martin Gardner in 1970 [19].

The numerical integration of a difference equation is relatively straightforward: starting from a state, the next state can be calculated by directly applying the difference equation itself. On the other hand, the numerical

integration of differential equations poses more difficulties; for example, we must take care of the integration step dt and it is necessary to use specific methods to reduce the integration errors. Possibly the most used method is Runge–Kutta's [20] one. The simulation of a continuous dynamical system is discussed in Sec. 2.4.

The aim of the theory of dynamical systems is to understand the behaviour of the orbits (trajectories). An orbit can be fixed, periodic, quasi-periodic and chaotic. The assessment of this characteristic is crucial to study the evolution of a system. A system can have different types of orbits depending on the initial conditions and/or the parameters that define the evolution equations. A fixed orbit is a set of phase space points that does not change when time changes. A periodic orbit is an orbit that can be written in the form $\mathbf{x}(t + T) = \mathbf{x}(t)$, where T is the period of the system. A quasi-periodic orbit is an orbit that can be written as a finite sum of periodic functions. The chaotic behaviour is presented in the following section.

2.2 Definition of a chaotic dynamical system

Despite chaos might be regarded as a well-known concept, its definition is indeed not unique. The most remarkable attempts to define chaos [21, 22] were provided by Li and Yorke [23], by Block and Coppel [24] and by Devaney [25].

According to Li and Yorke [23, 22]:

Definition 2.1 (*L/Y-chaos*)**.** A continuous map $\Phi : M \rightarrow M$ on a compact metric space (M, d) is said to be chaotic – in the sense of Li and Yorke, or *L/Y-chaotic* – if there exists an uncountable subset S (referred to as a scrambled set) of M having the following properties:

(i.) $\limsup_{n \to \infty} d(\Phi_n(x), \Phi_n(y)) > 0$ for all $x, y \in S$, $x \neq y$;

(ii.) $\liminf_{n \to \infty} d(\Phi_n(x), \Phi_n(y)) = 0$ for all $x, y \in S$, $x \neq y$;

(iii.) $\limsup_{n \to \infty} d(\Phi_n(x), \Phi_n(p)) > 0$ for all $x \in S$, $x \in M$, p periodic.

According to Block and Coppel [24, 22]:

Definition 2.2 (*B/C-chaos*)**.** A continuous map $\Phi : M \to M$ on a compact metric space M is said to be chaotic – in the sense of Block and Coppel, or *B/C-chaotic* – if there exists $m \in \mathbb{N}$ and a compact Φ_m-invariant subset Y of M such that $\Phi_m|_Y$ is semi-conjugate to the shift on Σ or, equivalently, if there exists a continuous surjection $\Psi : Y \to \Sigma$ satisfying

$$\Psi \circ \Phi_m = \sigma \circ \Psi \qquad \text{on} \quad Y.$$

In this work we use the definition of chaos given by Devaney [25, 22, 26, 27] and based on the topological approach of the problem, as follows:

Definition 2.3 (*D-chaos*)**.** Let M be a set and $\Phi : M \to M$ a map on this set. We say that Φ is chaotic – in sense of Devaney, *D-chaos* – on M if:

 (a.) Φ has a sensitive dependence on initial conditions;

 (b.) Φ is topologically transitive;

 (c.) periodic points are dense in M.

It is worth noting that the sensitive dependence on initial conditions, the topological transitiveness and the density of periodic points are not independent; for example, for a metric space, topological transitivity and denseness of periodic orbits imply sensitive dependence on initial conditions [28].

Usually, in Physics and applied sciences, people focus on point (a.), thus using the notion of chaos that describes sensitive dependence on initial conditions. The commonest method for distinguishing between regular and chaotic motion, and thus quantifying the sensitive dependence on initial conditions, is the evaluation of the maximum Lyapunov exponent χ_0: if $\chi_0 > 0$ the orbit is chaotic [27].

2.2.1   Sensitive dependence on initial conditions

Sensitivity to initial conditions is popularly known as "the butterfly effect", in the wake of a celebrated paper published by Edward Lorenz in 1972

for the American Association for the Advancement of Science (Washington, D.C., USA) and entitled "Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?" [29]. The flapping wing represents a small change in the initial condition of the system, which nevertheless causes a chain of events leading to large-scale phenomena. Had the butterfly not flapped its wings, the trajectory of the system might have been vastly different.

The sensitive dependence on initial conditions can be defined as follow:

Definition 2.4 (Sensitive dependence on initial conditions). $\Phi : M \to M$ has sensitive dependence on initial conditions if there exists $\epsilon > 0$ such that, for any $x \in M$ and any neighbourhood Δ of x, there exists $y \in \Delta$ and $n \geqslant 0$, such that $\|\Phi_n(x) - \Phi_n(y)\| > \epsilon$, where Φ_n denotes n successive applications of Φ.

This definition implies that there exist points arbitrarily close to x which eventually separate from x by at least ϵ under iterations of Φ. We point out that not all points near x need eventually to move away from x under iteration, but there must be at least one such point in every neighbourhood of x. Consequently, the slightest error ϵ in any initial condition x can lead to a macroscopic discrepancy Δ in the evolution of the dynamics of the system.

A consequence of the sensitivity to initial conditions is that, if we start with a finite amount of information about the system (as it is usually the case in practice), then, beyond a certain time, the system will no longer be predictable. The possibly most familiar example of this behaviour is weather, which is generally predictable only about a week ahead.

2.2.2 Topological transitivity

Topological transitivity (or topological mixing) is defined as the characteristic of a chaotic system to evolve over time so that any given region or open set of its phase space will eventually overlap with any other given region.

Definition 2.5. (Topological transitivity) $\Phi : M \to M$ is said to be topologically transitive if for any pair of open sets $U, W \subset M$ there exists $n > 0$ such that $\Phi_n(U) \cap W \neq \varnothing$.

This definition implies the existence of orbits that will eventually connect any arbitrarily small neighbourhood to any other one. Consequently, a dynamical system cannot be decomposed into two disjoint invariant, open sets. This mathematical concept of "mixing" corresponds to the standard intuition; mixing of coloured dyes or fluids is an example of a chaotic system.

Topological mixing is often omitted in papers on chaos, typically because chaos is linked exclusively to the sensitivity to initial conditions. However, sensitive dependence on initial conditions alone does not necessarily yield chaos. For example, the simple dynamical system produced by repeatedly doubling an initial value has a sensitive dependence on initial conditions, since any pair of nearby points will eventually become widely separated. However, such a dynamical system has no topological mixing (for example, there is no orbit linking negative with positive numbers), and shows therefore no chaos.

2.2.3   Density of periodic orbits

Density of periodic orbits means that every point in the space is approached arbitrarily by close periodic orbits. Topologically mixing systems not satisfying this condition may not display sensitivity to initial conditions, and thus they may be not chaotic.

2.3   Prototypical chaotic dynamical systems

In this section, chaotic dynamical systems widely used in the scientific literature are presented: Hénon map, logistic map, Ikeda map, Lorenz attractor, Rössler attractor, Mackey–Glass attractor, Duffing attractor and Chua circuit.

Hénon map, Logistic map and Ikeda map are discrete-time dynamical systems and are governed by difference equations. These systems are com-

monly used and studied because their integration is simpler and computationally less expensive than in case of continuous systems.

Lorenz attractor, Rössler attractor, Mackey–Glass attractor, Duffing attractor and Chua circuit are continuous dynamical systems; they are simulated by integrating the corresponding differential equations via a Runge–Kutta method (see Sec. 2.4 for more details). These systems are used to model many phenomena like weather, sun's activity, electric circuits, chemical reactions, acoustics, celestial mechanics and lasers [30].

It is worth noting that, while equations describing a chaotic behaviour might be often surprisingly simple [30], their solutions cannot be expressed analytically. In addition, trajectories are neither steady nor periodic and indeed appear to be very complex: when these motions are considered, adjectives like "wild", "turbulent", and "random" often occur.

2.3.1 Hénon map

In 1976, Michel Hénon [31] introduced the map – thereupon named after him – as a simplified model of the Poincaré section of the Lorenz model. This map is a discrete-time dynamical system and one of the most studied examples of dynamical systems that exhibit chaotic behaviour [32, 33, 34, 35]. Hénon map is governed by the following system of equations:

$$
\begin{aligned}
x_{n+1} &= 1 - a x_n^2 + y_n \,, \\
y_{n+1} &= b x_n \,,
\end{aligned}
\tag{2.3}
$$

where a and b are two constant parameters. The values of these parameters that are most used in the scientific literature and such that the Hénon map is chaotic are $a = 1.4$, $b = 0.3$. Depending on the values of a and b, the map can be chaotic, intermittent, or even converge to a periodic orbit. Fig. 2.1 shows a simulation of the Hénon map.
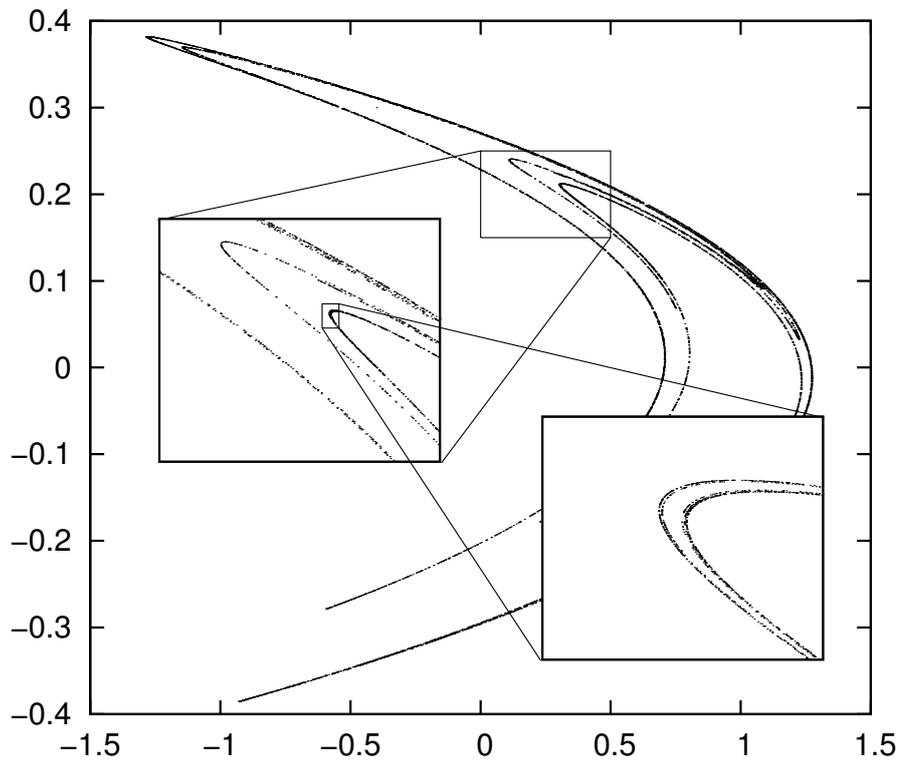
Figure 2.1: Simulation of Hénon map generated by meand of Eq. (2.3) and setting $a = 1.4$, $b = 0.3$; the starting point is $(1, 0)$. In this picture two zoom of the map are shown to highlight the chaotic behaviour and the fractal characteristic of this system.

2.3.2 Logistic Map

The logistic map was popularized in 1976 by the biologist Robert May [36]: his paper, entitled "Simple mathematical models with very complicated dynamics", discusses a discrete-time demographic model similar to the logistic equation first created by Pierre François Verhulst [37]. The logistic map was devised as a simple idealized ecological model for the yearly variations in the population of insects [30]. The map is a polynomial mapping of degree two, described by the following difference equation:

$$x_{n+1} = rx_n(1 - x_n),$$

r value	periodic point
$r \geqslant 0$	0
$r \geqslant 1$	$\frac{(r-1)}{r}$
$r \geqslant 4$	are dense in $[0, 1]$

Table 2.1: Periodic points of logistic map as a function of the parameter r.

where $x_n \in [0, 1] \in \mathbb{R}$ represents the ratio of existing population to the maximum possible population at year n (for example, x_0 represents the initial ratio of population to the maximum population at year 0), whereas r is a positive number that represents a combined rate for reproduction and starvation.

The logistic map exhibits periodicity for various values of the parameter r (see Tab. 2.1). For $r \geqslant 4$ the map has a chaotic behaviour.

The logistic map can be used to generate random numbers [38, 39]; Pareek *et al.* [40] used this map to develop cryptosystems and to encrypt images [41].

2.3.3 Ikeda map

The Ikeda map was first proposed by Ikeda *et al.* to model the dynamics of light in the ring cavity containing a nonlinear dielectric medium [42, 43]. It is modelled by the following equation:

$$z_{n+1} = p + \mu z_n \exp\left[i\kappa - \frac{i\alpha}{|z_n|^2 + 1} \right], \tag{2.4}$$

where p, μ, κ and α are the parameters of the equation and z_n is a complex number. In the model, z_n stands for the electric field inside the resonator at the n-th step of rotation in the resonator; p and κ are parameters that indicate the amount of light injected from the outside and the linear phase across the resonator, respectively. The parameter $\mu \leqslant 1$ is referred to as dissipation parameter and characterizes the resonator loss.

The Ikeda map is also used to describe a large number of systems of different nature [44, 45].

Fig. 2.2 shows an example of Ikeda map, obtained by using Eq. (2.4) a
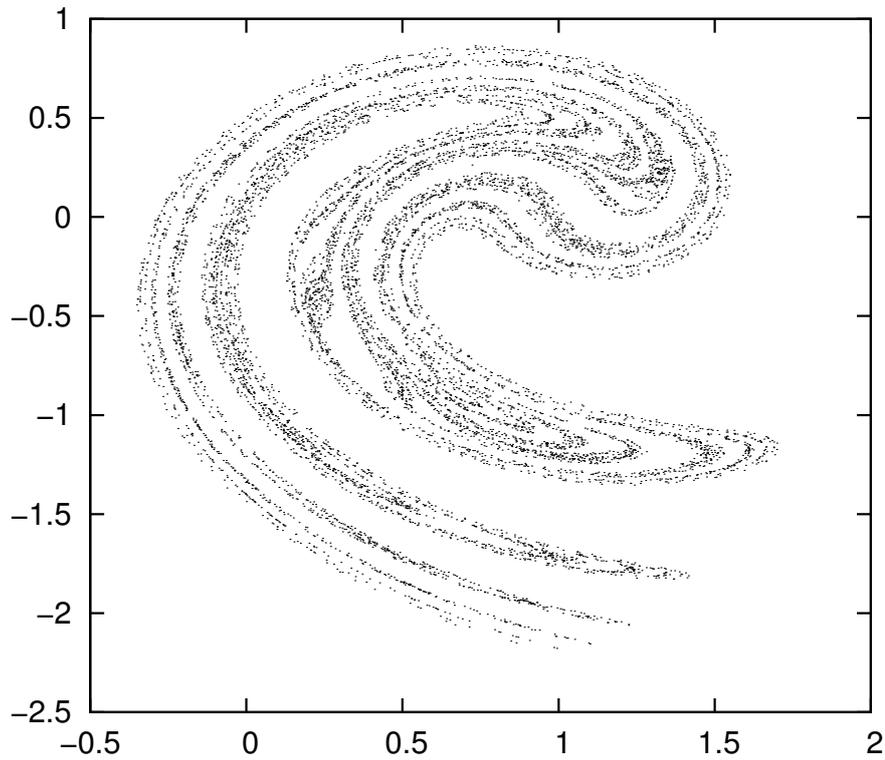
Figure 2.2: Simulation of the Ikeda map generated by Eq. (2.4) with using the parameter values $p = 1.0$, $\mu = 0.9$, $\kappa = 0.4$, and $\alpha = 6.0$ [46]. For the n-th point, the abscissa and the ordinate correspond to the real and the imaginary part of z_n.

setting the parameters as following: $p = 1.0$, $\mu = 0.9$, $\kappa = 0.4$, and $\alpha = 6.0$. A chaotic behaviour is clearly exhibited.

2.3.4   Lorenz attractor

The Lorenz attractor was introduced by, and thereupon named after, Edward Lorenz in 1963 [47], who derived it from the simplified equations of convection rolls arising in the equations of the atmosphere. In addition to its interest to the field of nonlinear mathematics, the Lorenz model has important implications for climate and weather forecast as well in simplified models for lasers [48, 49], dynamos [50, 51], thermosyphons [52], brushless DC motors [53], electric circuits [54], chemical reactions [55] and forward

osmosis [56].

Like other maps and attractors, the Lorenz system has chaotic solutions for certain parameter values and initial conditions; when plotted, it resembles a butterfly or figure "eight" (see Fig. 2.3). The Lorenz attractor is governed by the following system of equations [47]:

$$
\begin{aligned}
\frac{dx}{dt} &= \sigma(y - x)\,, \\
\frac{dy}{dt} &= x(\rho - z) - y\,, \\
\frac{dz}{dt} &= xy - \beta z\,,
\end{aligned}
\tag{2.5}
$$

where $(x, y, z) \in \mathbb{R}^3$ is the system state, σ is the so-called *Prandtl number* and ρ the so-called *Rayleigh number*. The parameters β, ρ, σ must be positive; widely-used values are $\beta = \frac{8}{3}, \sigma = 10$. The system exhibits chaotic behaviour for $\rho = 28$ but displays knotted periodic orbits for other values of ρ. In the original paper by Lorenz x, y and z are not space coordinates:

- x is proportional to the intensity of convective motion;

- y is proportional to the difference in temperature between the ascending and descending currents and denotes that warm fluid is rising and cold fluid is descending;

- z is proportional to the distortion of vertical temperature profile from linearity, a positive value indicating that the strongest gradients occur near the boundaries.

2.3.5 Rössler attractor

The Rössler attractor was proposed by Otto E. Rössler in 1976 [57]. Its behaviour is similar to the Lorenz attractor (see Sec. 2.3.4). The Rössler
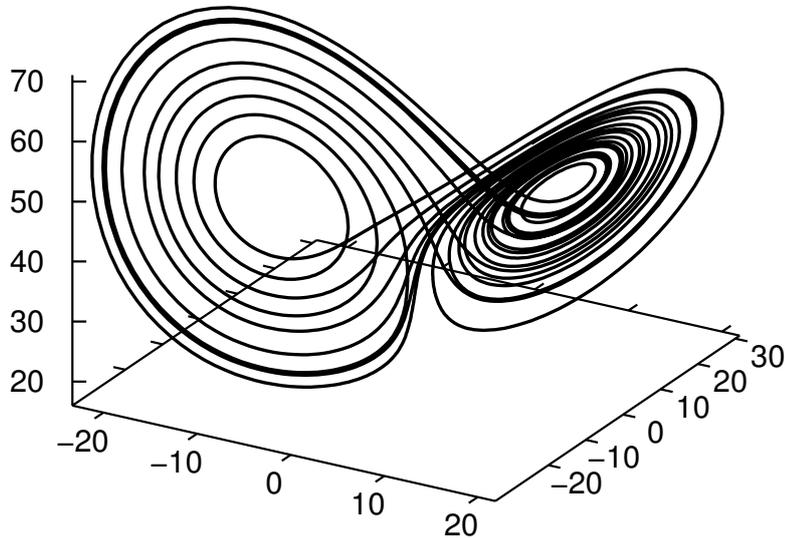
Figure 2.3: Simulation of the Lorenz attractor obtained by integrating Eq. (2.5) with $\sigma = 10$, $r = 45.92$, $b = 8/3$ and a sampling interval δt equal to 0.01.

attractor is governed by the following system of equations [57]:

$$
\begin{aligned}
\frac{dx}{dt} &= -y - z, \\
\frac{dy}{dt} &= x + ay, \\
\frac{dz}{dt} &= b + z(x - c),
\end{aligned}
\tag{2.6}
$$

where $(x, y, z) \in \mathbb{R}^3$ are again dynamical variables defining the phase space and $(a, b, c) \in \mathbb{R}^3$ are constant parameters.

Fig. 2.4 shows the integration of the system of equations (2.6). The orbit follows an outward spiral close to the x, y-plane and turning around an unstable fixed point; when the radius of the spirals becomes large enough,

a second fixed point influences the orbit, causing a rise and twist in the *z*-dimension. Fig. 2.5 shows the plots of *x*, *y* and *z* in the time domain: clearly, the variables *x* and *y* are "chaotically" oscillating within a fixed range of values.
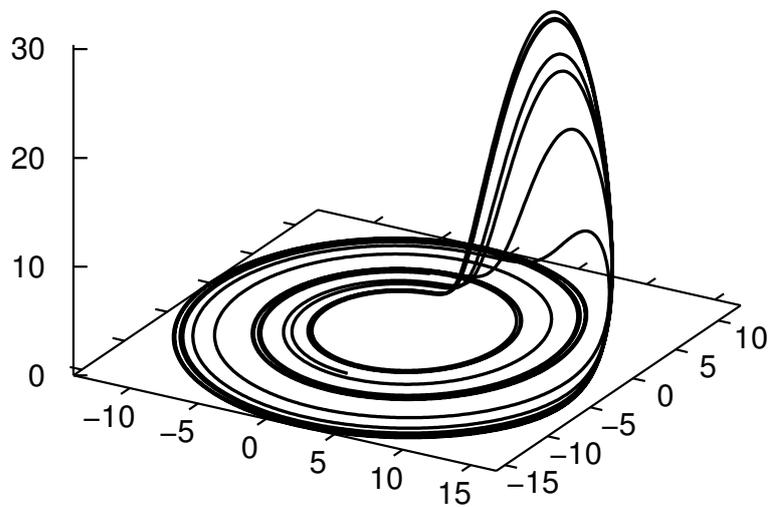
Figure 2.4: Simulation of Rössler attractor obtained by integrating Eq. (2.6) with $a = 0.15$, $b = 0.2$, $c = 10$ and a sampling interval δt equal to 0.125.

2.3.6 Mackey–Glass attractor

The Mackey–Glass attractor is a feedback system that models the concentration of electrolytes, oxygen, glucose, and blood cells in the blood, as well as blood pressure within the brain and other organs. Possibly the simplest differential equation representing a feedback system is

$$\frac{dx}{dt} = \lambda - \gamma x, \qquad (2.7)$$
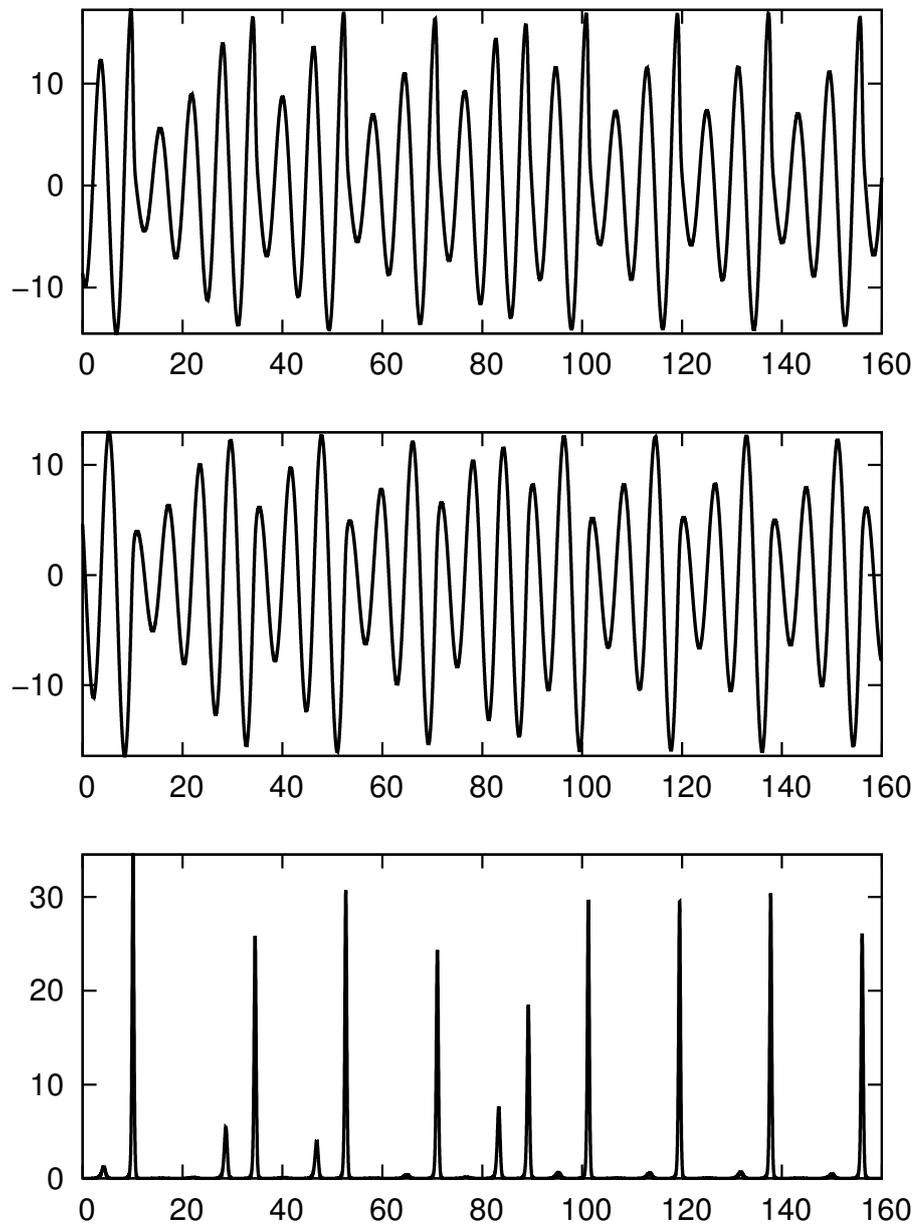
Figure 2.5: Time evolution of the components of the Rössler attractor shown in Fig. 2.4.

where λ and γ are positive constants and x is the variable of interest: λ and γx respectively correspond to the rate at which x is produced and destroyed. It is well known that the evolution of the solution of Eq. (2.7) for $t \to \infty$ is

given by $x \to \frac{\lambda}{\gamma}$, independently from the initial conditions. Thus, the related system has a stable fixed point.

In real feedback systems, however, there is a time lag between the sensing of the value of x and the feedback response. The Mackey–Glass equation allows for the modelization of real feedback systems. It is governed by the following equation [58]:

$$\frac{dx(t)}{dt} = a\frac{x(t-\tau)}{1 + [x(t-\tau)]^c} - bx(t), \quad a, b, c > 0, \tag{2.8}$$

where a, b, c and τ are constant parameters. In particular, τ is non-negative and represents a time delay.
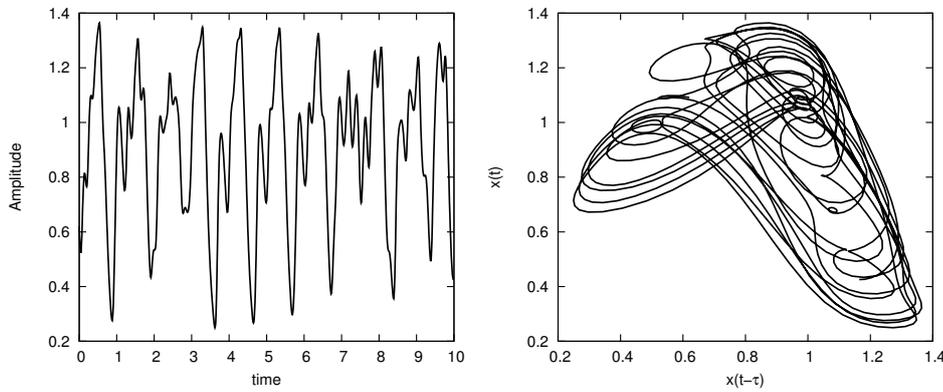
Figure 2.6: Simulation of the Mackey–Glass attractor obtained by integrating Eq. (2.8) with $a = 0.2$, $b = 0.1$, $c = 10$, $\tau = 30$ and a sampling interval δt equal to 1.5. On the left side, the system variable x is plotted as a function of time. On the right side, "delay" plot is reported.

Fig. 2.6 shows the result of an integration of Eq. (2.8). The chaotic behaviour is evident. Here it is worth citing the work by Namajunas *et al.* [59] developed an electric circuit to simulate the Mackey–Glass attractor.

2.3.7 Duffing attractor

In 1918, Georg Duffing, a German electrical engineer, introduced the original model of the oscillator named after him. Since then, the related equation has become, along with Van der Pol's equation, one of the most

common examples of nonlinear oscillation in textbooks and research articles. Duffing oscillator is an example of a periodically forced oscillator with a nonlinear elasticity, described by the following nonlinear second-order differential equation:

$$\ddot{x} + \delta\dot{x} + \beta x + \alpha x^3 = \gamma\cos(\omega t), \tag{2.9}$$

where α, β, δ, γ and ω are real parameters.

Besides the modelling of electrical and mechanical systems [60], the Duffing attractor is used to model biological systems [61], in the detection of weak signals [62] and to investigate in the framework of control and synchronization of dynamical systems [63, 64, 65, 66]. The integration of Eq. (2.9) occurs upon rewriting it as a system of first-order differential equations:

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= -\delta y - \beta x - \alpha x^3 + \gamma\cos(\varphi), \\ \frac{d\varphi}{dt} &= \omega. \end{aligned} \tag{2.10}$$

In Fig. 2.7 a simulation of Duffing attractor is shown.

2.3.8 Chua circuit

The most famous chaotic electronic circuits is possibly Chua's circuit [67], whose scheme is reported in Fig. 2.8. The circuital element crucial to the production of chaos is Chua's diode, a device with nonlinear current vs. voltage characteristics typically represented by the following piecewise linear equation (see also the rightmost plot of Fig. 2.8):

$$i_d(x) = m_0 x + \frac{1}{2}(m_1 - m_0)(|x + B_p| - |x - B_p|),$$

where m_0, m_1 and B_p are parameters. Chua's circuit has been widely studied due to its easy experimental implementation, robustness, and reproducibility of results. For these reasons, this circuit is also used as a test
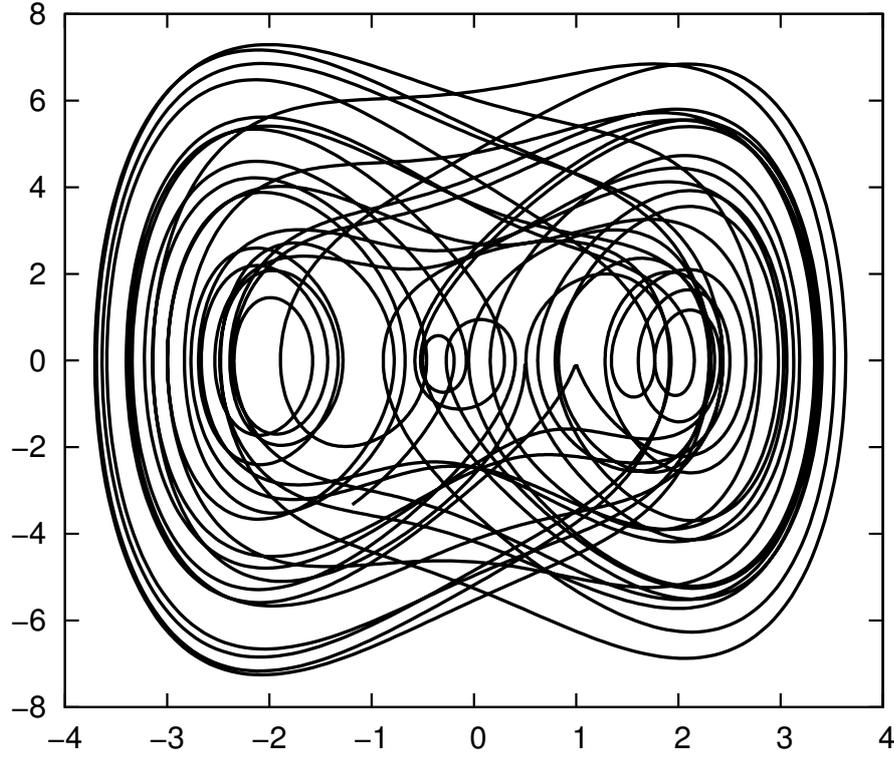
Figure 2.7: Simulation of the Duffing attractor obtained by integrating Eq. (2.10) with $\alpha = 1.0$, $\beta = -1.0$, $\gamma = 7.5$, $\delta = 0.05$ and $\omega = 1$.

platform for chaos control [68, 69, 70], synchronization [71, 72, 69], secure communications [73, 74] and signal encryption [75].

The differential equations governing the circuit are derived from Kirchhoff's laws: if $x(t)$, $y(t)$, and $z(t)$ represent the voltage across the capacitors C_1, the voltage across the capacitors C_2, and the electric current in the inductor L, respectively, the system of differential equations is [76]:

$$
\begin{aligned}
\frac{dx}{dt} &= \frac{y - x}{RC_1} - \frac{i_d(x)}{C_1}, \\
\frac{dy}{dt} &= \frac{x - y}{RC_2} - \frac{z}{C_2}, \\
\frac{dz}{dt} &= -\frac{y}{L} - z\frac{r_L}{L},
\end{aligned}
\tag{2.11}
$$
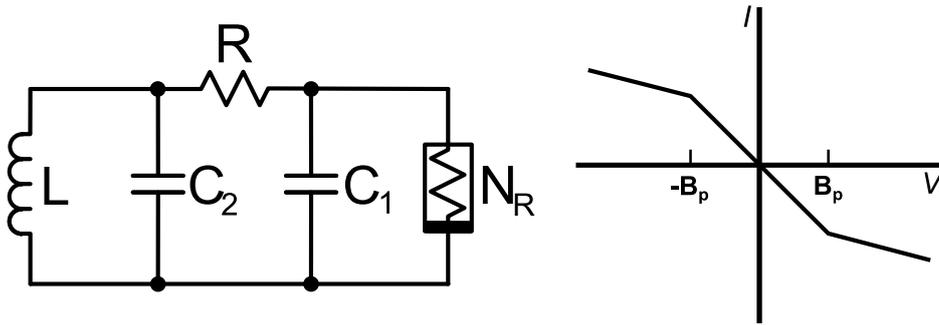
Figure 2.8: A standard Chua circuit (left side). The dynamical variables of Eq. (2.11) (x, y, z) correspond to the voltage across the capacitor C_1, the voltage across the capacitor C_1 and the current that flows through the inductor L, respectively. The characteristics of the Chua's diode is also plotted (right side).
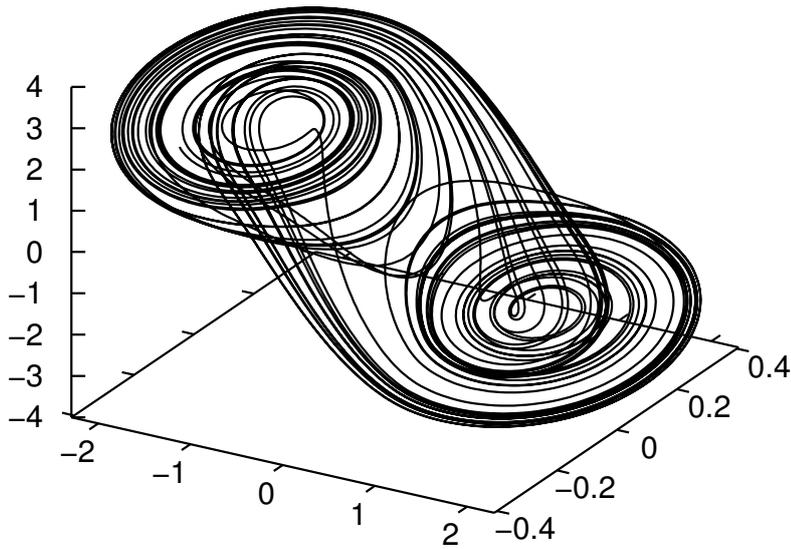
Figure 2.9: Simulation of Chua's circuit by integrating Eq. (2.11) with $R = 1$, $C_1 = 1$, $C_2 = 0.064$, $L = 0.04$, $B_p = 1$, $m_0 = -5/7$ and $m_1 = -8/7$.

where R, C_1, C_2 and L are passive linear elements and r_L is the inductor resistance. In Fig. 2.9 a simulation of Chua's attractor is shown.

2.4   Simulation of chaotic attractors

As mentioned above, the simulation of a discrete-time dynamical system is implemented by directly applying the difference equations that govern the system. On the other hand, to simulate a continuous-time dynamical system we must integrate the differential equation(s) over time. This process is subject to the discretization error, due to the finite time step of integration. In the simulation of a chaotic dynamical system, this kind of error make up a major issue, because of the sensitive dependence on initial conditions (see Sec. 2.2.1). In the scientific literature, the most used method to integrate chaotic dynamical systems is the 4th-order Runge–Kutta method [77, 78, 79, 80].

In this work, to integrate differential equations we used an 8th-order Runge–Kutta algorithm. In particular, we used the implementation available in the GNU Scientific Library (GSL) [81], and developed in 1981 by Dormand and Prince [82]. The Dormand–Prince algorithm is an adaptive stepsize integration algorithm (routine name: *gsl_odeiv2_step_rk8pd*). In the book "Solving Ordinary Differential Equation I: Nonstiff problems" [83, pp. 171-172] a comparison between different integration methods is discussed: the 8th-order [81] – an evolution of Dormand–Prince 5th-order [84] method – performs better than the others. The Dormand–Prince method is exposed in details in the book by Hairer, Nørsett, Wanner [83, p. 181].

Fig. 2.10 shows the integration of the Lorenz attractor, Eq. (2.5), by using a 4th-order Runge–Kutta method and the 8th-order Runge–Kutta method (Dormand–Prince algorithm). By assuming the latter to make up the reference solution, it is clear that the integration error increases during time [85]. Fig. 2.11 shows the logarithmic plot of the difference of the two solutions: the difference exponentially increases in the region $0 \leqslant t \leqslant 22$. Interestingly, the plot is very similar to the divergent exponent used to calculate the Lyapunov exponent via embedding, as explained in Sec. 4.1.2 (see also
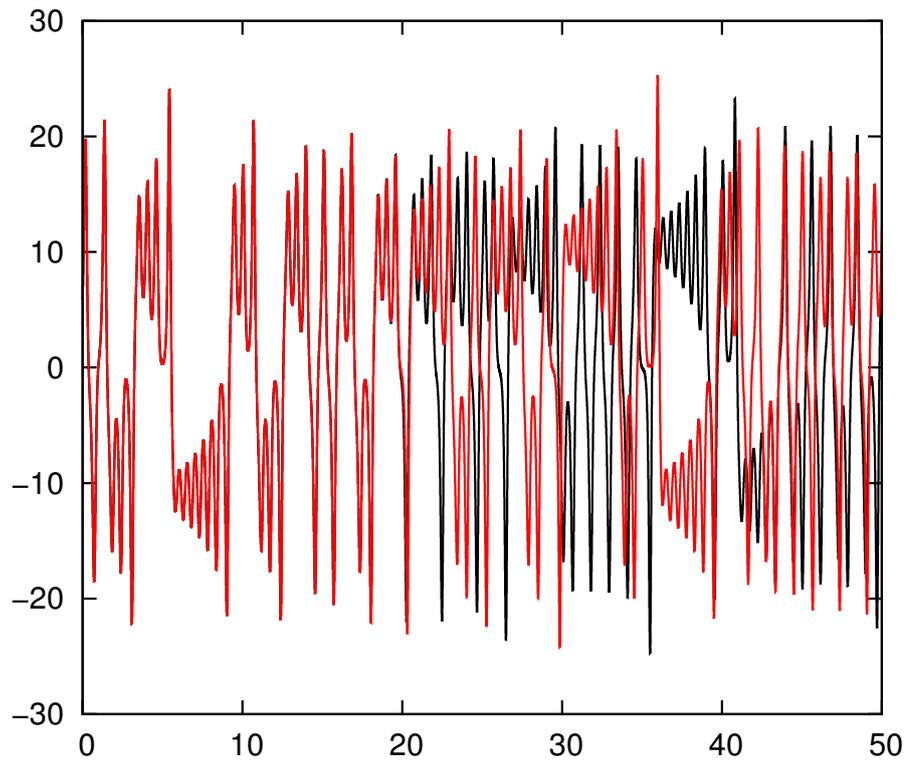
Figure 2.10: Evolution of the x-component of the Lorenz attractor using the Eq. (2.5) with $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 45.92$; the red line is integrated using a 4th-order Runge–Kutta, the black line is integrated using an 8th-order Runge–Kutta method developed by Dormand and Prince [82]

Fig. 4.4). In particular, the slopes of the growing sections of the plots are almost coincident: 1.22(1) for the present plot, to be compared with 1.23, a value that corresponds to the maximum Lyapunov exponent for the Lorenz attractor (given the parameters used in this work).

2.5 Reference systems

In following chapters, four systems are used as references:

- Hénon map, Eq. (2.3), with $a = 1.4$ and $b = 0.3$;

- Lorenz attractor, Eq. (2.5), with $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 45.92$ and time
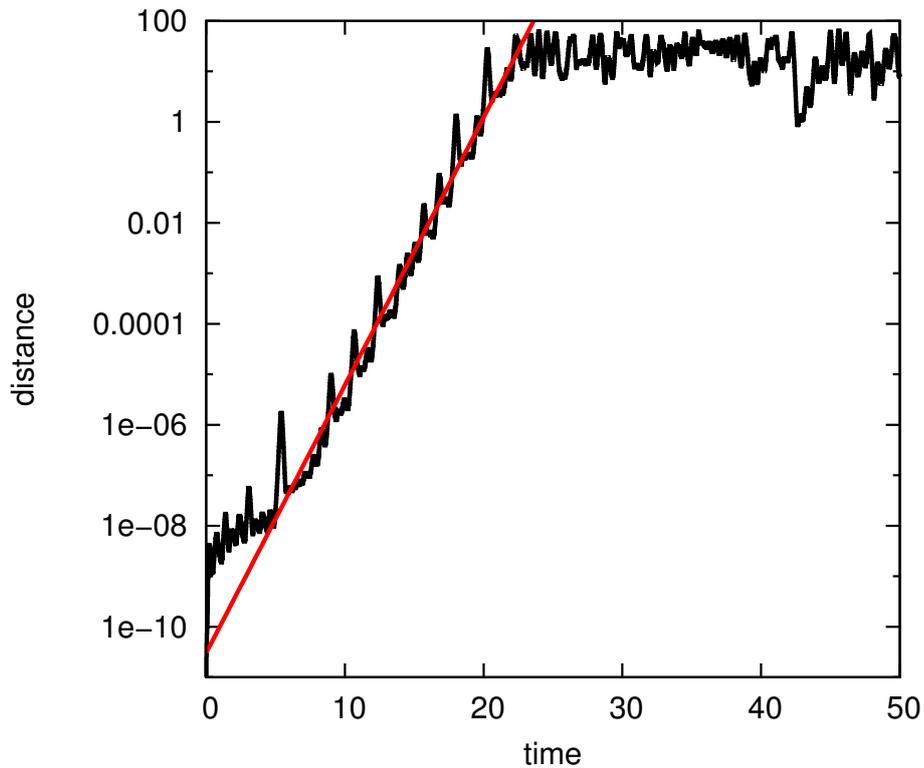
Figure 2.11: Logarithmic plot of the distance between the numerical integrations of the Lorenz attractor carried out by a 4th-order and a 8th-order Runge–Kutta method. An exponential dependence of the distance over time is highlighted by means of the red line.

step $\delta t = 0.03$;

- Rössler attractor, Eq. (2.6) with $a = 0.15$, $b = 0.2$, $c = 10$ and time step $\delta t = 0.125$;

- Mackey–Glass attractor, Eq. (2.8) with $a = 0.2$, $b = 0.1$, $c = 10$, $\tau = 30$ and $\delta t = 1.5$.

So, for example, the Lorenz attractor with the parameters set as above is henceforth referred to as the "reference Lorenz attractor".

Chapter 3

Embedding of time series

This chapter deals with the recording of experimental time series, the delay embedding and the Takens embedding theorem.

In Sec. 3.1 the sampling of a continuous-dynamical system is described and the delay embedding is defined. Sec. 3.2 is devoted to the Takens embedding theorem and its generalization. Finally, a discussion about the optimal embedding parameters is the topic of Sec. 3.3.

3.1 Basic concepts

Experiments are characterized by continuous dynamical systems that can be modelled by means of differential equations. At any time, the state of a system is described by a point $\mathbf{x}$ lying on a d-dimensional manifold $M \subset \mathbb{R}^d$. The time evolution of the system is described by a tuple $\mathcal{S} \equiv (\mathbb{R}, M, \Phi)$ (see Sec. 2.1); fixing the initial condition at time $t = 0$, a state $\mathbf{x}(t)$ at time t is evaluated as

$$\mathbf{x}(t) = \Phi_t \left[\mathbf{x}(0) \right] .$$

During an experiment, the state $\mathbf{x}$ is typically sampled at integer multiples of a time interval (sampling time) δt; the measuring process discretizes Φ_t, by creating a time series $\{ \Phi_{n \cdot \delta t} \left[\mathbf{x}(0) \right] \}$ where $n = 0, 1, 2, \ldots, n \in \mathbb{N}$. We can

set $\mathbf{x}_n \equiv \mathbf{x}(n \cdot \delta t)$ so that the previous expression becomes

$$\mathbf{x}_n = \Phi_{n \cdot \delta t}(\mathbf{x}_0).$$ (3.1)

An example of this measurement process applied to a continuous dynamical system with $d = 1$ is represented in Fig. 3.1.

The measurement process corresponds to a map $\varphi : M \to Y$, where $Y \subset M$. Typically, we do not have access to all components of the state $\mathbf{x}$, but, rather, we can only observe a projection of the state on a lower-dimensional space. The most extreme situation occurs when a single scalar component of the dynamical system is observed; in this case $\varphi : M \to \mathbb{R}$ and the recorded time series is $\{\varphi(\mathbf{x}_n)\}$. Henceforth we set $\varphi_n \equiv \varphi(\mathbf{x}_n)$.

The observed measurement φ_n is one-dimensional, whereas $\Phi_{n \cdot \delta t}$ lies in $\mathbb{R}^d$, with d generally larger than unity. At first, it might thus appear that $\{\varphi_n\}$ contains relatively little information about the behaviour of Φ. In other words, the question is to which extent $\{\varphi_n\}$ can provide information on the original continuous dynamical system.

Many statistical approaches and techniques, like the discrete Fourier transform ot the Wavelet transform, can be used to extract information from an experimental sequence $\{\varphi_n\}$. Unfortunately, most methods are not useful if the time series has a chaotic behaviour. On the other hand, Takens embedding theorem [6, 86, 87] gives us a technique to reconstruct Φ from the observed time series $\{\varphi_n\}$.

The importance of Takens embedding theorem can be highlighted by looking at the number of publications per year that cite the original paper by Takens [6]. Fig. 3.2 shows the plot of the yearly citation rate. In the last fifteen year, the paper gets about 140 citations per year.

Takens embedding theorem provides the theoretical foundation for the analysis of time series generated by nonlinear deterministic dynamical systems as well as for the majority of numerical methods used [88, 89]. The success of this theorem is related to the simplicity of the technique proposed: the delay embedding can be indeed applied to any experimental time series. The delay embedding is implemented in many computational libraries. A
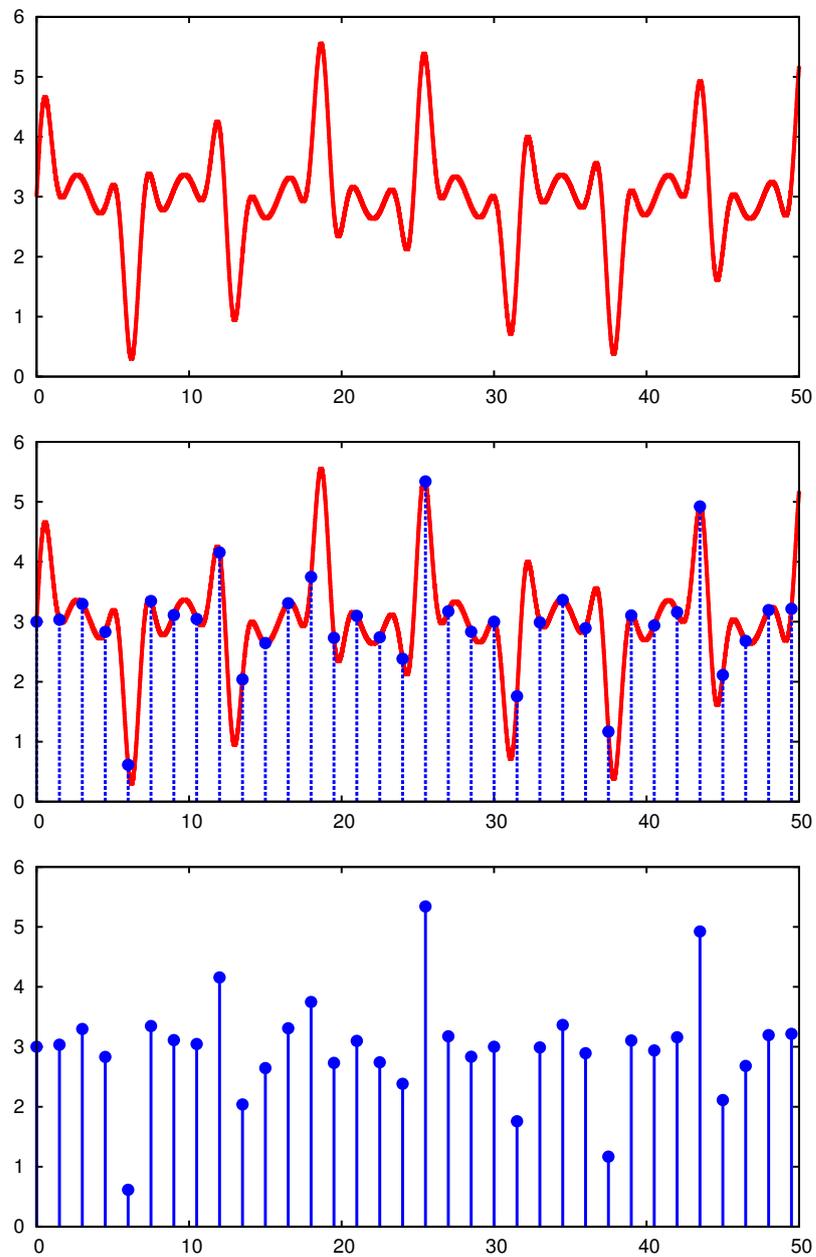
Figure 3.1: On the top, the real continuous signal is plotted with a red line. In the middle, the sampling process is represented; on the bottom the time series recorded is plotted with blue dots.
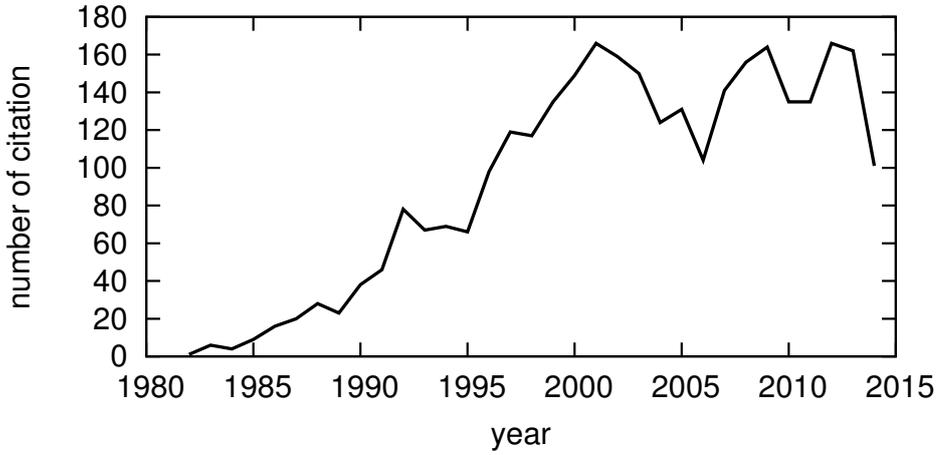
Figure 3.2: Plot of the citation of Takens embedding theorem [6] per year; this data come from the Web of Science databases.

complete software tool on this topic is, for example, TISEAN, developed by Hegger, Kantz and Schreiber [90].

A comprehensive review about Takens embedding theorem is contained in a work by Stark *et al.* [7]. These authors explained how to apply the theorem to experimental data, starting from an ideal situation where only the chaotic dynamics is present, to a real experimental situation in which forced dynamics and stochastic noise are added to the signal of interest [91, 92]. In the following sections we introduce Takens embedding theorem and discuss the major experimental issues linked to this technique.

3.2 Takens embedding theorem

Let m and L (both natural numbers) be the embedding dimension and lag respectively. The delay embedding map $\Psi_{\Phi,\varphi} : M \to \mathbb{R}^m$ is then defined as

$$
\begin{aligned}
\Psi_{\Phi,\varphi}(\mathbf{x}) &\equiv (\varphi(\mathbf{x}), \varphi(\Phi_{L\delta t}(\mathbf{x})), \dots, \varphi(\Phi_{(m-1)L\delta t}(\mathbf{x}))) = \\
&= (\varphi(\mathbf{x}), \varphi(\mathbf{x}_L), \dots, \varphi(\mathbf{x}_{(m-1)L})) \,.
\end{aligned}
\tag{3.2}
$$
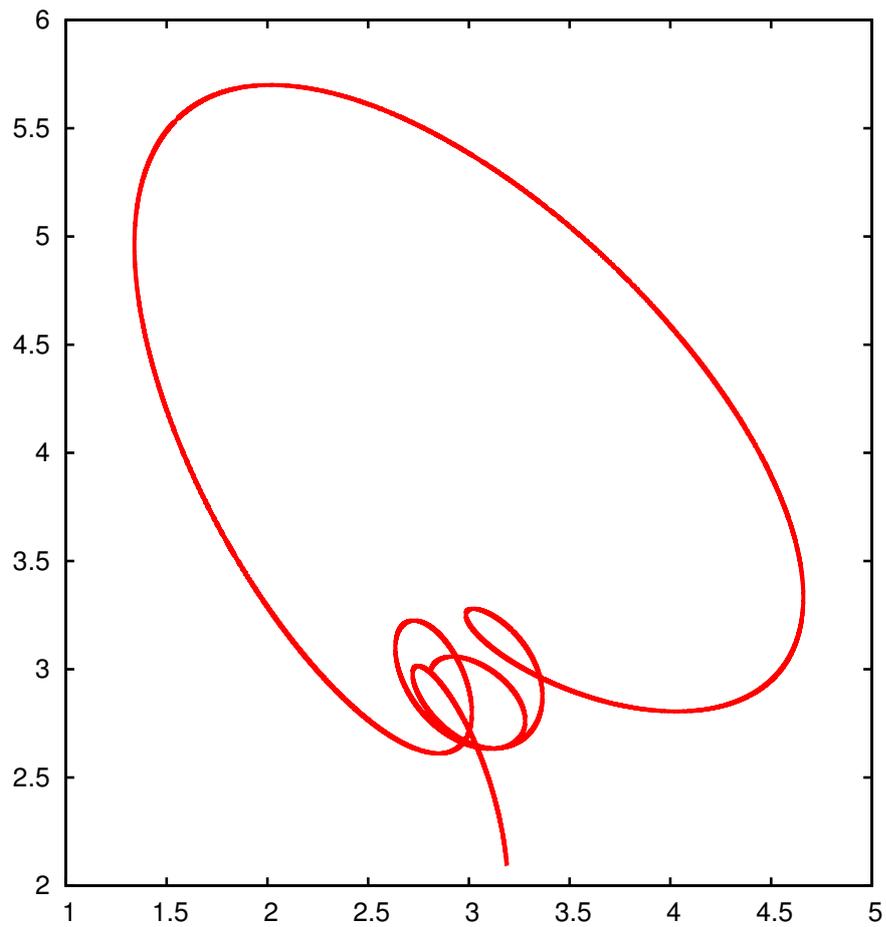
Figure 3.3: Example of embedding $(m, L) = (2, 4)$ of time series in Fig. 3.1.

Upon this definition, Takens embedding theorem can be stated as follows [6, 87, 7]:

Theorem 3.2.1. *Let* M *be a compact* d*-dimensional manifold. If* $m \geqslant 2d + 1$, *the set of* (Φ, φ), *for which the map* $\Psi_{\Phi, \varphi}$ *is an embedding, is open and dense in* $\mathcal{D}^r(M) \times \mathcal{C}^r(M, \mathbb{R})$ *for* $r \geqslant 1$, *where* $\mathcal{D}^r(M)$ *and* $\mathcal{C}^r(M, \mathbb{R})$ *are the set of* $\mathcal{C}^r$ *diffeomorphisms of* M *and the set of observation functions on* M, *respectively.*

To lighten the notation, we henceforth set $\Psi \equiv \Psi_{\Phi, \varphi}$.

The standard Takens theorem implies that, knowing the dimension of the manifold d, for typical Φ and φ, the image $\Psi(M)$ is equivalent to the original manifold M, modulo a smooth invertible coordinate change given by the embedding Ψ. Furthermore, a function $\Theta : \mathbb{R}^m \to \mathbb{R}^m$ can be defined as $\Theta \equiv \Psi \circ \Phi \circ \Psi^{-1}$, as explained in the following scheme:

$$
\begin{array}{ccc}
M & \quad & M \\
& & \\
& \equiv & \\
& & \\
\mathbb{R}^m \xrightarrow{\ \Theta\ } \mathbb{R}^m & &
\end{array}
\qquad
\begin{array}{ccc}
M & \xrightarrow{\ \Phi\ } & M \\
\Psi^{-1} \uparrow & & \downarrow \Psi \\
\mathbb{R}^m & & \mathbb{R}^m
\end{array}
$$

In other words, Θ and Φ have the same dynamical properties. A crucial corollary is that the system's invariants are the same, no matter which representation is used to describe or to infer them. The set of invariants includes the numbers and topological types of fixed points, periodic orbits, the correlation dimension and the Lyapunov exponents. So, for example, we can measure the maximum Lyapunov exponents of Θ and be certain that these are the same as the original system description Φ.

One might argue that achieving a complete knowledge of Θ is by no means easier than achieving a complete knowledge of Φ. However, Θ can be estimated by observing finite sequences of the time series $\{\varphi_n\}$. So, according to the standard Takens theorem, studying these sequences can lead to an assessment of the invariants of the dynamical system.

3.2.1 Time evolution of the delay embedded time series

Let $\mathbf{z}_n \in \mathbb{R}^m$ be an delay embedding point at time n: $\mathbf{z}_n = (\varphi_n, \ldots, \varphi_{n+m-1})$. We have:

$$
\begin{aligned}
\mathbf{z}_n &= (\varphi(\Phi_n(\mathbf{x}_0)), \ldots, \varphi(\Phi_{n+m-1}(\mathbf{x}_0))) \\
&= (\varphi(\mathbf{x}_n), \varphi(\Phi(\mathbf{x}_n)), \ldots, \varphi(\Phi_{m-1}(\mathbf{x}_n))) \\
&= \Psi(\mathbf{x}_n)
\end{aligned}
$$

By applying Θ to $\mathbf{z}_n$ we obtain:

$$
\begin{aligned}
\Theta(\mathbf{z}_n) &= \Psi \circ \Phi \circ \Psi^{-1}(\mathbf{z}_n) \\
&= \Psi \circ \Phi \circ \Psi^{-1}(\Psi(\mathbf{x}_n)) \\
&= \Psi \circ \Phi(\mathbf{x}_n) \\
&= \Psi(\mathbf{x}_{n+1}) \\
&= \mathbf{z}_{n+1}
\end{aligned}
\tag{3.3}
$$

Thus

$$
\Theta(\varphi_n, \ldots, \varphi_{n+m-1}) = (\varphi_{n+1}, \ldots, \varphi_{n+m}).
$$

According to this last expression, the dynamics of Θ consists of "shifting" by one time step along the ordered dimensions of the m-dimensional embedding space a number m of successive elements of the time series $\{\varphi_n\}$. It is worth noting that, while the original dynamics is possibly unobservable, the dynamics in terms of Θ and $\mathbf{z}_n$ is expressible in terms of the observed time series $\{\varphi_n\}$.

Eq. (3.3) implies that the time series $\{\varphi_n\}$ is deterministic: by iterating Θ a number k of times, the future value φ_{n+k} can be predicted starting from φ_n. If Φ, and hence Θ, is chaotic or noisy there is an upper limit on how precisely and far into the future we can predict.

Thus, Takens embedding theorem allows us to reconstruct the unknown dynamical system by simply recording a sequence and constructing a new state space by suitably combining – i.e. embedding – successive observations

of the system. Putting this procedure into operation requires the overcoming of many issues. A major one is the fact that the theorem does not provide any information about the embedding dimension m that we should use if we do not know the dimension of the manifold d. Another problem is the choice of the sampling time δt (and consequently of the lag L), which can significantly affect the performance of the techniques used to reconstruct the dynamics.

In addition, Takens embedding theorem requires a deterministic and autonomous dynamics (a system is said to be autonomous if it is isolated with respect to other systems). Both these requirements are not satisfied in many experimental systems. Therefore, it is important to study the generalizations of Takens embedding theorem applicable to forced systems and to stochastic ones [7, 93, 94]. In the following section, we introduce the generalization of Takens embedding theorem in the case of deterministic forcing systems and in the case of stochastic systems.

3.2.2   Deterministically forced system

Stark [93] generalized the Takens embedding theorem to dynamical system forced by a deterministic system; this new theorem is called *Forced Takens Theorem*. This generalization is very important to reduce the gap between the ideal, mathematical world and the experimental applications.

Besides the dynamical system described by a state $\mathbf{x}_i$ and the discrete time evolution given by Eq. (3.1), let us consider an additional, "forcing" deterministic system

$$\Upsilon : N \to N \, ,$$

where $N \subset \mathbb{R}^k$ is the manifold of the forcing dynamics. Similarly to Eq. (3.1), the discrete evolution of a state $\mathbf{y}_i$ is given by

$$\mathbf{y}_{i+1} = \Upsilon(\mathbf{y}_i) \, .$$

Assuming now a "feeding" (one-way coupling) of the second dynamics into the first one, we get a so-called deterministically forced system. The overall

dynamics, which occurs within a new phase space contained in $\mathbb{R}^d \times \mathbb{R}^k$ is given by:

$$
\begin{aligned}
\mathbf{x}_{i+1} &= \Phi(\mathbf{x}_i, \mathbf{y}_i), \\
\mathbf{y}_{i+1} &= \Upsilon(\mathbf{y}_i).
\end{aligned}
\tag{3.4}
$$

The *Forced Takens Theorem* allows us to reconstruct the joint dynamics of a forcing and a forced system from a time series of an observable of the forced system alone. This theorem is applicable to cases in which the forcing dynamics is unknown; such cases correspond to many practical applications.

On the other hand, if the forcing dynamics is known – ex. gr. in many experimental situations, the experimenter has an independent knowledge of the state of the forcing system – it is not necessary to reconstruct it out of measurements carried out on the forced system. Consequently, it is possible to simplify the reconstruction of the dynamical system without wasting the resources necessary to reconstruct also the forcing dynamics Υ.

The *Forced Takens Theorem* provides the justification for the majority of numerical methods used in experimental systems and gives a theoretical background to popular techniques like the lock-in amplifier. The theorem can be for example used to analyse signals stemming from plasma discharge experiments and electroencephalographic (EEG) recordings: in plasma discharge the gas system is stimulated by using a "known dynamics" made of pulses or other pre-set waveforms; in EEG recordings, forcing dynamics can be represented by neural signals arising from the cardiac system, from the eye movements and, more generally, from other voluntary/involuntary sources of artefacts.

3.2.3 Stochastically forced system

Stark *et al.* [7, 94] generalized the Takens embedding theorem also for the situation in which the dynamics is forced by a stochastic systems. This formulation of Takens Theorem is referred to as *Takens theorem for stochastic systems*. The framework is typical of all real systems: the noise is everywhere

and typically influences the dynamics of a system. Similarly to Eq. (3.4), the discrete time evolution of a dynamical system forced by a stochastic system can be written as follows:

$$\mathbf{x}_{i+1} = \Phi(\mathbf{x}_i, \omega_i),$$

where $\mathbf{x}_i \in \mathbb{R}^d$ and ω_i is the noise component at step i. This formalism includes also the case of a deterministic system Φ subject to "dynamical" noise, namely $\Phi = f(\mathbf{x}_i) + \omega_i$.

The Takens theorem for stochastic systems states that $\Psi_{\Phi, \varphi, \omega}$ is an embedding. Consequently, the dynamics of the original system can be reconstructed, or better estimated, with a degree of uncertainty given by the size of ω.

Stark *et al.* [94] generalized the Takens theorem for stochastic systems also when the noise is of "additive" kind, a situation that typically occurs in a measurement process. Therefore, the embedding delay technique can be used in any experimental situation.

3.3   Optimal embedding parameters

The main drawback of Takens embedding theorem is the fact that it gives no clues on how to choose the embedding parameter m and L for a "experimental" time series. Because the choice of the embedding parameters is crucial to the subsequent analysis, many embedding methods were developed to tackle the issue of determining the optimal embedding parameters. The most used and cited methods are: Gao–Zheng's method [8, 9, 10] (Sec. 6.1), Schuster's method [11] (see Sec. 6.2), the method of characteristic length, the method of global false-nearest-neighbours and autocorrelation function, the method of global false-nearest-neighbours and mutual information [12, 13]. Each motivates the choice of (m, L) by addressing one or more particular features of the dynamics of the system. A review of these methods is given by Cellucci *et al.* [14].

To compare embedding methods we must define a measure of the quality

of the embedding parameter. In 1992, Buzug *et al.* [95] proposed to use the deviation of the estimation of the correlation dimension of the reconstructed manifold from the correlation dimension of the original manifold. Cellucci *et al.* [14] compared delay embedding methods by analysing two dynamical properties: the maximum Lyapunov exponent (MLE) and the robustness to noise. Robustness to noise is determined in terms of impact on the cumulative distribution of interpoint distances in the embedding space. However, the embedding parameters (m, L) are found to vary when noise is added to the time series and it is not clear whether a wrong embedding choice produces a lesser degree of robustness.

The determination of the MLE provides a more decisive way to rank the quality of different embedding methods. As discussed in the next chapter (Sec. 4.1), the maximum Lyapunov exponent of a dynamical system $\mathcal{S}$ can be evaluated independently from the embedding parameters, by using the so-called *standard method* [15, 16] $\chi_{\mathcal{S}}$. Given an (m, L)-embedded time series, i.e., embedded by using a particular parameter pair (m, L), the MLE can be estimated by determining the local divergence rate of nearby trajectories [8, 96, 97], $\Gamma_{\mathcal{S}}(m, L)$. So, according to Cellucci *et al.*, method A is considered to be superior to method B if $\Gamma_{\mathcal{S}}(m_A, L_A)$, calculated by using the embedding pair assessed via A, approaches $\chi_{\mathcal{S}}$ better than $\Gamma_{\mathcal{S}}(m_B, L_B)$, calculated by using the embedding pair assessed via B [17].

In Chapter 5 we discuss the comparison method proposed by Cellucci *et al.* by analysing the statistical error of $\Gamma_{\mathcal{S}}(m, L)$ [17]. We will point out that not all the chaotic dynamical systems provide good benchmarks to assess the reliability of different embedding methods [98, 17].

Chapter 4

Characteristics of a chaotic dynamical system

In this chapter we introduce the main tools to study a chaotic dynamical system. Sec. 4.1 deals with the Lyapunov characteristic exponents of a dynamical system; in addition, the *standard method* used to calculate the exponents and the *divergent rate method* used to calculate the maximum Lyapunov exponent out of a time series are discussed. In Sec. 4.1.2 we present our implementation of the *divergent rate method* to automatize the procedure and to calculate not only the maximum Lyapunov exponent but also its standard error. The dimension of an attractor is presented in Sec. 4.2; four methods to estimate it are presented, namely the box counting, the information dimension, the correlation dimension and the Kaplan–Yorke conjecture. Another important characteristic of a chaotic attractor is entropy, which is the topic of Sec. 4.3; this section also contains a discussion on the relation between the Lyapunov characteristic exponents and the Kolmogorov entropy.

4.1 Lyapunov characteristic exponents

The Russian mathematician Alexandr Lyapunov, in his PhD Thesis in 1892, introduced the concept of Lyapunov exponents [99] to measure the stability of a dynamical system by averaging the rate of growth of small

perturbations to the orbits.

The Lyapunov characteristic exponents (LCEs) provides a quantitative characterization of the dynamics: for a d-dimensional dynamical system, there are d Lyapunov exponents. They measures the average rate at which a volume element in the phase space expands or contracts along its trajectory. A positive maximum Lyapunov exponent (MLE) characterizes an exponentially-growing divergence of orbits, and is linked to chaotic dynamics; a negative MLE characterizes a stable motion, because nearby trajectories converge.

Let us consider the evolution of two states x, y of a scalar continuous dynamical system. According to Eq. (2.2) and considering the divergence $\Delta \equiv y - x$ it follows:

$$\Delta(dt) = \Delta(0) \left(1 + \frac{f[y(0)] - f[x(0)]}{\Delta(0)} dt \right) .$$

By iterating the expression above for a time $t = n \cdot dt$, assuming the coefficient of dt to have a slower evolution than x, y, and averaging, we get

$$\Delta(t) \approx \Delta(0) e^{\lambda t} ,$$

where λ is a time-independent constant.

This argument can be expanded to systems whose evolution occurs in multidimensional spaces. So, heuristically, two trajectories in phase space with initial separation $\delta Z(0)$ diverge as (see also Fig. 4.1):

$$|\delta Z(t)| \approx e^{\lambda t} |\delta Z(0))| .$$

The constant λ is referred to as the *maximum Lyapunov exponent* (MLE).

More precisely, the MLE is defined as follows:

$$\lambda_{max} = \lim_{t \to \infty} \lim_{\delta Z(0) \to 0} \left\langle \frac{1}{t} \ln \frac{|\delta Z(t)|}{|\delta Z(0)|} \right\rangle ,$$

where the average is taken on the set of pairs of initially nearby trajectories,
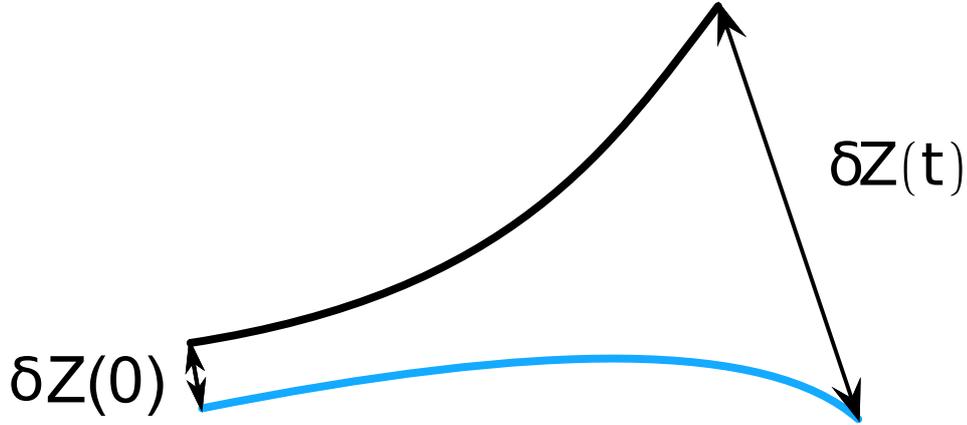
Figure 4.1: Evolution of two initially nearby orbits.

whose separation is, in the limit, vanishing.

The MLE is a measure linked to the sensitive dependence to initial conditions. It was introduced by Hénon and Heiles [100] and was further used by several authors (ex. gr. , [101, 102]). In the study by Hénon and Heiles, two initial points were chosen "very close" to each other, so that their phase space separation was about 10^{-7}–10^{-6}. They found that, in the case of a chaotic system, the distance exhibits an exponential increase in time.

As an example, Fig. 4.2 shows the evolution of the distance between two nearby trajectories of the reference Lorenz attractor. This system is described by Eq. (2.5), with the set of parameters expressed in Sec. 2.5. The trajectories start at the points

$$\mathbf{a} = (5.84613084500473, 11.2335817205288, 25.7823621037852) ,$$
$$\mathbf{b} = (5.84613084500472, 11.2335817205288, 25.7823621037852) ;$$

the difference only regards the last digit of the x-coordinate and is equal

to 10^{-14}. For $t \leqslant 22$ the distance increases exponentially (linearly in the logarithmic plot). The angular coefficient of the straight line fitting the points – a quantity related to the MLE – is in this example 1.22(1). This value coincides, within the uncertainty, with the MLE for the reference Lorenz attractor, namely 1.2346(6) (see also Tab. 4.1).
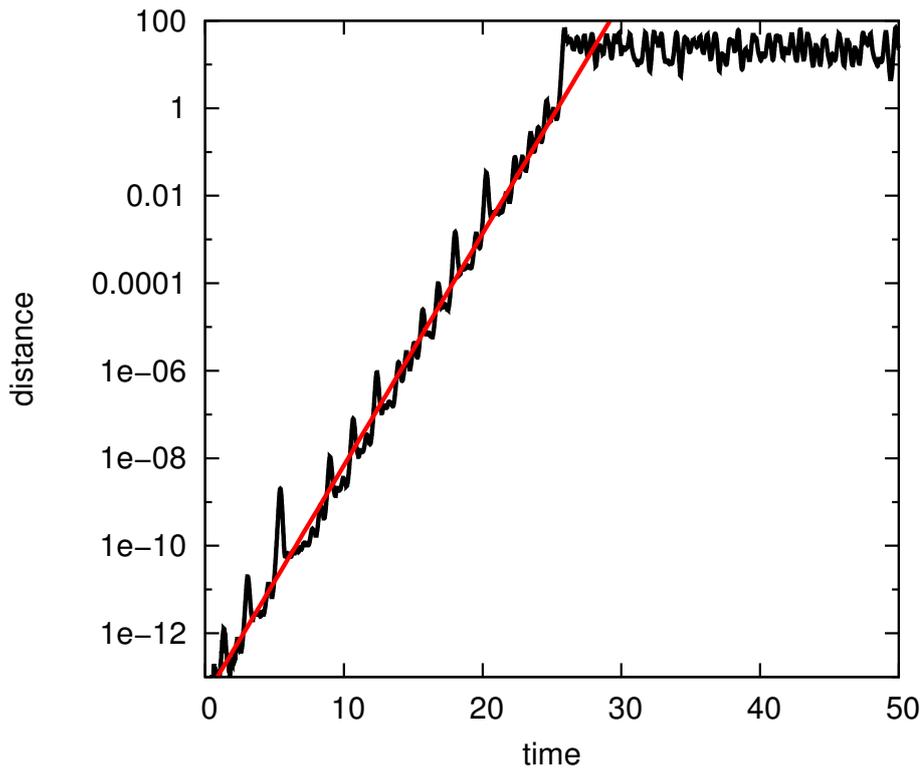
Figure 4.2: Evolution of the distance between two nearby trajectories generated by using Eq. (2.5) with $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 45.92$. The two trajectories start from points that differ by $1 \cdot 10^{-14}$ on the x direction. The slope of the linear red line ($t \leqslant 22$), is equal to 1.22(1).

A dynamical system "living" in a d-dimensional manifold is characterized by a set of d Lyapunov exponents. To assess these exponents, several methods have been developed: the *standard method* [15, 16], the singular value decomposition and the QR decomposition procedures [27].

In the experimental analysis, the most important exponent is the MLE because it represents the main behaviour of dynamical systems. Starting

from a time series, the MLE can be evaluated by means of the *divergent rate method* [8, 96, 97]. A comprehensive review on this method is contained in the work by Kantz *et al.* [103].

4.1.1 The standard method

The *standard method* was developed by Benettin *et al.* [15, 16, 104] in 1980 and it is considered a mainstay of the LCEs computation [105, 106, 14, 27].

Considering a continuous-time dynamical system, whose evolution is described by Eq. (2.2), the corresponding evolution of a deviation vector **s** (also often referred to as δ**x**) is given by its tangent map:

$$\frac{d\mathbf{s}}{dt} = \mathbf{J},\qquad(4.1)$$

where **J** is the Jacobian matrix $J_{ij} = \frac{\partial f_i}{\partial x_j}$.

As an example, in the case of the Lorenz attractor (see Eq. (2.5)), we have:

$$
\begin{aligned}
\frac{ds_x}{dt} &= \sigma(s_y - s_x),\\
\frac{ds_y}{dt} &= s_x(\rho - z) - x s_z - s_y,\\
\frac{ds_z}{dt} &= s_x y + x s_y - \beta s_z.
\end{aligned}
\qquad(4.2)
$$

At time t_1, let us consider a p-dimensional parallelepiped whose edges are given by p linearly independent deviation vectors, namely $\mathbf{s}_1(t_1)$, $\mathbf{s}_2(t_1)$, ... , $\mathbf{s}_p(t_1)$. Let $\mathrm{vol}_p(t_1)$ be volume of this parallelepiped. Furthermore, let $\mathbf{s}_i(t_2)$ be the i-th deviation vector ($i = 1, \ldots, p$) that results from the evolution till time t_2 of the vector $\mathbf{s}_i(t_1)$. The vector $\mathbf{s}_i(t_2)$ can be evaluated by (numerically) integrating Eq. (4.1). Let $\mathrm{vol}_p(t_2)$ be the volume of the new parallelepiped.

Besides the definitions above, it is worth introducing three additional, important definitions. First, let the expansion $\eta(t_1 \to t_2)$ be defined as

$$\eta(t_1 \to t_2) \equiv \frac{\mathrm{vol}_p(t_2)}{\mathrm{vol}_p(t_1)}.$$

Second, let $X_p(t)$ be defined as

$$X_p(t) \equiv \frac{1}{t} \ln \eta(0 \to t) \, . \tag{4.3}$$

Third, given the p greater Lyapunov exponents χ_1 (i.e. the MLE), χ_2, …, χ_p, let p-MLE be their sum:

$$\text{p-MLE} \equiv \sum_{i=1}^{p} \chi_i \, .$$

A crucial theorem proved by Benettin, Galgani, Giorgilli, and Strelcyn in 1980 [15] states that

$$\text{p-MLE} = \lim_{t \to \infty} X_p(t) \, . \tag{4.4}$$

This theorem can be exploited to iteratively compute the LCEs: setting $p = 1$ directly provides the MLE; the second largest Lyapunov exponent is then evaluated by i) setting $p = 2$, ii) assessing 2-MLE, and iii) subtracting the MLE; and so on. However, as shown below, the computation can be considerably, and "free of charge", shortened.

The volume vol_p of a parallelepiped whose edges are $\mathbf{s}_1$, …, $\mathbf{s}_p$, can be computed by exploiting the Gram–Schmidt orthonormalization method. Starting from the set $\mathbf{s}_1$, …, $\mathbf{s}_p$, a new set of orthonormal vectors $\hat{\mathbf{s}}_1$, …, $\hat{\mathbf{s}}_p$ is constructed as follows:

$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{s}_1, \\
\gamma_1 &= \|\mathbf{u}_1\|, \\
\hat{\mathbf{s}}_1 &= \frac{\mathbf{u}_1}{\gamma_1}, \\
\mathbf{u}_k &= \mathbf{s}_k - \sum_{j=1}^{k-1} \langle \mathbf{s}_k, \hat{\mathbf{s}}_j \rangle \, \hat{\mathbf{s}}_j \, , \\
\gamma_{ki} &= \|\mathbf{u}_k\|, \\
\hat{\mathbf{s}}_k &= \frac{\mathbf{u}_k}{\gamma_k},
\end{aligned}
\tag{4.5}
$$

for $1 < k \leqslant p$. In the previous expressions, $\langle \mathbf{a}, \mathbf{b} \rangle$ is the scalar product of the

vectors **a** and **b**. The volume vol_p of the parallelepiped is then equal to the volume of the new, p-dimensional box having the vectors $\mathbf{u}_j$ as edges. This last volume is simply given by the product of the norms of the vectors that represent the new edges:

$$\mathrm{vol}_p = \prod_{i=1}^{p} \gamma_i \, .$$

By means of this last expression and by taking the starting set $\mathbf{s}_1(0)$, $\mathbf{s}_2(0), \ldots , \mathbf{s}_p(0)$ to be orthonormal, $X_p(t)$ (see Eq. (4.3)) can be re-written as follows:

$$X_p(t) = \frac{1}{t} \sum_{i=1}^{p} \ln \gamma_i \, . \tag{4.6}$$

In principle, the next step should consist in taking the limit $t \to \infty$ – i.e., from the computational point of view, taking the number of Runge-Kutta integration steps $k = t/\delta t$ sufficiently large – and then applying Eq. (4.4). However, if the system under study is chaotic, there are two main issues to cope with:

- the norm of the deviation vectors $\|\mathbf{s}_i(t)\|$ increases exponentially with time t, leading to numerical overflow;

- when at least two deviation vectors are considered, the angles between their directions become too small to numerically compute.

These difficulties can be overcome by a powerful lemma [16]: the expansion of any p-dimensional parallelepiped under the action of a linear map (as it is the case of the tangent map, Eq. (4.1); see, for example, Eq. (4.2)) does not depend on the initial volume.

Consequently, let the evolution from step 0 to step t be partitioned in cycles of length τ, henceforth referred to as "renormalization period", and ending with a Gram-Schmidt orthonormalization. Because of the lemma, Eq. (4.3) can be rewritten as:

$$X_p(t) = \frac{1}{t} \sum_{n=1}^{t/\tau} \ln \eta \left[(n-1)\tau \to n\tau \right] \, . \tag{4.7}$$

However, because the starting set of deviation vectors is supposed to be orthonormal and the Gram-Schmidt orthonormalization delivers, at each cycle, an orthonormal set, the expansion $\eta\left[(n-1)\tau \to n\tau\right]$ at the n-th cycle can be evaluated by means of Eq. (4.6), as follows:

$$\eta\left[(n-1)\tau \to n\tau\right] = \prod_{i=1}^{p} \gamma_{n,i}. \tag{4.8}$$

The period τ has to be large enough to observe the effects of time evolution – namely, a "distortion" of the starting parallelepiped – but small enough so that the two numerical issues expressed above do not occur. Typically, $\tau \approx 100\,\delta t$, i.e., every $\tau/\delta t \approx 100$ Runge-Kutta steps, the Gram-Schmidt orthonormalization is carried out and the related expansion evaluated. Again, the resulting orthonormal set of vectors serves as a starting point for a new run of $\tau/\delta t \approx 100$ Runge-Kutta steps, and so on for typically $t/\tau \approx 10^5$ cycles.

By merging Eq. (4.7) and Eq. (4.8), we get:

$$X_p(t) = \frac{1}{t} \sum_{n=1}^{t/\tau} \sum_{i=1}^{p} \ln \gamma_{n,i} = \frac{1}{t} \sum_{i=1}^{p} \sum_{n=1}^{t/\tau} \ln \gamma_{n,i}.$$

Thus, the theorem of Eq. (4.4) becomes:

$$\sum_{i=1}^{p} \chi_i = \lim_{t \to \infty} \frac{1}{t} \sum_{n=1}^{t/\tau} \sum_{i=1}^{p} \ln \gamma_{n,i} = \sum_{i=1}^{p} \left[\lim_{t \to \infty} \frac{1}{t} \sum_{n=1}^{t/\tau} \ln \gamma_{n,i} \right],$$

so that,

$$\chi_p = \left(\sum_{i=1}^{p} \chi_i \right) - \left(\sum_{i=1}^{p-1} \chi_i \right) = \lim_{t \to \infty} \frac{1}{t} \sum_{n=1}^{t/\tau} \ln \gamma_{n,p}. \tag{4.9}$$

The Lyapunov exponents calculated using the standard method satisfy the inequalities $\chi_1 \geqslant \chi_2 \geqslant \cdots \geqslant \chi_p$. Thus, the MLE corresponds to χ_1. Fig. 4.3 illustrates the computing of the two largest LCEs χ_1, χ_2 according to the standard method and by means of the Gram–Schmidt orthonormalization method.
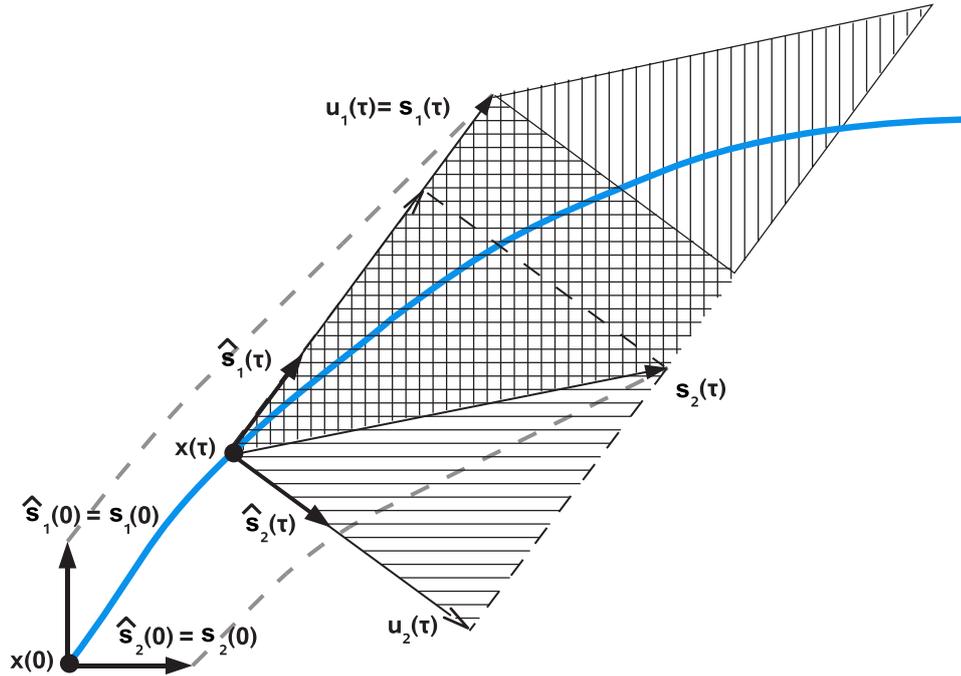
Figure 4.3: Computation of the two largest LCEs χ_1, χ_2 according to the standard method. The deviation vectors $\mathbf{s}_1(0)$, $\mathbf{s}_2(0)$ are evolved according to Eq. (4.1) for a time τ. The evolved vectors $\mathbf{s}_1(\tau)$ and $\mathbf{s}_2(\tau)$ undergo a Gram–Schmidt procedure yielding a set of orthonormal vectors $\hat{\mathbf{s}}_1(\tau)$ and $\hat{\mathbf{s}}_2(\tau)$. These last vectors are again evolved and the same procedure is iteratively applied. For each cycle, the quantities $\gamma_1 = \|\mathbf{u}_1\|$, $\gamma_2 = \|\mathbf{u}_2\|$ are computed and χ_1, χ_2 are estimated from 4.9. Source: [27].

Implementation

We implemented the standard method to calculate the LCEs by using a code written in C++. In this section we report part of the developed code, both to further illustrate how Benettin's method works and to provide an example of the extensive programming activity that was carried out in this thesis work. To integrate the differential equations that describe the evolution of the dynamical system and the deviation vectors, we use the 8th-order Runge–Kutta algorithm discussed in Sec. 2.4. The steps to perform the calculation of the LCEs are:

1. The renormalization period τ and the number of renormalizations $N = t/\tau$ are set.

2. The dynamical system $\mathbf{x}(0)$ and the deviation vectors $\mathbf{s}_1(0), \dots, \mathbf{s}_p(0)$ are initialized. The initial state $\mathbf{x}(0)$ is randomly set. The deviation vectors are set according to the rule $s_{k,i}(0) = \delta_{ki}$ (the Kronecker delta); so, for example, if $p = 3$, $\mathbf{s}_1(0) = (1, 0, 0)$, $\mathbf{s}_2(0) = (0, 1, 0)$ and $\mathbf{s}_3(0) = (0, 0, 1)$.

3. The following steps 4. and 5. are repeated N-times.

4. The dynamical system and the deviation vectors are evolved for a renormalization period τ. In the following box we reported the differential equations for the Lorenz attractor and its tangent map; lines 9 to 11 contain the implementation of Eq. (2.5); lines 14 to 18 contain the implementation of Eq. (4.2).

```
1    int lorenz (double t, const double r[], double f[], void *params) {
         static int i, Dim, dim;
         mackeyGlassParameters* mgpPtr = (mackeyGlassParameters*) params;
         vector<double> a;
         Dim = mgpPtr->rowDim;
6        dim = mgpPtr->colDim;
         a.clear();
         a = mgpPtr->a;

         f[X] = a[0] * (r[Y] - r[X]) + a[4];          // dx/dt = sigma(y-x)   a[0] = sigma
11       f[Y] = r[X] * (a[2] - r[Z]) - r[Y] + a[5];   // dy/dt = x*(rho-z)-y  a[2] = rho
         f[Z] = r[X] * r[Y] - a[1] * r[Z] + a[6];     // dz/dt = x*y-beta*z   a[1] = beta

         for (i = 1; i < Dim; i++) {
             f[SX] = a[0] * (r[SY] - r[SX]);                       // dsx/dt = sigma(sy-sx)
16           f[SY] = r[SX] * (a[2] - r[Z]) - r[X] * r[SZ] - r[SY]; // dsy/dt = sx*(rho-z)-x*sz-sy
             f[SZ] = r[SX] * r[Y] + r[X] * r[SY] - a[1] * r[SZ];   // dsz/dt = sx*y+x*sy-beta*sz
         }
         return GSL_SUCCESS;
    }
```

5. The Gram–Schmidt orthonormalization (see Eq. (4.5)) is performed: the deviation vectors fed to this procedure are replaced with the new orthonormal set; the factors $\gamma_{n,p}$ are stored for further processing. The code used to implement the Gram–Schmidt orthonormalization is the following:

```
void gram_schmidt(vector< vector<double> >&r, vector<double>&lyap) {
    int i, j;
```

```
     double x;
     for(i = 1; i <= lyap.size(); i++) {
5        for(j = 1 ; j < i; j++) {
             x = inner_product(r[i].begin(), r[i].end(), r[j].begin(), 0.0);
             substract_vector(r[i].begin(), r[i].end(), r[j].begin(), x);
         }
         x = sqrt(inner_product(r[i].begin(), r[i].end(), r[i].begin(), 0.0));
10       normalize(r[i].begin(), r[i].end(), x);
         lyap[i-1] += log(x);
     }
 }
```

6. The values of the Lyapunov exponents are obtained by using Eq. (4.9).

Given a system $\mathcal{S}$, we assume the MLE calculated by means of the standard method to be the reference value – henceforth referred to as $\chi_\mathcal{S}$ – for the maximum Lyapunov exponent of that system. The values of $\chi_\mathcal{S}$ for the four reference dynamical systems described in Sec. 2.5 are reported in Tab. 4.1. Each value corresponds to the sample mean and the sample standard deviation of the results of 10 runs, each made of 10^7 integration steps.

Table 4.1: Standard value $\chi_\mathcal{S}$ of the MLE, calculated for each of the four dynamical systems of Sec. 2.5 by means of the standard method. The digit in parentheses corresponds to the uncertainty $\sigma_{\chi,\mathcal{S}}$ on the least significant digit.

Dynamical system ($\mathcal{S}$)	$\chi_\mathcal{S}$
Hénon	0.41924(9)
Lorenz	1.2346(6)
Rössler	0.08889(9)
Mackey–Glass	0.00742(2)

4.1.2 Divergent rate method

In 1993-1994, Gao and Zheng [8], Rosenstein *et al.* [96], and Kantz *et al.* [97] independently proposed the same method to evaluate the MLE from an experimental time series. The method is referred to as *divergent rate method* [103] and is based on the embedding of the time series. In this

section, we present the method and its implementation as discussed in the paper "Statistical properties of the maximum Lyapunov exponent calculated via divergence rate method" [17], published by our research group.

Given a sample time sequence $\{x_n\}$ of a scalar, real variable x, we consider the m-dimensional embedding $\mathbf{X}_i \equiv \Psi(x_i)$ as defined in Sec. 3.2 (see Eq. (3.2)); the embedding lag is L. We then consider a set of N randomly-selected (without replacement) pairs $\mathbf{X}_i$, $\mathbf{X}_j$. Each $\mathbf{X}_i$, $\mathbf{X}_j$ pair must satisfy the two following conditions [14]:

i. $\|\mathbf{X}_i - \mathbf{X}_j\| \leqslant r$, with r corresponding to a given percentile p of the distribution of all Euclidean distances $\|\mathbf{X}_i - \mathbf{X}_j\|$;

ii. $|i - j| \geqslant c_o$, where c_o is a constraint linked to the autocorrelation function.

In the literature, the constraint c_o is set in different way: Theiler [107, 108] suggests to take c_o equal to the autocorrelation time, whereas Gao and Zheng [8] suggest to take $c_o = (m-1)L$. We decide to merge these two suggestions and set c_o to the first zero of the autocorrelation function multiplied times $m(L + 1)$.

The time-dependent divergence exponent $\Lambda(k)$ is then defined as

$$\Lambda(k) \equiv \left\langle \ln \frac{\|\mathbf{X}_{i+k} - \mathbf{X}_{j+k}\|}{\|\mathbf{X}_i - \mathbf{X}_j\|} \right\rangle , \qquad (4.10)$$

where k is an non-negative, integer delay and the average $\langle \ldots \rangle$ is taken on the N-dimensional set of $\mathbf{X}_i$, $\mathbf{X}_j$ pairs. From the definition it follows $\Lambda(0) = 0$.

An important step in our algorithm is that, rather than considering a single calculation for each value of k, we determined each point of $\Lambda(k)$ and the respective uncertainty $\sigma_\Lambda(k)$ as the point-wise sample mean and sample

standard deviation of a set of M different calculations, respectively:

$$\Lambda(k) \ \equiv \ \frac{1}{M} \sum_{i=1}^{M} \Lambda_i(k) \,, \tag{4.11a}$$

$$\sigma_\Lambda(k) \ \equiv \ \frac{1}{M-1} \sum_{i=1}^{M} [\Lambda_i(k) - \Lambda(k)]^2 \,. \tag{4.11b}$$
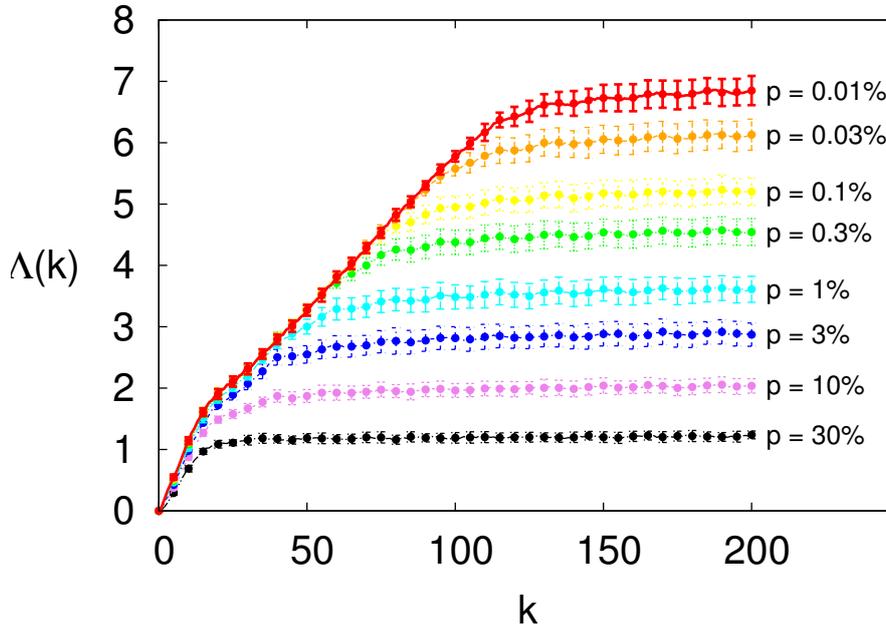
Figure 4.4: Time-dependent divergence exponent $\Lambda(k)$ for the reference Lorenz attractor, obtained by using the embedding pair (3, 3). The different plots correspond to eight different values of p: 0.01%, 0.03%, 0.1%, 0.3%, 1%, 3%, 10%, 30%. For the sake of clarity, only one point every five is plotted.

Fig. 4.4 shows the time-dependent divergence exponent $\Lambda(k)$ for the reference Lorenz attractor and different values of the value p. Typically, there is a k-range within which the dependence of Λ on k turns out to be linearly growing and independent of the boundary conditions. The corresponding value of the MLE is then given by the slope of $\Lambda(k)$ divided by δt ($\delta t = 1$ for maps). The slopes of $\Lambda(k)$ tend to increase for progressively smaller p. The

reason is that the smaller p, the smaller the probability that two close points belong to trajectories that are still approaching to each other.

Finally, to determine the MLE $\Gamma_S(m, L)$ of the system S with embedding (m, L), we implement the following automatic procedure:

1. if $k \to \infty$, $\Lambda(k)$ tends to a constant value that is estimated by averaging the 10 largest available k values; k_{slope} is then determined as the lowest k such that $\Lambda(k) > 0.9\Lambda_\infty$;

2. given the growing section of $\Lambda = \Lambda(k) \mid k \leqslant k_{slope}$, we determine the point P of abscissa k_P that allows for the longest straight-line segment, centered on P, that fits the data with a reduced χ^2 not larger than $1 + \sqrt{8/(2\ell - 1)}$, where 2ℓ is the length of the segment (in other words, a fit is assumed to be valid if the corresponding χ^2 does not exceed its expected value by twice its standard deviation);

3. if more adjacent points satisfy the requirement of step 2., point P is selected by considering the minimum χ^2;

4. to increase accuracy, a final straight-line fit is carried out; the fitting segment is centered in P and has a length equal to $2\ell'$, where $\ell' \equiv \lceil \ell/2 \rceil$; the fit result is taken to be valid if $\ell' \leqslant 2$, i. e. if the new length is at least 4;

5. $\Gamma_S(m, L)$ is set to the straight-line slope divided by the time step δt of the sample time sequence; accordingly, the uncertainty $\sigma_{\Gamma,S}(m, L)$ on $\Gamma_S(m, L)$ is set as the error on the slope divided by δt:

$$\Gamma_S(m, L) = \frac{1}{\delta t} \frac{\overline{k\Lambda(k)} - \overline{k} \cdot \overline{\Lambda(k)}}{\overline{k^2} - \overline{k} \cdot \overline{k}}, \tag{4.12a}$$

$$\sigma_{\Gamma,S}(m, L) = \frac{1}{\delta t} \left[\left(\overline{k^2} - \overline{k} \cdot \overline{k} \right) \cdot \sum_{k=k_P-\ell'}^{k_P+\ell'} \frac{1}{\sigma_\Lambda^2(k)} \right]^{-1/2}, \tag{4.12b}$$

where, for a generic $f(k)$,

$$\overline{f(k)} \equiv \left[\sum_{k=k_P-\ell'}^{k_P+\ell'} \frac{f(k)}{\sigma_\Lambda^2(k)} \right] \cdot \left[\sum_{k=k_P-\ell'}^{k_P+\ell'} \frac{1}{\sigma_\Lambda^2(k)} \right]^{-1} . \quad (4.13)$$

4.1.3 Divergent rate method applied to a stochastic time series

By applying the divergent rate method to a stochastic time series, generated by white noise ($\sigma = 1$), we obtain the time-dependent divergence exponent $\Lambda(k)$ represented in Fig. 4.5. In this example the embedding pair used is $(10, 1)$.

We note that the function $\Lambda = \Lambda(k)$ shows no linear increasing slope. However, as in the case of the Lorenz attractor discussed in the previous section, the "plateau" value, i.e. the constant value Λ_{pl} taken on by $\Lambda(k)$ for $k \geqslant m = 10$, depends on the percentile p of the distribution of the Euclidean distances (see Sec. 4.1.2).

The behaviour of $\Lambda(k)$ in Fig. 4.5 for $k < m = 10$ can be easily explained: given k, and because $L = 1$, a number of $(m - k)$ components of the vector $\mathbf{X}_{i+k}$ coincide with the same number of components of the vector $\mathbf{X}_i$; the same occurs for $\mathbf{X}_{j+k}$ and $\mathbf{X}_j$. For this reason, the lower k the more similar are, on average, the numerator and the denominator of the ratio of Eq. (4.10). Consequently, the average ratio tends to one and its average logarithm $\Lambda(k)$ drops to zero as k decreases from m to 0.

For a stochastic time series, it is possible – and interesting – to estimate the plateau value Λ_{pl}. The distance $\|\mathbf{X}_i - \mathbf{X}_j\|$ can be approximated as follows:

$$\|\mathbf{X}_i - \mathbf{X}_j\| = \left[\sum_{l=0}^{m-1} (x_{i+lL} - x_{j+lL})^2 \right]^{\frac{1}{2}} \approx \left[2\chi_m^2 \right]^{\frac{1}{2}} .$$

For this approximation we assume, first, x_{i+lL}, x_{j+lL} to be independent. In addition we exploit the fact that the difference of two independent, standard normal random variables is normal with variance equal to 2 (the white noise
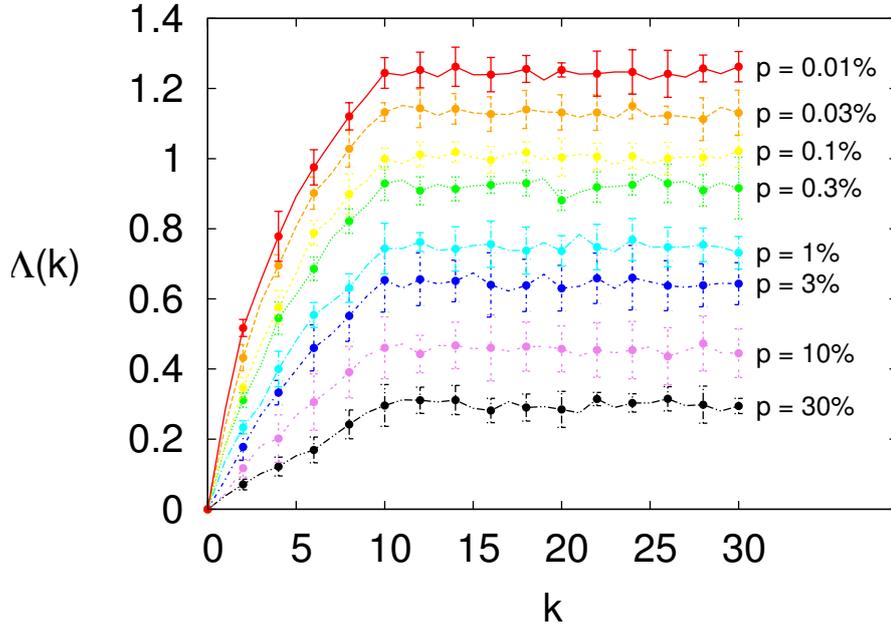
Figure 4.5: Time-dependent divergence exponent $\Lambda(k)$ for a purely noisy time series; the embedding pair is $(10, 1)$. The different plots correspond to eight different values of the percentile p: 0.01%, 0.03%, 0.1%, 0.3%, 1%, 3%, 10%, 30%.

– and thus each x_i – is assumed to have unitary variance), and the fact that the sum of m independent, standard normal random variables is distributed as χ_m^2.

The link between the percentile p and the maximum length r can be assessed by noting that $\|\mathbf{X}_i - \mathbf{X}_j\| \leqslant r$ implies $\chi_m^2 \approx \frac{1}{2}\|\mathbf{X}_i - \mathbf{X}_j\|^2 \leqslant \frac{r^2}{2}$ and writing:

$$\int_0^{\frac{r^2}{2}} f_m\left(\chi_m^2\right) d\chi_m^2 \approx p\,, \tag{4.14}$$

where $f_m\left(\chi_m^2\right)$ is the probability density function of χ^2 with m degrees of freedom.

The average $\langle \ln\|\mathbf{X}_i - \mathbf{X}_j\|\rangle$ can be evaluated by taking into account the probability density function of the squared distance, namely a χ^2 distribution

truncated at $r^2/2$ and suitably re-normalized by a factor p^{-1}:

$$\langle \ln \|\mathbf{X}_i - \mathbf{X}_j\| \rangle \approx \frac{1}{2p} \int_0^{\frac{r^2}{2}} f_m\left(\chi_m^2\right) \ln(2\chi_m^2) d\chi_m^2 \,. \tag{4.15}$$

The average $\langle \ln \|\mathbf{X}_{i+k} - \mathbf{X}_{j+k}\| \rangle$ can be evaluated in similar way:

$$\langle \ln \|\mathbf{X}_{i+k} - \mathbf{X}_{j+k}\| \rangle \approx \frac{1}{2} \int_0^\infty f_m\left(\chi_m^2\right) \ln(2\chi_m^2) d\chi_m^2 \,. \tag{4.16}$$

In the previous expression the upper integration limit is ∞ as the k-step-evolution "removes" from the pair $\mathbf{X}_{i+k}, \mathbf{X}_{j+k}$ the effects of the constraint on the starting vectors $\mathbf{X}_i, \mathbf{X}_j$ (and no re-normalization is required).

The plateau value Λ_{pl} is given by the difference of Eq. (4.16) and Eq. (4.15):

$$\Lambda_{pl}(m, p) \approx \frac{1}{2} \int_0^\infty f_m\left(\chi_m^2\right) \ln(2\chi_m^2) d\chi_m^2 - \frac{1}{2p} \int_0^{\frac{r^2}{2}} f_m\left(\chi_m^2\right) \ln(2\chi_m^2) d\chi_m^2 \,. \tag{4.17}$$

By knowing that, for a positive x,

$$f_m\left(x\right) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} e^{-\frac{x}{2}} x^{\frac{k}{2}-1} \,, \tag{4.18}$$

and integrating by parts, we get the following expression for the leftmost integral of Eq. (4.17):

$$L_m \equiv \int_0^\infty f_m\left(\chi_m^2\right) \ln(2\chi_m^2) d\chi_m^2 \,,$$

$$\Downarrow$$

$$L_m = \frac{1}{m-2} + L_{m-2} = \frac{1}{m-2} + \frac{1}{m-4} + \cdots + \begin{cases} \frac{1}{3} + 1 + L_1 & \text{if } n \text{ is odd} \\ \frac{1}{4} + \frac{1}{2} + L_2 & \text{if } n \text{ is even} \end{cases} \,,$$

where

$$L_1 = -\frac{\gamma}{2} \,,$$
$$L_2 = -\frac{\gamma}{2} + \log 2 \,,$$

and $\gamma = 0.57721\,56649\ldots$ is the Eulero Mascheroni constant given by

$$\gamma \equiv \lim_{n\to\infty}\left[\left(\sum_{k=1}^{n}\frac{1}{k}\right) - \log(n)\right].$$

By using the definition of n-th harmonic number,

$$H_n \equiv \sum_{k=1}^{n}\frac{1}{k},$$

the previous expressions for the integral L_m become (for $m > 1$):

$$L_{2m+1} = H_{2m} - \frac{1}{2}H_m - \frac{\gamma}{2},$$
$$L_{2m} = \frac{1}{2}H_{m-1} - \frac{\gamma}{2} + \log 2.$$

Finally, by using the approximation $H_n \gtrsim \gamma + \log(n)$, it is easy to show that

$$L_m \simeq \frac{1}{2}\log 2\,(m-1).$$

If $p \ll 1$, Eq. (4.14) can be easily integrated by setting the exponential to 1 in Eq. (4.18), yielding:

$$\frac{r^2}{2} \approx 2\left[p\Gamma\left(\frac{m}{2}+1\right)\right]^{\frac{2}{m}}.$$

By using the same approximation (exponential set to 1), the rightmost integral of Eq. (4.17) becomes

$$-\frac{1}{2p}\int_0^{\frac{r^2}{2}} f_m\left(\chi_m^2\right)\ln(2\chi_m^2)d\chi_m^2 \approx -\log 2 + \frac{1}{m}\left\{1 - \log\left[p\Gamma\left(\frac{m}{2}+1\right)\right]\right\}.$$

By adding the two integrals, we finally get the following approximation to Eq. (4.17):

$$\Lambda_{pl}(m,p) \approx \frac{1}{2}\log\frac{m-1}{2} + \frac{1}{m}\left\{1 - \log\left[p\Gamma\left(\frac{m}{2}+1\right)\right]\right\}. \tag{4.19}$$

Fig. 4.6 shows the goodness of the argument exposed above in explaining the plateau value in the case of a stochastic time series.
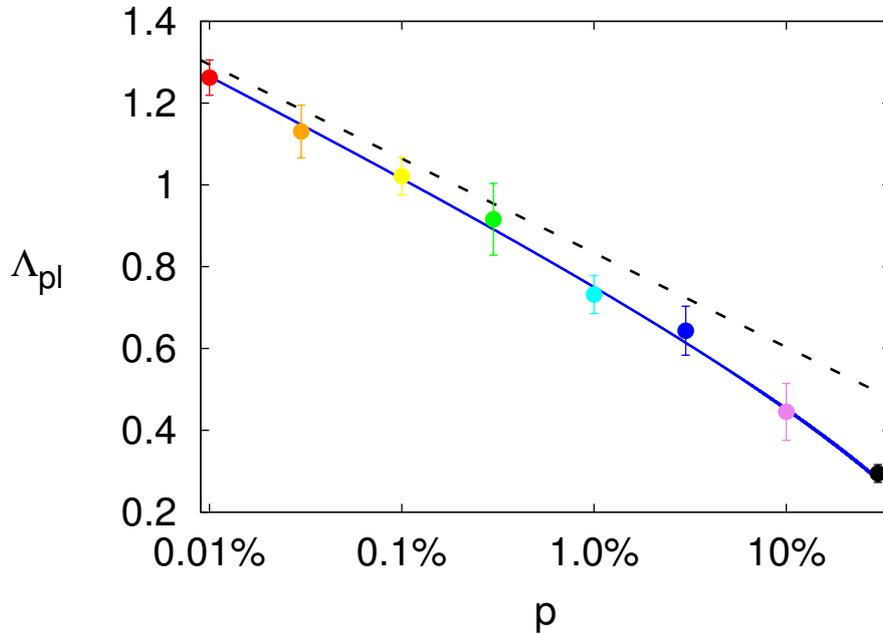
Figure 4.6: Plateau value Λ_{pl} of the divergence exponent as a function of the percentile p for a noise time series. The embedding pair used is $(10, 1)$. The abscissa of each point – and the related colour – corresponds to one of the percentile values reported in Fig. 4.5; the ordinate of each point, and the related errorbar, corresponds to the rightmost point ($k = 30$), with the respective errorbar, in Fig. 4.5. The blue line corresponds to Eq. (4.17), the dashed line to Eq. (4.19).

4.1.4 Distinguishing stochastic noise from other kinds of dynamics in a time series

Fig. 4.7 shows the plateau value Λ_{pl} of the divergence exponent as a function of the percentile p for the reference Lorenz attractor. The diagram refers to the same data as those used in Fig. 4.4. The behaviour is completely different from that one shown in Fig. 4.6. In particular, each point lays well above the blue line (and the dashed line) corresponding to the plateau behaviour in case of a Gaussian time series.
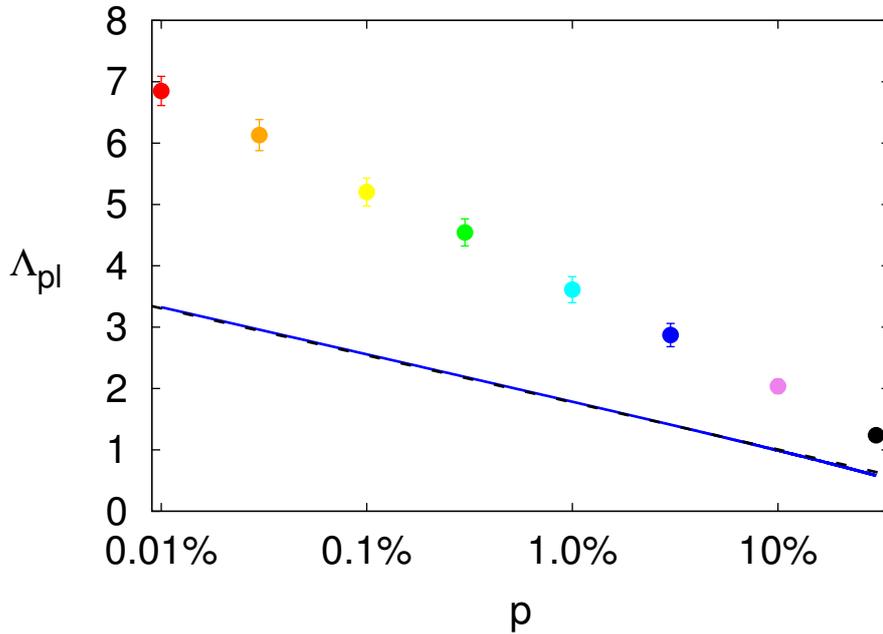
Figure 4.7: Plateau value Λ_{pl} of the divergence exponent as a function of the percentile p for the reference Lorenz attractor. The embedding pair used is $(3, 3)$. The abscissa of each point – and the related colour – corresponds to one of the percentile values reported in Fig. 4.4; the ordinate of each point, and the related errorbar, corresponds to the rightmost point ($k = 30$), with the respective errorbar, in Fig. 4.4. The blue line corresponds to Eq. (4.17), the dashed line to Eq. (4.19).

This observation, along with the discussion of the previous section, suggests that looking at the plateau value Λ_{pl} makes up a new tool to tackle the problem of distingushing stochastic noise from other kinds of dynamics in a time series (see Chapter 1), or at least to find evidence of a non-stochastic behaviour.

A further evidence of the last statement is the following. If a purely deterministic chaotic system is contaminated by a progressively growing amount of stochastic noise, one should observe a progressive shift from the non-stochastic behaviour of Λ_{pl} as in Fig. 4.7 to a stochastic one as in Fig. 4.6. To test this prediction, rather than simply adding a Gaussian term to the solution of the noiseless equation, we dynamically fed Gaussian noise

into the evolution of the Lorenz attractor, by modifying Eq. (2.5) as follows:

$$
\begin{aligned}
\frac{dx}{dt} &= \sigma(y - x) + \alpha \cdot \eta_x / \sqrt{3}, \\
\frac{dy}{dt} &= x(\rho - z) - y + \alpha \cdot \eta_y / \sqrt{3}, \\
\frac{dz}{dt} &= xy - \beta z + \alpha \cdot \eta_z / \sqrt{3},
\end{aligned}
$$

where η_x, η_y, η_z, are three standard normal random variables and α is the noise amplitude (instantaneous standard deviation). Fig. 4.8 shows the results, that are in agreement with the prediction.
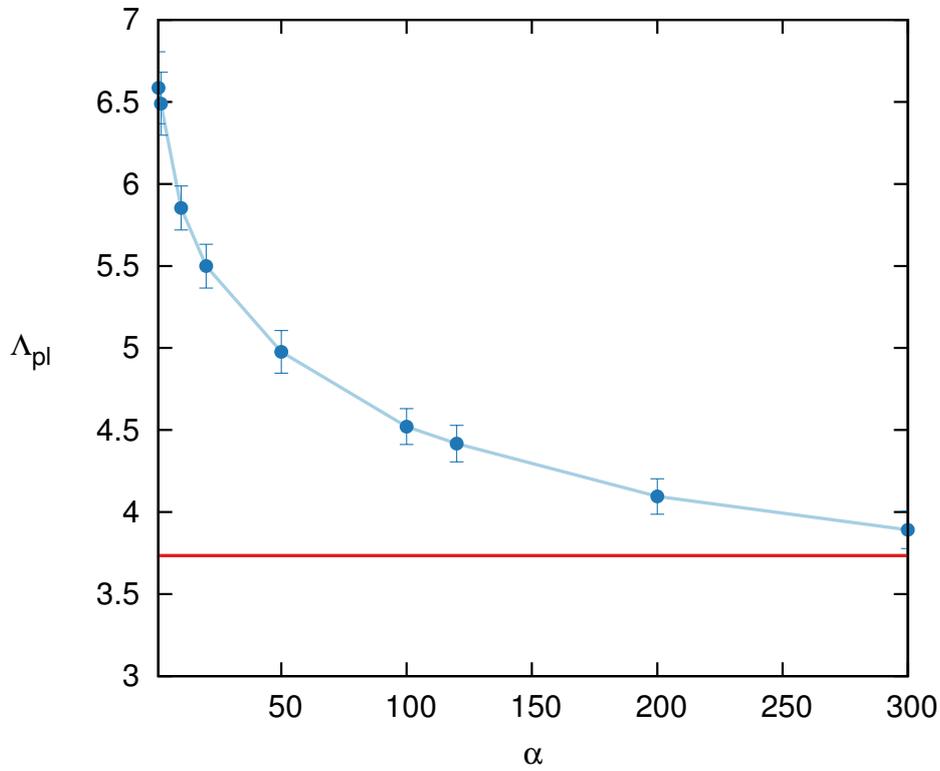
Figure 4.8: Dependence of the plateau value Λ_{pl} of the divergence exponent on the amplitude α of the dynamical noise added to the evolution of the reference Lorenz attractor. The embedding pair used is $(4, 4)$ and the percentile p is set to 0.01 %. The red line corresponds to the value $\Lambda_{pl}(4, 10^{-4})$ provided by Eq. (4.17).

4.2   Fractal dimension

The concept of dimension of a dynamical system is linked to the number of coordinates needed to specify a state. From the geometric point of view, the dimension is linked with how volumes scale as a function of length:

$$V \propto r^D \tag{4.20}$$

where V is the volume, r is a length and D is the dimension. Eq. (4.20) can be used to define the dimension D as follows:

$$D \equiv \lim_{r \to 0} \frac{\ln V}{\ln r} \,.$$

For example, D is equal to 2 for planar areas, and 3 for volumes in real space.

In contrast to linear systems, which always have an integer dimension, chaotic attractors have a fractal dimension. This last quantity can be used to detect chaotic dynamics and to distinguish chaos from noise. A large number of reviews on dimension and its estimation are available [109, 108, 12].

In this section, we shortly discuss four procedures used to estimate the dimension of an attractor: the box counting, the information dimension, the correlation dimension and the Kaplan–Yorke conjecture.

4.2.1   Box counting dimension

Let us consider an attractor "living" in a space $\mathbb{R}^n$. Given a length ϵ we consider the number $N(\epsilon)$ of hypercubes of edge ϵ needed to cover the attractor. The box counting dimension is then defined as:

$$D_\square \equiv \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln \left(\frac{1}{\epsilon}\right)} \,.$$

As an example, let us assume an attractor that densely covers a hypercube of edge L. The volume of the hypercube is L^n; the volume of the hypercubes of edge ϵ is ϵ^n, so that $N(\epsilon) = (L/\epsilon)^n$. It is straightforward to show that in this case $D_\square = n$.

4.2.2 Information dimension

Again, we consider an attractor "living" in a space $\mathbb{R}^n$ and, given a length ϵ, we partition the space in hypercubes of edge ϵ. Let $P_i(\epsilon)$ is the probability that the i-th hypercube is "populated", i.e. that a state of the system lays within it. The normalization condition holds:

$$\sum_{\forall i} P_i(\epsilon) = 1 \,.$$

Let S be the *Shannon information* (entropy) [110] linked to the probability distribution $P_i(\epsilon)$:

$$S(\epsilon) = -\sum_{\forall i} P_i(\epsilon) \ln P_i(\epsilon) \,.$$

The information dimension is then defined as follows [109]:

$$\sigma \equiv \lim_{\epsilon \to 0} \frac{S(\epsilon)}{\ln\left(\frac{1}{\epsilon}\right)} \,.$$

If, again, the attractor densely covers a hypercube of edge L, we have $P_i(\epsilon) = \frac{1}{N(\epsilon)}$, with $N(\epsilon)$ defined as in Sec. 4.2.1. It is straightforward to show that in this case $\sigma = n = D_\square$. From arguments linked to entropy it follows: $\sigma \leqslant D_\square$.

4.2.3 Correlation dimension

In 1983, Grassberger and Procaccia [111, 112] introduced the concept of correlation dimension. Given a scalar time series

Given a sample time sequence $\{x_n\}$ of a scalar, real variable x, as in Sec. 4.1.2 we consider an m-dimensional embedding with lag L, yielding a set of m-dimensional vectors $\mathbf{X}_i$. Let the correlation integral [111] $C(r)$ be defined as:

$$C(r) \equiv \lim_{N \to \infty} \frac{2}{N(N-1)} \sum_{\forall (i,j)} \theta(r - |\mathbf{X}_i - \mathbf{X}_j|) \,,$$

where θ is the Heaviside step function.

Remarkably, for small r, $C(r)$ behaves as a power of r:

$$C(r) \propto r^\nu. \tag{4.21}$$

The exponent ν is the so-called correlation dimension. If the time series is given by Gaussian noise with variance α^2, the correlation dimension ν can be evaluated by noting that Eq. (4.21) corresponds to the cumulative distribution of distances. Thus, by exploiting Eq. (4.14) and Eq. (4.18), we can write

$$
\begin{aligned}
C(r) = p(r) &= \int_0^{\frac{r^2}{2\alpha^2}} f_m\left(\chi_m^2\right) d\chi_m^2 \\
&\approx \frac{1}{2^m \alpha^m \Gamma\left(\frac{m}{2} + 1\right)} r^m \quad \text{if} \quad r \to 0.
\end{aligned}
$$

Consequently, the correlation dimension is equal to the embedding dimension, i.e. $\nu = m$. On the other hand, if the system is deterministic (ex.gr. chaotic), the behaviour given by Eq. (4.21) is embedding-independent provided that $m > \nu$. The correlation dimension can be indeed used to distinguish between chaotic and stochastic behaviour [112]. It can be shown that $\nu \leqslant \sigma \leqslant D$ [111].

4.2.4 Kaplan–Yorke dimension

As shown in Sec. 4.1.1, once the evolution laws of a dynamical system are known, the Lyapunov exponents can be calculated via the standard method. In 1979, Kaplan and Yorke introduced the dimension D_{KY} in terms of the χ_i's [113]:

$$D_{KY} = k + \sum_{i=1}^{k} \frac{\chi_i}{|\chi_{k+1}|},$$

where k is the maximum integer such that the sum of the k largest exponents is non-negative. In the scientific literature the dimension D_{KY} is often referred to as Lyapunov dimension. It can be shown that D_{KY} makes up an upper bound for the information dimension σ of the system. According to

dynamical system	k	Sano and Sawada [114]	present work
Hénon	1	1.26	1.26
Lorenz	2	2.06	2.08
Rössler	2	2.1	2.009
Mackey–Glass	3	3.58	3.55

Table 4.2: Kaplan–Yorke dimension D_{KY} for the reference dynamical systems (see Sec. 2.5). Our values are compared with those provided by Sano and Sawada [114].

the so-called Kaplan-Yorke conjecture, $D = D_{KY}$.

Tab. 4.2 reports the values of D_{KY} obtained for the reference dynamical systems of Sec. 2.5.

4.3   Kolmogorov entropy

In this section we briefly mention, for the sake of completion, another important measure to analyse dynamical systems, namely Kolmogorov entropy [115], which is a measure of how rapidly we loose the ability to predict the evolution of a dynamical system [116].

As above, we consider an attractor "living" in a space $\mathbb{R}^n$ and, given a length ϵ, we partition the space in hypercubes of edge ϵ. Let us consider $n + 1$ of these cells and label them $i_0, i_1, \ldots, i_n$. Taking into account an orbit $\mathbf{x} = \mathbf{x}(t)$ from time 0 to time $n\tau$, let $P_{i_0, i_1, \ldots, i_n}$ be the probability that $\mathbf{x}(0)$ lies in the i_0-th cell, $\mathbf{x}(k\tau)$ in the i_k-th cell, ..., $\mathbf{x}(n\tau)$ in the i_n-th cell. Let K_n be defined as follows:

$$K_n \equiv - \sum_{\{i_0,\ldots,i_n\}} P_{i_0,\ldots,i_n} \ln P_{i_0,\ldots,i_n} \, ,$$

where the sum is taken on all possible n-tuplets of cells.

The difference $K_{n+1} - K_n$ is linked to the additional information required to specify which cell will be occupied by $\mathbf{x}(n\tau + \tau)$, i.e. the amount of information that is lost from $n\tau$ to $(n + 1)\tau$. The Kolmogorov $\mathcal{K}$-entropy is

then defined as [116]:

$$\mathcal{K} \equiv \lim_{\epsilon \to 0} \lim_{\tau \to 0} \lim_{N \to \infty} \frac{1}{N\tau} \sum_{n=0}^{N-1} (K_{n+1} - K_n) =$$

$$= -\lim_{\tau \to 0} \lim_{\epsilon \to 0} \lim_{N \to \infty} \frac{1}{N\tau} \sum_{i_0,\ldots,i_N} P_{i_0,\ldots,i_N} \ln P_{i_0,\ldots,i_N} .$$

The Kolmogorov $\mathcal{K}$-entropy is useful for distinguishing regular from chaotic behaviour as well as deterministic chaos from stochastic noise [117]. A linear system has $\mathcal{K} = 0$ because there is no loss of information during the time evolution. For chaotic systems $\mathcal{K} > 0$: on average, initially nearby points separate exponentially, and thus joint probabilities for cell occupations decrease exponentially with time. For a stochastic system, initially nearby points will spread independently and with vanishing correlation time all over the allowed volume, so that $\mathcal{K} \to \infty$ for pure randomness.

Intuitively, the growth of uncertainty is due to the sensitive dependence on initial conditions, and it is therefore linked to the divergence of nearby trajectories. A relation between the Kolmogorov entropy and the Lyapunov exponents can be expected. This relation was first discussed by Presin [118], and later by Ruelle [119] who showed that $\mathcal{K}$ is bounded by the sum of positive Lyapunov exponents:

$$\mathcal{K} \leqslant \sum_{\forall i \,|\, \chi_i > 0} \chi_i .$$

Chapter 5

Determination of the maximum Lyapunov exponent and its dependence on the embedding

This chapter deals with the maximum Lyapunov exponent (MLE) and its uncertainty, both evaluated by means of the divergent rate method (see Sec. 4.1.2).

In Sec. 3.3 the major issue of Takens embedding theorem is mentioned: despite its power, the theorem gives no clue on how to set the embedding dimension m and lag L for an experimental time series. As discussed in the next chapter, several methods have been developed to tackle this issue. Unfortunately, none of these methods has been yet proven to be superior to the others. Cellucci *et al.* [14] proposed a technique to rank the embedding methods based on the evaluation of the MLE.

A main point of the proposed ranking technique lays in the possibility of determining the MLE χ_S of a dynamical system S independently from the embedding parameters, by using the *standard method* [15, 16] described in Sec. 4.1.1. As discussed in the previous chapter, given an (m, L)-embedded time series, i.e., embedded by using a particular parameter pair (m, L), the MLE can be estimated out of a time series by means of the divergent rate method, i.e. by determining the local divergence rate $\Gamma_S(m, L)$ of nearby

trajectories [8, 96, 97]. So, according to Cellucci *et al.*, method A is considered to be superior to method B if $\Gamma_S(m_A, L_A)$, calculated by using the embedding pair assessed via A, approaches χ_S better than $\Gamma_S(m_B, L_B)$, calculated by using the embedding pair assessed via B [17].

In this chapter, we address the validity of the method proposed by Cellucci *et al.*; we show that the crucial parameter is the uncertainty $\sigma_{\Gamma,S}(m, L)$ of the MLE $\Gamma_S(m, L)$. In particular, if different embedding points provide $\Gamma_S(m, L)$ values that differ less than $\sigma_{\Gamma,S}(m, L)$, then the dynamical system turns out to be unsuitable for the sake of comparing different embedding methods. Grassberger, Badii and Politi [18] discussed the origin of the statistical uncertainty, or noise, that affects finite time MLE evaluations by using an argument based on the central limit theorem. They showed that, for attractors with short correlation time, the noise is normally distributed and its amplitude depends on the number of time steps used to evaluate the MLE. Prasad and Ramaswamy [120] showed that, in the case of intermittent attractors, the noise distribution has a positive skewness due to exponential tails.

As a consequence of the Takens embedding theorem (see Sec. 3.2), if the embedding choice is correct, the intrinsic noise behaviour predicted by Grassberger, Badii and Politi should be reflected in the distribution of the MLE $\Gamma_S(m, L)$ estimated via the divergence rate method. Consequently, the mean of distributions corresponding to different, correct embedding choices have necessarily to approach the same value χ_S. For the same reason, the uncertainty $\sigma_{\Gamma,S}(m, L)$ corresponding to different, correct embedding choices have to tend to the same value.

This chapter is organized as follows. Sec. 5.1 presents the notation and the procedure used to analyse the four reference dynamical systems under study. In Sec. 5.2 the relation between $\Gamma_S(m, L)$ and $\sigma_{\Gamma,S}(m, L)$ is discussed. The distribution of the MLE uncertainty is the topic of Sec. 5.3. In Sec. 5.4 we present a way to visualize MLE as a function of the embedding point (m, L) as well as embedding pairs of the (m, L) lattice that can be deemed to be "good".

5.1 Notation and MLE analysis

The four reference dynamical system studied in this chapter, namely Hénon map, Lorenz attractor, Rössler attractor and Mackey–Glass attractor, are described in Sec. 2.5. The MLEs used as reference values, obtained by the standard method (see Sec. 4.1.1), are reported in Tab. 4.1; we remind that each of these values corresponds to the sample mean and the sample standard deviation of the results of 10 runs, each made of 10^7 integration steps.

For each reference dynamical system $\mathcal{S}$, the MLE $\Gamma_{\mathcal{S}}(m, L)$ and the related uncertainty $\sigma_{\Gamma,\mathcal{S}}(m, L)$ are evaluated by means of the divergent rate method discussed in Sec. 4.1.2. The MLE was calculated on the lattice $2 \leqslant m \leqslant 10$, $1 \leqslant L \leqslant 10$. The maximum value of the delay k taken into account for the evaluation of $\Lambda = \Lambda(k)$ is 100 for the Hénon map, 200 for the Lorenz attractor, and 400 for both the Mackey–Glass and the Rössler attractors. The radius r is set so as to correspond to the percentile $p = 0.01\%$. For each system and embedding pair, a number $M = 50$ (see Eq. (4.11)) of calculations was carried out.

Upon evaluation, the value $\Gamma_{\mathcal{S}}(m, L)$ and its uncertainty $\sigma_{\Gamma,\mathcal{S}}(m, L)$ are normalized to the related standard value $\chi_{\mathcal{S}}$, namely:

$$\mu_{\mathcal{S}}(m, L) \equiv \frac{\Gamma_{\mathcal{S}}(m, L)}{\chi_{\mathcal{S}}},$$

$$\sigma_{\mathcal{S}}(m, L) \equiv \frac{1}{\chi_{\mathcal{S}}} \left[\sigma^2_{\Gamma,\mathcal{S}}(m, L) + \frac{\Gamma^2_{\mathcal{S}}(m, L)\sigma^2_{\chi,\mathcal{S}}}{\chi^2_{\mathcal{S}}} \right]^{1/2}, \qquad (5.1)$$

where $\sigma_{\chi,\mathcal{S}}$ is the uncertainty on $\chi_{\mathcal{S}}$.

5.2 Uncertainty vs. MLE

The results of the evaluation of $\mu_{\mathcal{S}}$ and $\sigma_{\mathcal{S}}$ are shown in Fig. 5.1. For each dynamical system, the diagram contains the scatter plot of the points $(\mu_{\mathcal{S}}, \sigma_{\mathcal{S}})$ each corresponding to an embedding pair of the lattice $2 \leqslant m \leqslant 10$,

$1 \leqslant L \leqslant 10$. In addition, the uncertainty is normalized to a value u_S, whose meaning is discussed below. Points are clearly not uniformly distributed; they form, rather, clusters; in particular, points for which $\mu_S(m, L) \approx 1$ (or equivalently, $\Gamma_S(m, L) \approx \chi_S$), are characterized by a lower uncertainty. The investigation of this correlation is the topic of the following section.
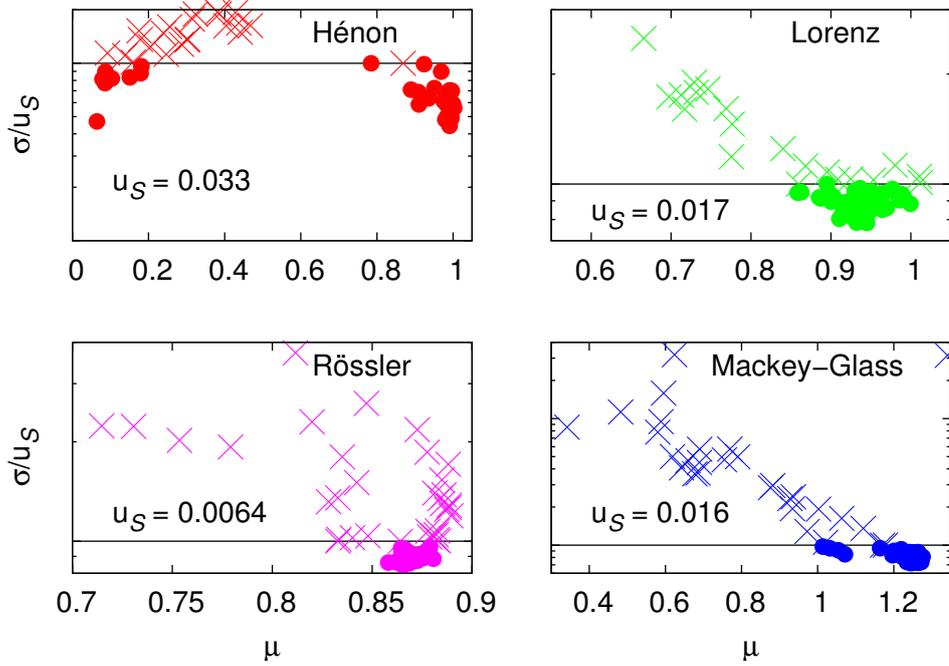
Figure 5.1: Diagram of the uncertainty $\sigma_S(m, L)$ (normalized to the respective u_S; see Sec. 5.3), vs. normalized MLE $\mu_S(m, L)$. Each point corresponds to an embedding pair. Black lines correspond to $\sigma = u_S$. Points below (above) the black lines are represented with dots (crosses). Ordinate axis are logarithmically scaled.

5.3 Distribution of uncertainty

Because $\sigma_{\Gamma,S}^2(m, L) \gg \sigma_{\chi,S}^2$, the statistical properties of $\sigma_S(m, L)$ are mainly linked to $\sigma_{\Gamma,S}(m, L)$ (see Eq. (5.1)).

The $\Gamma_S(m, L)$ is calculated by means of a straight-line fit (see Eq. (4.12a)). The squared uncertainty is essentially due to the sum of contributions stemming from the variances of the points belonging to the time-dependent

divergence exponent $\Lambda(k)$ (see Eq. (4.12b)). In the case of Lorenz, Mackey–Glass, and Rössler attractor, the number of degrees of freedom ν is of order of the length of the segment $\ell \gtrsim 50$. Consequently, for each embedding pair (m, L), the squared uncertainty is expected to follow a normal distribution as a consequence of the central limit theorem. The normal approximation should also work satisfactorily with regard to the Hénon map, for which $\ell \lesssim 10$. If one assumes homoscedasticity with respect to k in Eqs. (4.12), (4.13), so that $\sigma_\Lambda(k) = \sigma_\Lambda$ for the k-values belonging to the $\Lambda(k)$ slope, the mean value of the uncertainty can be estimated as follows:

$$\sigma_{\Gamma,\mathcal{S}}(m, L) \approx \frac{\sigma_\Lambda}{\Delta} \sqrt{\frac{12}{\ell}}, \tag{5.2}$$

where $\Delta \equiv \ell \delta t$ is the length of the time interval on which the linear fit is carried out.

As discussed in the introduction to this chapter, the distribution of a finite-time MLE is approximately normal. Consequently, its squared uncertainty is expected to be χ^2 distributed. Let us consider the embedding pairs where $\Lambda(k)$ plots deliver MLE values close to the standard one: in this case the MLE values are expected to follow the same distribution, independently of the values of m and L; the same occurs for the corresponding squared uncertainties. This happens because the $\Lambda(k)$ plots that deliver "good" MLE values approximately have the same slope and the same length Δ. This observation leads to the fact that, once observational conditions are fixed (ex. gr. the sampling time δt), a dynamical system is characterized by a typical MLE distribution as well as by a typical uncertainty distribution, and thus by a typical average uncertainty. The following discussion concerns the verification of this assumption.

Given the system $\mathcal{S}$ and a positive number $u_\mathcal{S}^2$, let $V_\mathcal{S}(u_\mathcal{S})$ be the set of the sample variances (the squared uncertainties) that do not exceed $u_\mathcal{S}^2$: $V_\mathcal{S}(u_\mathcal{S}) \equiv \{\sigma_\mathcal{S}^2(m, L) \mid \sigma_\mathcal{S}^2(m, L) \leqslant u_\mathcal{S}^2\}$. By exploiting the Shapiro–Wilk normality test [121] and using a significance threshold $\alpha = 0.01$, we can determine the maximum value $u_\mathcal{S}^2$, if any, such that the set $V_\mathcal{S}(u_\mathcal{S})$ is compatible with a normal distribution. The value of $u_\mathcal{S}$ for each of the four dynamical

systems is reported in Tab. 5.1 along with the root-mean-square value Σ_S of the uncertainties belonging to $V_S(u_S)$. The same values are also reported in Fig. 5.1.

Table 5.1: The central column reports, for each dynamical system, the maximum value u_S such that the set $V_S(u_S)$ of the squared uncertainties that do not exceed u_S^2 is compatible with a normal distribution. The compatibility is assessed via Shapiro–Wilk test ($p - value \geqslant 0.01$). The rightmost column reports, again for each dynamical system, the root-mean-square value Σ_S of the uncertainties belonging to $V_S(u_S)$.

Dynamical system (S)	u_S	Σ_S
Hénon	0.033	0.023
Lorenz	0.017	0.015
Rössler	0.0064	0.0057
Mackey–Glass	0.016	0.013

The root-mean-square Σ_S depends on the observational conditions, namely the values of the parameters that concur to define the dynamical system; for example, in the case of the Lorenz attractor, the parameters σ, β, ρ (see Sec. 2.5). However, the root-mean-square Σ_S does not depend on the embedding choice. Thus, it turns out to be a parameter typical of the dynamical system S: it describes the uncertainty with which the MLE can be estimated by means of the divergence rate method.

The distribution of the variance for each of the four dynamical systems is shown in Fig. 5.2. As predicted, the red-coloured histograms, which correspond to the sets $V_S(u_S)$, are bell-shaped.

Fig. 5.3 shows the distribution of the normalized MLE μ_S for each dynamical system. The histograms marked in red corresponds to embedding pairs for which $\sigma_S(m, L) \leqslant u_S$; these values tend to cluster around 1 or in regions nearby. In the case of Hénon map, there is an additional red cluster at $\mu_S(m, L) \approx 0.1$. This second cluster can be explained by noting that the Hénon map is a discrete-time dynamical system with low correlation time; while the optimal embedding pairs have $L \geqslant 2$, increasing the lag parameter ($L \geqslant 3$) generates aliasing, i.e. the embedding "samples" a trajectory too
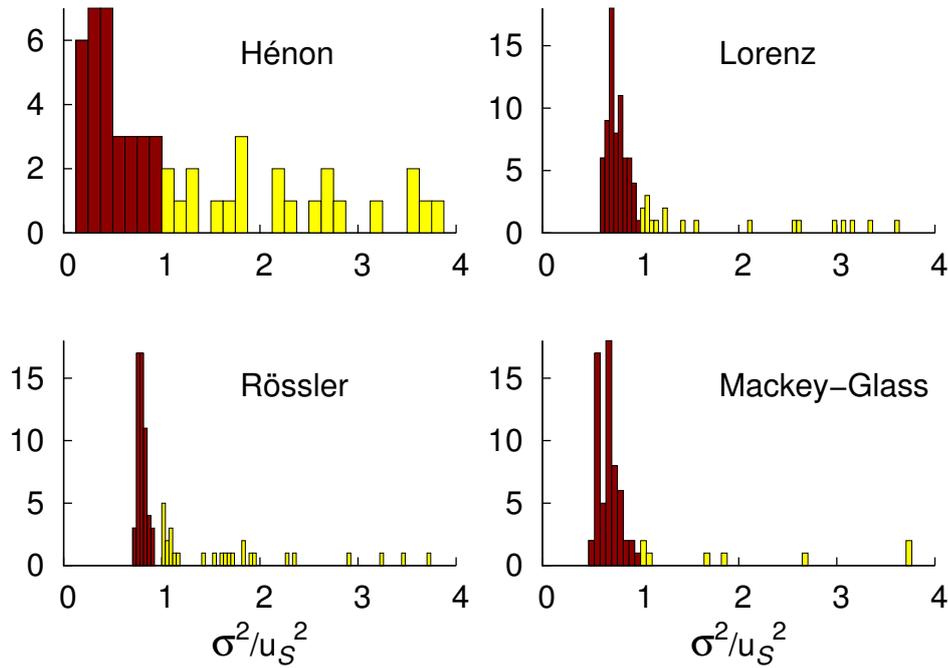
Figure 5.2: Distribution of variance for each of the four dynamical systems. Red-coloured bins refer to $\sigma_S(m, L)$ values that do not exceed the respective threshold u_S, whereas yellow-coloured bins refer to $\sigma_S(m, L)$ values that exceed the respective threshold u_S. For the sake of clarity, only data corresponding to $\sigma_S^2(m, L)/u_S^2 \leqslant 4$ are shown.

slowly, in fact observing a different system.

Fig. 5.3 also shows how potentially "good" points – namely embedding points such that $\mu_S(m, L) \approx 1$ – are distributed in comparison with the typical uncertainty Σ_S (see Tab. 5.1). For example, in the case of the Rössler attractor, potentially "good" MLE values have a peak around 0.87 and a standard deviation of 0.005. The normalized MLE μ of the Rössler attractor is clearly slightly underestimated, probably because of the intermittency, i.e. the extremely long correlation time, of this attractor. More importantly, the width (standard deviation) of the distribution is very close to the Σ_S value for the Rössler dynamical system (see Tab. 5.1), and therefore well within the corresponding 99% non-rejection region, given by $5.15 \cdot \Sigma_S$. Consequently, the Rössler attractor is not useful to compare different embedding methods by using the MLE calculation as a gauge: different embedding choices essentially yield the same result. On the other hand, the other three dynamical
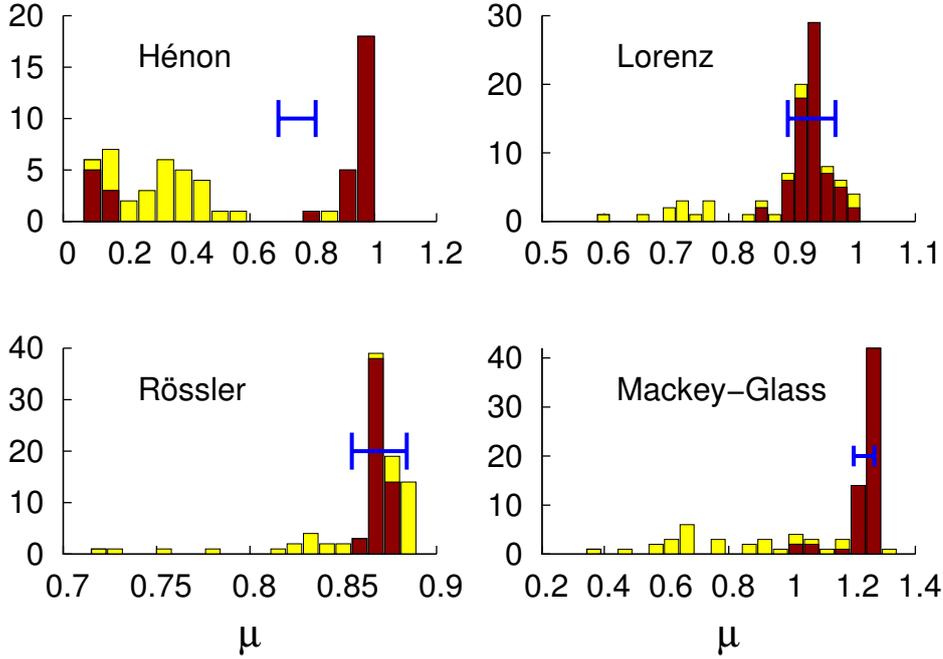
Figure 5.3: Distributions of the normalized MLE for each of the four dynamical systems: each yellow-coloured histogram refers to all available data; the superimposed, red-coloured histogram refers to normalized MLE values whose corresponding σ_S does not exceed the respective threshold u_S (i.e. $\sigma_S^2(m, L) \in V_S(u_S)$). Each horizontal errorbar corresponds to $5.15 \cdot \Sigma_S$, i.e. the width of the two-tailed 99% non-rejection region (level of significance equal to 0.01). The errorbars are centered around the mean values of the red histograms.

systems – and especially the Mackey–Glass attractor – appear to be more appropriate to be used as test benches.

Identifying normally-distributed clusters of MLE and its uncertainty is thus an additional method to investigate an unknown dynamical system by analysing time sequences generated by the system itself. In particular, the identification of normally-distributed clusters of MLE and its uncertainty can provide a valuable tool to optimally choose the embedding pair.

Finally, it is worth noting that the statistical properties of the finite-time MLE evaluated by means of our algorithm complies with the theory discussed by Grassberger, Badii and Politi [18]. In the case of continuous systems with short correlation times, if the sampling frequency δt^{-1} is sufficiently high, the time interval $\Delta = \ell \delta t$ is independent from the sampling

time δt and therefore from the number of points ℓ (for the definition of ℓ see Sec. 4.1.2). Consequently, according to Eq. (5.2), the uncertainty scales as $\ell^{-1/2}$, in agreement with the results of these authors. On the other hand, in the case of intermittent systems, where correlations are long-living, the length Δ is expected to be dependent on the sampling frequency, and therefore on ℓ. Thus, the power-law scaling of uncertainty $\sigma_S(m, L) \sim \ell^{-p}$ is modified, so that the exponent p departs from the standard value 0.5 [120, 122].

5.4 MLE on the embedding lattice: a graphical analysis

In this section, each one of Figs. 5.4, 5.5, 5.6, 5.7 shows two diagrams: in the top one, the reference MLE value is given – as in the rest of the chapter – by the standard method; in the bottom one, the reference value MLE is set to be equal to the average of the $\mu_S(m, L)$ values of the points for which $\sigma_S \leqslant u_S$. Although the MLE is defined on a discrete lattice, for the sake of clarity a bilinear interpolation is used in each diagram to generate a continuous map. The dark-green shade represents $\mu_S(m, L)$ values that differ from unity (the ideal value) by less than $2.576\,\Sigma_S$; the (m, L)-pairs contained in this area can be considered to be "good" embedding choices. Shade colours correspond to different values of $|\mu_S - 1|$ (see palette), whereas the contour lines are: $\mu = 0.8 \leftrightarrow$ dashed orange; $\mu = 0.9 \leftrightarrow$ orange; $\mu = 1.1 \leftrightarrow$ cyan; $\mu = 1.2 \leftrightarrow$ dashed cyan.

With regard to the points, the type depends on whether the variance σ_S^2 does (does not) belong to $V_S(u_S)$. The point size refers to how close the normalized MLE $\mu_S(m, L)$ is to its ideal value, i.e. unity. If the point type is a circle and if $\mu_S(m, L)$ differs from unity by less than $2.576\,\Sigma_S$ ($p - value > 0.01$) then the circle is coloured in red; given the values in Tab. 5.1, the maximum value of $2.576\,\Sigma_S$ occurs for the Hénon map and is equal to $2.576 \times 0.023 \simeq 0.06$. Red points can be considered to represent "good" embedding choices. A summary of this point code is contained in Tab. 5.2.

Finally, in each figure additional symbols, namely triangles, squares,

diamonds, correspond to the embedding pairs proposed in the scientific literature by using different methods to find an optimal embedding. In addition, a violet pentagon is shown, which corresponds to our estimate of optimal embedding pair. This pair is assessed by using the method discussed in Chapter 6.

Table 5.2: Point code used in Figs. 5.4, 5.5, 5.6, 5.7.

circles	● ●	$\sigma_S^2 \in V_S(u_S)$
crosses	✕	$\sigma_S^2 \notin V_S(u_S)$
• ✕		$0.2 < \lvert \mu_S(m, L) - 1 \rvert$
• ✕		$0.1 < \lvert \mu_S(m, L) - 1 \rvert \leqslant 0.2$
● ✕		$\lvert \mu_S(m, L) - 1 \rvert \leqslant 0.1$
●		$\lvert \mu_S(m, L) - 1 \rvert \leqslant 2.576\,\Sigma_S$

The top part of Fig. 5.4 shows, in the case of the Hénon map, that some low variance points (circles) are located in the brown area, corresponding to a relatively large difference of μ_S from 1; these points belong to the spurious cluster located at low MLE value ($\mu_S \approx 0.1$) in Fig. 5.3; the origin of this spurious effect is discussed in Sec. 5.3.

Because all the dynamical systems discussed here show a great number of "good" (m, L) pairs, we can conclude that the method proposed by Cellucci *et al.* [14] to rank the embedding methods does not provide valuable information if it is combined with these reference dynamical systems. Moreover, the methods to find an optimal embedding point discussed in the scientific literature provide "good" result for the Hénon map, the Lorenz attractor and the Mackey–Glass attractor. However, in the case of the Rössler attractor, two methods provide $(3, 8)$ and $(3, 9)$ as optimal embedding pairs; both these points cannot be deemed to be "good". On the other hand, for each system, the optimal embedding pair provided by our method corresponds to a "good" embedding point.
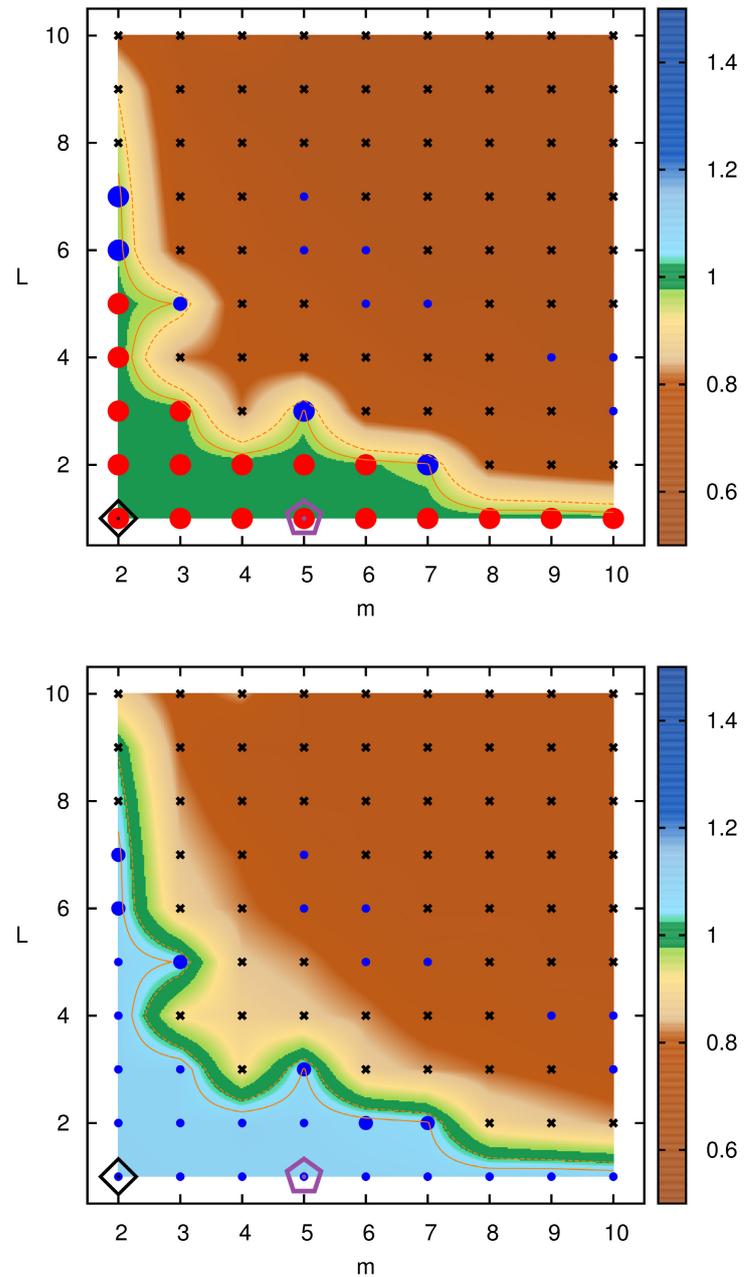
Figure 5.4: Normalized MLE $\mu_s(m, L)$ for the Hénon map. The type, colour and size of the points, as well as the colour of contour lines and shades, are explained in Table 5.2 and in the main text.
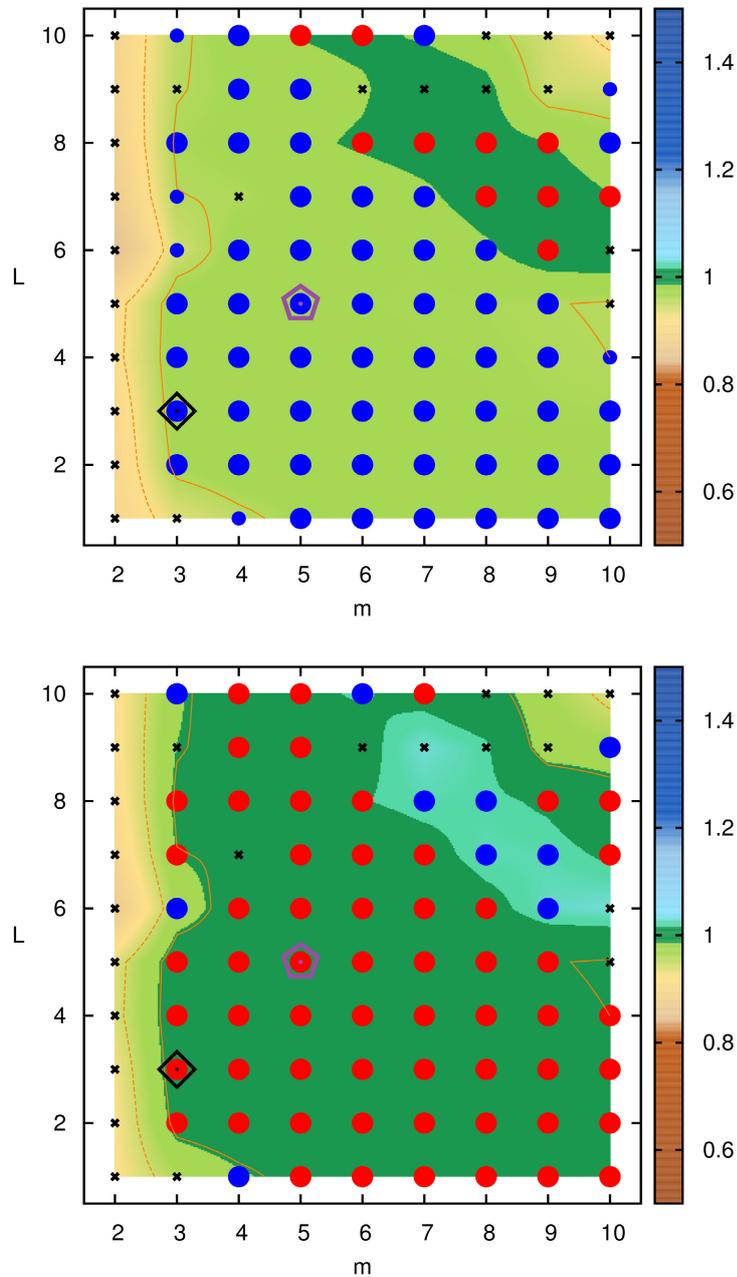
Figure 5.5: Normalized MLE $\mu_S(m, L)$ for the Lorenz attractor. The type, colour and size of the points, as well as the colour of contour lines and shades, are explained in Table 5.2 and in the main text.
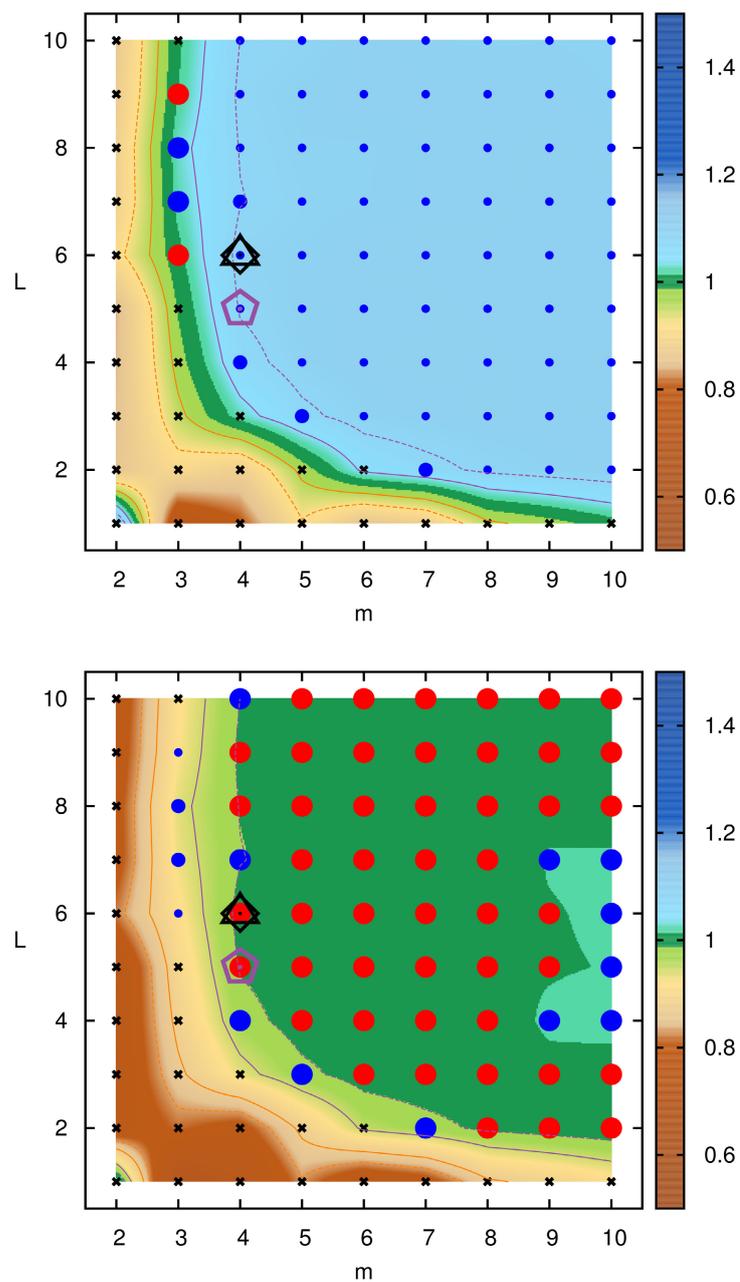
Figure 5.6: Normalized MLE $\mu_S(m, L)$ for the Mackey–Glass attractor. The type, colour and size of the points, as well as the colour of contour lines and shades, are explained in Table 5.2 and in the main text.
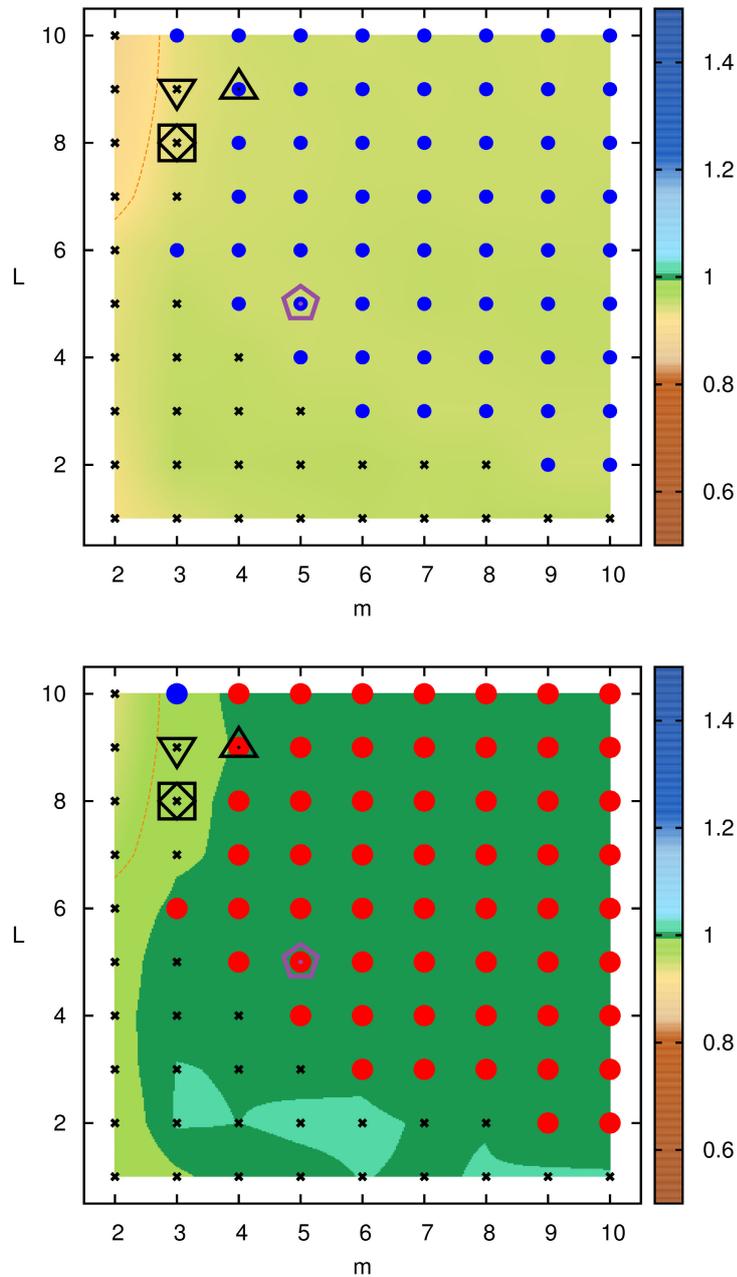
Figure 5.7: Normalized MLE $\mu_S(m, L)$ for the Rössler attractor. The type, colour and size of the points, as well as the colour of contour lines and shades, are explained in Table 5.2 and in the main text.

Chapter 6

A Kolmogorov–Smirnov approach to calculate the optimum embedding dimension and lag

The major issue of the Takens embedding theorem is the fact that it gives no clue to identify the correct embedding pair (m, L). To tackle this issue, several methods to have been developed. The most used and cited methods are: Gao–Zheng method [8, 9, 10], method of Schuster [11], the method of characteristic length, the method of global false-nearest-neighbours and autocorrelation function, the method of global false-nearest-neighbours and mutual information [12, 13].

In this chapter we analyse in detail two of these methods, as they were a starting point for the development of a new one, which is the topic of the second part of the chapter. A general review of the other available methods is reported in the paper by Cellucci *et al.* [14].

Sec. 6.1 and Sec. 6.2 present the Gao–Zheng method and the method of Schuster, respectively. In Sec. 6.3 we discuss a new embedding method that relies on a Kolmogorov–Smirnov approach. The results are discussed in Sec. 6.4.

6.1   Gao–Zheng method

In 1993, Gao and Zheng [8, 9, 10] proposed a method to find the optimal embedding pair. The method is based on the assumption that a good embedding point minimizes the number of false-nearest-neighbours, i.e. vectors of the phase space that appear to be close to each other, although their apparent proximity is actually due to crossing orbits.

Given a scalar time series $\{x_n\}$, let $\mathbf{X}_i$ be an embedding vector in $\mathbb{R}^m$ with a lag L. To verify the real nearest-neighbourhood of two embedding vectors $\mathbf{X}_i$ and $\mathbf{X}_i$ in $\mathbb{R}^m$, Gao and Zheng assumed that, if $\mathbf{X}_i$ is a neighbour of $\mathbf{X}_j$ then, for a "small" k, $\mathbf{X}_{i+k}$ must be a neighbour of $\mathbf{X}_{j+k}$. Gao and Zheng proposed to fix k equal to the autocorrelation time, defined as the time required for the autocorrelation function to drop to $1/e$ of its initial value.

The two authors constructed the following measure

$$\Lambda(m,\,L;\,k) \equiv \left\langle \ln \frac{\|\mathbf{X}_{i+k} - \mathbf{X}_{j+k}\|}{\|\mathbf{X}_i - \mathbf{X}_j\|} \right\rangle,$$

where $\langle \cdots \rangle$ denotes the average on all $i,\,j$ pairs that satisfy the following two requirements:

- $|j - i| > w$, where w is set equal to the first minimum of the autocorrelation function;

- the distance between the starting embedding vectors must satisfy $\|\mathbf{X}_i - \mathbf{X}_j\| \leqslant r$.

Given k, the measure $\Lambda(m,\,L;\,k)$ depends on the embedding pair $(m,\,L)$. According to Gao and Zheng, the larger $\Lambda(m,\,L;\,k)$, the larger the number of false-nearest-neighbours. Based on this argument, they concluded that an embedding is deemed to be successful if the value of Λ does not "significantly" decrease when the embedding dimension m is increased; the optimal lag L corresponds to the value that, once given m, minimizes $\Lambda(m,\,L;\,k)$. An example is shown in Fig. 6.1: by using the Gao–Zheng method, we obtain as an optimal embedding point $(4,\,8)$ for the reference Rössler attractor. For the

same attractor, Gao and Zheng reported an optimal embedding pair equal to (3, 8).
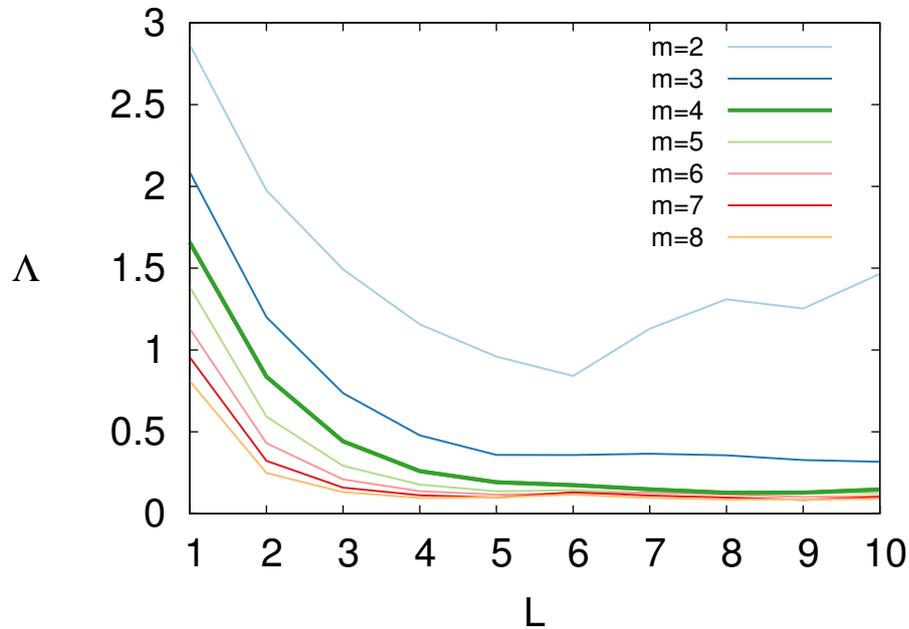
Figure 6.1: $\Lambda(m, L; k)$ as a function of the embedding pair (m, L) for the reference Rössler attractor (see Eq. (2.6) and Sec. 2.2).

The Gao–Zheng procedure suffers from two main drawbacks:

- it is strongly dependent on the delay k;

- there is no quantitative parameter to evaluate what "significantly" means (see above).

6.2   Method of Schuster

In 1991, Liebert, Pawelzik and Schuster proposed a procedure for estimating an optimal embedding dimension based on the examination of the relation between sets of nearest-neighbours in successive embeddings [11]. This procedure is referred to as method of Schuster in the scientific literature.

Given a scalar time series $\{x_n\}$, let $\mathbf{X}_i$ be an embedding vector in $\mathbb{R}^m$ with a lag L. Let $\mathbf{X}_{i,1}$, $\mathbf{X}_{i,2}$, ..., $\mathbf{X}_{i,N}$, be the first N nearest neighbours of $\mathbf{X}_i$ so that the distance $\|\mathbf{X}_{i,k} - \mathbf{X}_i\|$ increases with k. The method of Schuster considers the impact of increasing m to $m+1$ on the set of the nearest neighbours.

Let $\mathbf{X}'_i \in \mathbb{R}^{m+1}$ be given by $\mathbf{X}'_i \equiv (\mathbf{X}_i, x_{i+mL})$; $\mathbf{X}'_i \in \mathbb{R}^{m+1}$ is said to correspond to $\mathbf{X}_i \in \mathbb{R}^m$.

Let $\mathbf{X}'_{i,k} \in \mathbb{R}^{m+1}$ correspond to $\mathbf{X}_{i,k}$ in $\mathbb{R}^m$.

Let $\mathbf{Y}'_{i,k}$ denote the k-th nearest neighbour of $\mathbf{X}'_i$, where the nearest neighbours are ordered in the same way as before. Consequently, the vector $\mathbf{Y}'_{i,k}$ is linked to $\mathbf{X}'_i$ and it does not necessarily correspond to $\mathbf{X}_{i,k}$.

Liebert, Pawelzik and Schuster hypothesized that in the case of an optimal embedding, the nearest neighbours are preserved, i. e. $\mathbf{Y}'_{i,k} = \mathbf{X}'_{i,k}$, $\forall i, k$. They proposed a metric to quantify the extent at which this ideal relation fails to be true.

In the case of an ideal embedding, $\mathbf{Y}'_{i,k} = \mathbf{X}'_{i,k}$, the following equation holds:

$$\frac{\|\mathbf{X}'_i - \mathbf{X}'_{i,1}\|}{\|\mathbf{X}'_i - \mathbf{Y}'_{i,1}\|} = 1 \,.$$

If $\mathbf{Y}'_{i,1} \neq \mathbf{X}'_{i,1}$, this ratio is greater that 1, because $\mathbf{Y}'_{i,1}$ is the closest neighbour of $\mathbf{X}'_i$. The product

$$\mathcal{F} \equiv \prod_{k=1}^{N} \frac{\|\mathbf{X}'_i - \mathbf{X}'_{i,k}\|}{\|\mathbf{X}'_i - \mathbf{Y}'_{i,k}\|}$$

is a measure of the degree of correspondence between the sets $\mathbf{Y}'_{i,k}$ and $\mathbf{X}'_{i,k}$. A large value of $\mathcal{F}$ indicates a breaking of the nearest neighbour relation $\mathbf{Y}'_{i,k} = \mathbf{X}'_{i,k}$, which is possibly due to an unsuitable value of m.

By considering the projection of $\mathbf{Y}'_{i,k}$ back into the original $\mathbb{R}^m$ space, so that $\mathbf{Y}_{i,k}$ indicates the first m-component of the vector $\mathbf{Y}'_{i,k}$, the product $\mathcal{B}$ is defined:

$$\mathcal{B} \equiv \prod_{k=1}^{N} \frac{\|\mathbf{X}_i - \mathbf{X}_{i,k}\|}{\|\mathbf{X}_i - \mathbf{Y}_{i,k}\|} \,.$$

The product $\mathcal{F} \cdot \mathcal{B}$, defined as

$$W_i(m, L) \equiv \mathcal{F} \cdot \mathcal{B} = \prod_{k=1}^{N} \left(\frac{\|\mathbf{X}_i - \mathbf{X}_{i,k}\|}{\|\mathbf{X}_i - \mathbf{Y}_{i,k}\|} \right) \left(\frac{\|\mathbf{X}_i' - \mathbf{Z}_{i,k}'\|}{\|\mathbf{X}_i' - \mathbf{X}_{i,k}'\|} \right),$$

is "typical" of vector $\mathbf{X}_i$. By randomly selecting a set of vectors in the $\mathbb{R}^m$ embedding space and averaging the related values $W_i(m, L)$, the measure $W(m, L)$ is obtained:

$$W(m, L) \equiv \langle \ln W_i(m, L) \rangle,$$

where $\langle \ldots \rangle$ indicates the averaging process.

Citing Cellucci *et al.* [14], "*The best choice of embedding corresponds to the smallest value of m that produces the limiting behaviour of $W(m, L)$*".

The method of Schuster suffers from a main drawback:

- there is no quantitative parameter to define "the limiting behaviour".

6.3   A Kolmogorv–Smirnov approach

The common drawback of the two optimal embedding methods discussed above is the lack of clear quantitative parameters that define the optimal choice. This drawback also affects the other methods described in the paper by Cellucci *et al.* [14]. Citing from this paper:

- with regard to the method of characteristic length, "*Further increases in m do not, however, result in further increases $C(m, L)$*" (p. 7, left column);

- with regard to the method of global false-nearest-neighbours and autocorrelation function, "*The value of m is increased until false nearest neighbours are no longer observed or until the frequency of false nearest neighbours is below an acceptable value.*" (p. 7, right column);

- with regard to the method of global false-nearest-neighbours and mutual information, the problem is the same as in the previous method

because"*This procedure differs from the immediately preceding method in the criterion used to determine the lag*" (p. 7, right column).

In other words, all these methods require a human adjusting.

In this section, we describe a new optimal embedding method that evaluates the optimal (m, L)-pair by finding extrema of a potential-like quantity defined on the (m, L) lattice. The potential-like quantity – or, simply, the potential – is determined by evaluating, for each (m, L) pair, the distribution of the Euclidean distances between all pairs of embedding vectors $\mathbf{X}_i$, $\mathbf{X}_j$, and by using a metrics based on the Kolmogorov–Smirnov test.

6.3.1 Hypothesis

The method of Schuster (see Sec. 6.2) takes into account the relation between sets of nearest neighbours in successive embeddings, ex. gr. (m, L) and $(m + 1, L)$. Our approach considers not only the relation between embeddings that differ by one step in m, but also embeddings that differ by one step in L: (m, L), $(m, L + 1)$.

The method is based on two assumptions:

- a potential-like quantity $P_{m, L}$ is defined on a plain lattice of points (m, L) that are characterized by the embedding dimension m and the lag L;

- optimal embedding points coincide with local extrema (maxima or minima) of $P_{m, L}$.

To define $P_{m, L}$, a Kolmogorov–Smirnov approach is used, as follows.

Given a scalar time series $\{x_n\}$, let $\mathbf{X}_i$ be an embedding vector in $\mathbb{R}^m$ with a lag L. Given two vectors $\mathbf{X}_i$, $\mathbf{X}_j$, let d be their Euclidean distance normalized to $\sqrt{m}$:

$$d_{i,j} = \frac{1}{\sqrt{m}}\|\mathbf{X}_i - \mathbf{X}_j\|, \qquad i \neq j.$$

By considering a suitable number of vectors, the cumulative distribution

$F_{(m, L)}(d)$ of the distance d between pairs of distinct vectors $\mathbf{X}_i$, $\mathbf{X}_j$ can be estimated.

Given two embedding points (m_1, L_1), (m_2, L_2), let us consider the Kolmogorov–Smirnov (KS) distance D between the two related cumulative distributions $F_{m_1, L_1}(d)$, $F_{m_2, L_2}(d)$ (see Fig. 6.2):

$$D_{(m_1, L_1) \leftrightarrow (m_2, L_2)} = \sup_{\{d\}} |F_{m_1, L_1}(d) - F_{m_2, L_2}(d)| .$$
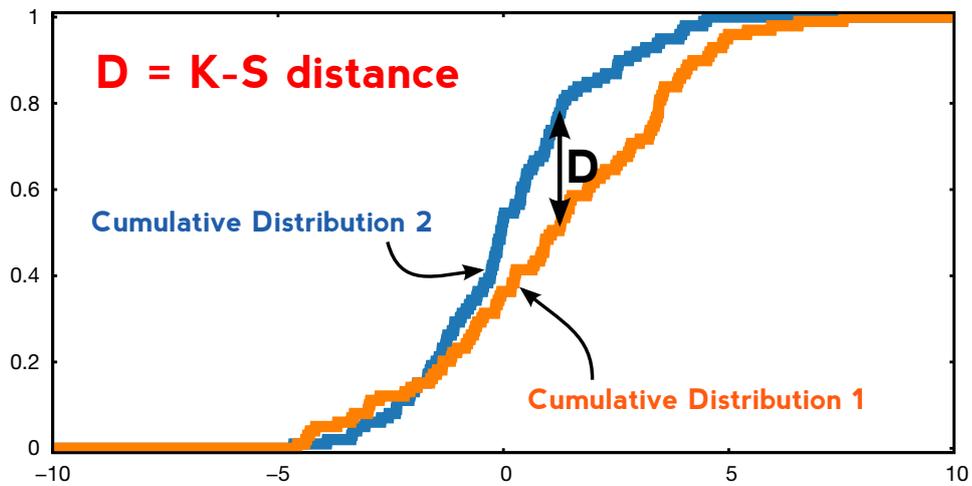
Figure 6.2: Kolmogorov–Smirnov distance between two cumulative distributions.

We assume the KS distance D to correspond to the potential gap between the two embedding points:

$$|P_{m_1, L_1} - P_{m_2, L_2}| \approx D_{(m_1, L_1) \leftrightarrow (m_2, L_2)} .$$

The approximation symbol is used because the cumulative distributions $F_{m_1, L_1}(d)$, $F_{m_2, L_2}(d)$, and consequently the KS distance $D_{(m_1, L_1) \leftrightarrow (m_2, L_2)}$, are sample estimates rather that population-like entities.

Unfortunately, the KS distance only measures the modulus of the potential gap between the two embeddings. The reconstruction of the potential–like surface $P_{m, L}$ is therefore similar to the problem of mapping the altitude of an area by using an altimeter that only provides the modulus of the

difference of altitude between two points of a grid.

Before proceeding with the reconstruction algorithm, it is proper to justify the chosen approach: first, the fact that an optimal embedding is identified by a minimum variation of a measure when adjacent points are considered, is a common trait of all methods proposed in the scientific literature; second, the measure must depend on the whole time series available; third, the measure should be linked to false-nearest-neighbours (as it is the case of Gao–Zheng method and the method of Schuster) and thus take into account distances.

6.3.2   Reconstruction

To reconstruct the potential on the lattice of Fig. 6.3, we first set the potential value at the corner point $(1, 0)$ to zero: $P_{1,0} = 0$.
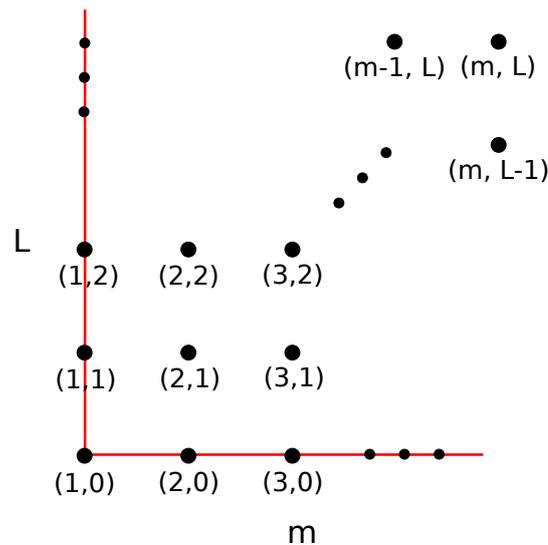
Figure 6.3: Lattice of embedding points (m, L), where m, L represent the embedding dimension and lag, respectively. The red line corresponds to the equipotential line with $P = 0$. The value $P_{m, L}$ is calculated from $P_{m, L-1}$, $P_{m-1, L}$ and the two KS distances $D_{(m, L) \leftrightarrow (m, L-1)}$, $D_{(m, L) \leftrightarrow (m-1, L)}$.

It is easy to show that the cumulative distribution $F_{(m, 0)}(d)$ for embedding vectors with lag $L = 0$ coincides with the cumulative distribution $F_{(1, 0)}(d)$ as all the elements of an embedding vector of any dimension m but

with zero lag are equal. Consequently, $P_{m,0} = P_{1,0}$, $\forall m$. In addition, since in an one-dimensional embedding the lag has no effect, we have $P_{1,L} = P_{1,0}$, $\forall L$. The bottom and left boundaries of the lattice are then equipotential lines for P, with $P = 0$ (see Fig. 6.3).

Starting from the point $(2, 1)$, the potential $P_{m,L}$ is iteratively evaluated from $P_{m,L-1}$, $P_{m-1,L}$, as follows. There are two possible estimates of $P_{m,L}$:

$$
\begin{aligned}
P'_{m,L} &\approx P_{m-1,L} + s_m\, D_{(m,L)\leftrightarrow(m-1,L)}\,, \\
P''_{m,L} &\approx P_{m,L-1} + s_L\, D_{(m,L)\leftrightarrow(m,L-1)}\,,
\end{aligned}
$$

where s_m, s_L are two signs, equal to either $+1$ or -1. The two signs are determined by choosing the combination that minimizes $|P'_{m,L} - P''_{m,L}|$. Once the values of s_m, s_L are given, the potential of $P_{m,L}$ is set to $[P'_{m,L} + P''_{m,L}]/2$. The algorithm relies on a sort of Stokes' theorem, as represented in Fig. 6.4.
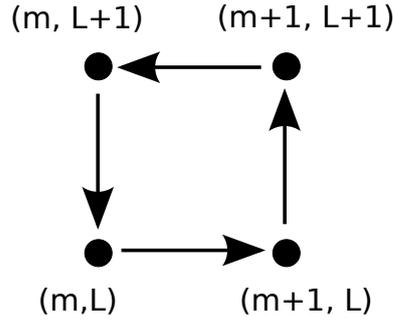
Figure 6.4: Similarly to Stokes' theorem, differences of the potential have to add up to zero in case of a closed loop.

For the very first point, namely $(2, 1)$, there are two possible values for the potential $P_{2,1}$, i.e. $\pm[P'_{m,L} + P''_{m,L}]/2$. However, since we are interested in the extrema of P, we can exploit this degree of freedom by setting the potential of the point $(2, 1)$ to be positive.

The final ingredient of the algorithm regards the multiplicity of maxima and minima and thus the choice of the optimal extremum. The crucial point is again the fact that the cumulative distributions of distances, and thus the KS distances, are estimates, and therefore uncertainty-affected. This uncertainty cumulates and affects the reconstructed value $P_{m,L}$ proportionally the

number of steps necessary to reach the point (m, L) from the point $(1, 0)$, namely $m+L-1$. For this reason, it appears reasonable to choose the optimal extremum as the closest one – in terms of the L_1, or city-block, distance – to the starting point $(1, 0)$.

We note that other algorithms could have been chosen to determine the potential P. For example, one could have taken triangles instead of unitary squares as basic mesh elements to iteratively evaluate P. However, the choice is the only one that implies the evaluation of distances D between adjacent embeddings.

Fig. 6.5 and Fig. 6.6 show the potential P reconstructed out of a sequence of 10^5 real numbers obtained by integrating the reference Mackey–Glass attractor Eq. (2.8) (see Sec. 2.5). According to our algorithm, the lattice point $(4, 5)$ is deemed to be the optimal embedding point for the attractor under study.

6.4 Results

Our Kolmogorov–Smirnov approach was applied to each of the four reference dynamical systems (see Sec. 2.5. The results are reported in Tab. 6.1. The embedding points generated by our approach are "good" in the sense discussed in the previous chapter. However, again as discussed in that chapter, the success of the MLE evaluation cannot be considered conclusive with regard to the issue of finding the best method to assess the optimal embedding. Nevertheless, our approach does not require any human interpretation, and seems therefore to be more reliable than the other methods described in the scientific literature.

Possible ways to improve our method are taking into account the distribution of the KS distance D, and implementing a reconstruction method that acts globally rather than locally, as it is presently the case.

Table 6.1: MLE, expressed in terms of μ_s and its uncertainty σ_s (see Sec. 5.1), for each of the four reference dynamical systems (see Sec. 2.5). Data are taken from the works available in the scientific literature, or computed by means of the techniques developed in this work. In particular, our data, highlighted in bold, were evaluated from a time series by using the divergent rate method (Sec. 4.1.2). MLE values taken from other works could not be verified; remarkably, in a case, the same authors (Gao and Zheng, with regard to the Rössler attractor) provides two different values. The time series sequence was embedded by using the parameters (m, L) deemed to be optimal by the different methods. Finally, "Gao–Zheng" corresponds to Gao–Zheng method, "Schuster" to method of Schuster, "CL" to the method of characteristic length, "GFNN-A" to the method of global false- nearest-neighbours and autocorrelation function, and "GFNN-MI" to the method of global false-nearest-neighbours and mutual information.

optimal embedding method	optimal (m, L)	MLE μ_s (σ_s)	
		original ref.	our calculation
Hénon			
Gao–Zheng [10]	(2, 1)	1.00(1)	**1.00(2)**
ours	(5, 1)		**1.00(2)**
Lorenz			
Gao–Zheng [10]	(3, 3)	1.20(2)	**0.94(1)**
ours	(5, 5)		**0.92(1)**
Rössler			
Gao–Zheng [10]	(3, 8)	0.75(7)	**0.840(7)**
Gao–Zheng [14]	(3, 8)	1.001	**0.840(7)**
Schuster [14]	(3, 9)	1.057	**0.833(6)**
CL [14]	(3, 8)	1.001	**0.840(7)**
GFNN-A [14]	(4, 9)	0.967	**0.858(6)**
GFNN-MI [14]	(4, 12)	0.979	not evaluated
ours	(5, 5)		**0.866(6)**
Mackey–Glass			
Gao–Zheng [10]	(4, 5)	0.97(8)	**1.21(1)**
Schuster [10]	(4, 5)	0.957	**1.21(1)**
ours	(4, 5)		**1.21(1)**
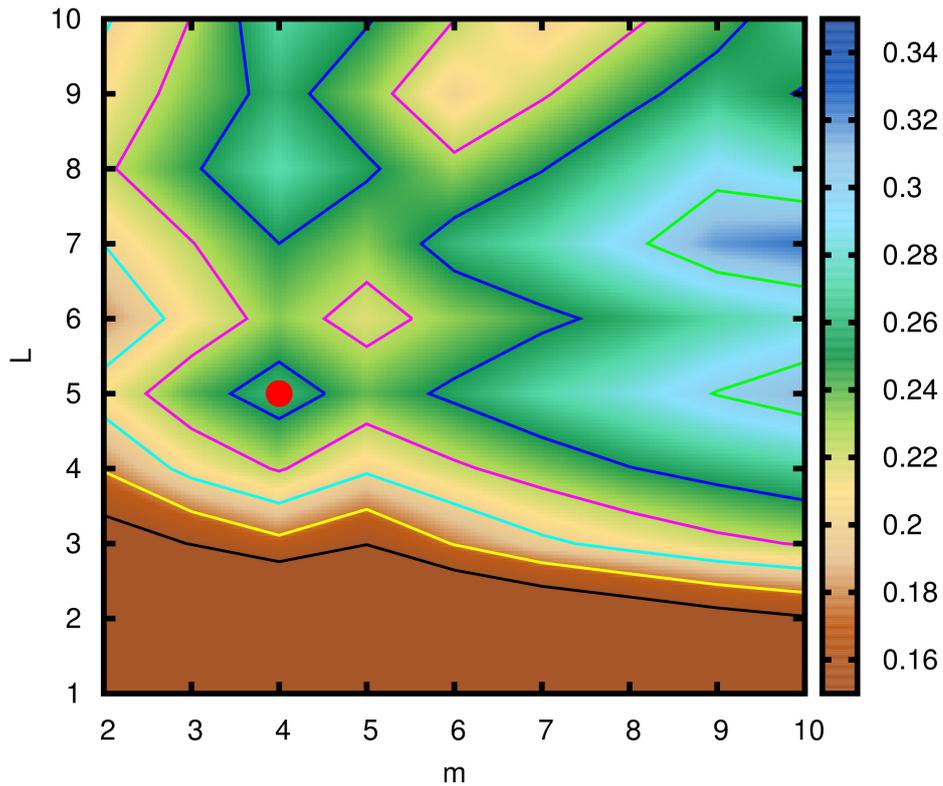
Figure 6.5: Reconstruction of the potential for a Mackey–Glass attractor (see Sec. 2.5). The embedding point (4, 5) is the local extremum that is "L_1-closest" to the starting point 1, 0. As such, according to our method, it corresponds to the optimal embedding point.
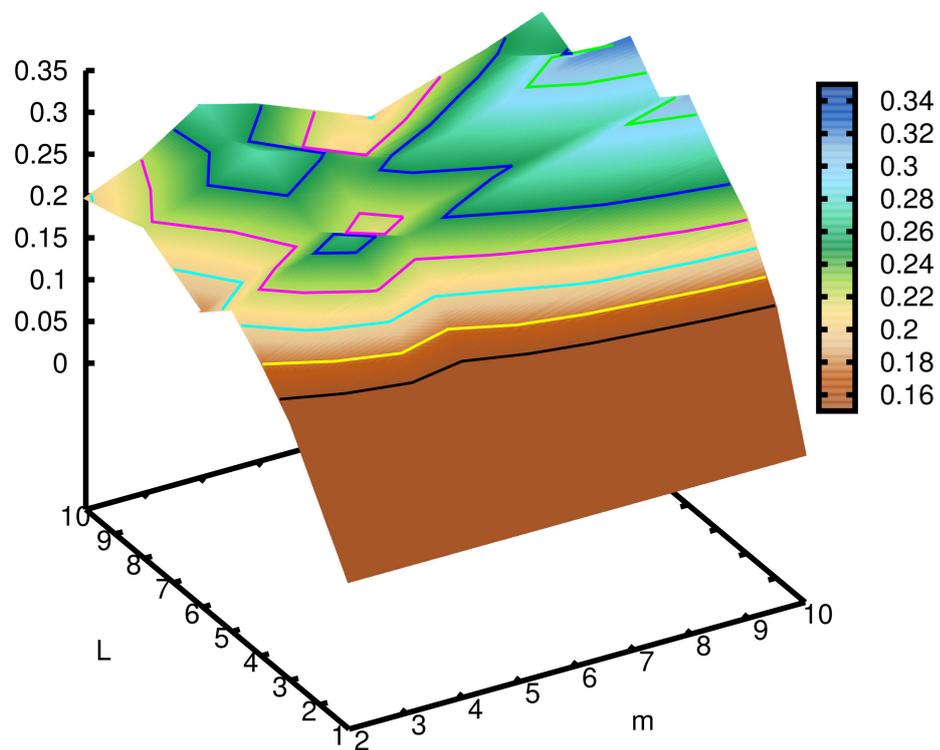
Figure 6.6: 3-D view of the reconstructed potential for the reference Mackey–Glass attractor of Fig. 6.5.

Chapter 7

A review of electroencephalography. Preliminary investigations on real EEG signals

This chapter deals with electroencephalographic (EEG) recordings and their analysis, with a focus on the chaotic behaviour of brain signals. The final part of the chapter describes the preliminary results obtained by applying the techniques presented in Chapters 4 and 5 to EEG signals recorded in a resting-state condition.

This chapter is organized as follows. In Sec. 7.1 the basic concepts as well as a historical overview of electroencephalography are presented. Sampling and clinical classification of EEG rhythms are the topic of Sec. 7.2. Sec. 7.3 presents pre-processing methods to remove artefacts from signals. Sec. 7.4 is an overview of nonlinear analysis applied to the EEG signals. Finally, Sec. 7.4 shows the first applications of the techniques developed in this work to the nonlinear analysis of EEG recordings.

7.1 Introduction

The brain is composed by glial cells and neurons that build up a network of connections; each neuron is connected to about 7000 other neurons, thus forming a highly complex system. The communication between two neurons is carried out by an ion current that flows from the axon of one cell to the dendrites of the other cell. In the last decade an enormous amount of resource and energy has been devoted to projects aimed at understanding brain functions and activities, and to find new ways to treat, cure, and prevent brain disorders. The current major projects are: the *Human brain project*, and the *BRAIN Initiative*. These two projects have been launched in 2013 and are located in Europe and in the United States, respectively. The objective of *Human brain project* is to simulate the brain and its activity by using a supercomputer [123]. *BRAIN Initiative* is a collaborative research initiative with the goal of mapping the activity of every neuron in the human brain [124].

Different techniques are used to collect data from the brain – for example EEG, magnetoencephalography (MEG) and functional magnetic resonance imaging (f-MRI). EEG and MEG are direct measurement of the current flowing in the brain mapped by using the secondary current and the magnetic field, respectively. f-MRI is an indirect measure based on the correlation between brain activity and blood flow: a neural activity leads to an increase of the blood flow to the area where the neural activity occurs; f-MRI looks at the variation of the oxygen concentration linked to the varying blood flow. The most used technique to analyse nonlinear dynamics in brain is electroencephalography. This is due to two main reasons: first, EEG has a temporal resolution of order 1 ms; second, it is less expensive and more "patient-friendly" than other techniques; the comfort of patients during the recordings is crucial to avoid stress to neural and cognitive functions and, consequently, to maintain the neural dynamics unchanged.

In 19th century, Carlo Matteucci and Emil Du Bois-Reymond registered the electrical signal emitted from muscle nerves, establishing the concept of neurophysiology. In 1875, the first recording of the brain activity was

carried out by Caton [125]. Until 1970, EEG interpretation was principally heuristic and qualitative [126, 127, 128], i.e. basically relying on the subjective interpretation of recordings. 1985 is a crucial year for the brain study, because of the publication of two seminal papers analysing EEG signals by means of nonlinear algorithms. Rapp *et al.* [129] described their results by using the words "chaos analysis" of spontaneous neural activity in the motor cortex of a monkey. Babloyantz *et al.* [3] reported the first observations of the correlation dimension in EEGs sampled during human sleep. These pioneering works were not only made possible by virtue of the availability of computers, but also because of the progress in the physics and mathematics of nonlinear dynamical systems [4]. Since then, a great number of clinical issues have been investigated using EEG signal analysis [130, 131]:

- monitoring alertness, coma, and brain death;

- locating areas of damage following head injury, stroke, and tumour;

- monitoring cognitive engagement;

- producing biofeedback situations;

- controlling anaesthesia depth;

- investigating epilepsy and locating seizure origin;

- testing epilepsy drug effects;

- assisting in experimental cortical excision of epileptic focus;

- monitoring the brain development;

- testing drugs for convulsive effects;

- investigating sleep disorders and physiology;

- investigating mental disorders;

- providing a hybrid data recording system together with other imaging modalities.

Studying the chaotic behaviour in EEG is supposed to be of help in identifying disorders [132] like epilepsy [133, 134], autism [135, 136, 137], schizophrenia, insomnia, dementia and in monitoring of anaesthesia depth. It is also considered a tool to understand the brain functions and to develop instrumentation that can be mind-controlled.

7.2 Sampling and clinical classification

An EEG signal is a time-domain measurement of the currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. When neurons are activated, the synaptic currents are produced within the dendrites. This current generates a secondary electrical field over the scalp measurable by EEG systems [131].

7.2.1 Tools to sample EEG signals

EEG recording systems consist of a number of electrodes, a set of differential amplifiers followed by filters, and a recording system. All these parts are crucial to minimize signal distortion. Presently, the signals are digitalized to further suppress noise and to perform data computational analysis and storage. The digitalization process is carried out by analogue-to-digital converters (ADC). Conventional setups can acquire signals with sampling frequencies up to 5 kHz. In order to maintain diagnostic information, the ADC resolution of EEG signals is at least 16 bit.

Electrodes make up a crucial part of the sampling setup. The most common types of electrodes are [131]:

- disposable (gel-less, and pre-gelled types);

- reusable disc electrodes (gold, silver, stainless steel, or tin);

- headbands and electrode caps;

- saline-based electrodes;

- needle electrodes.
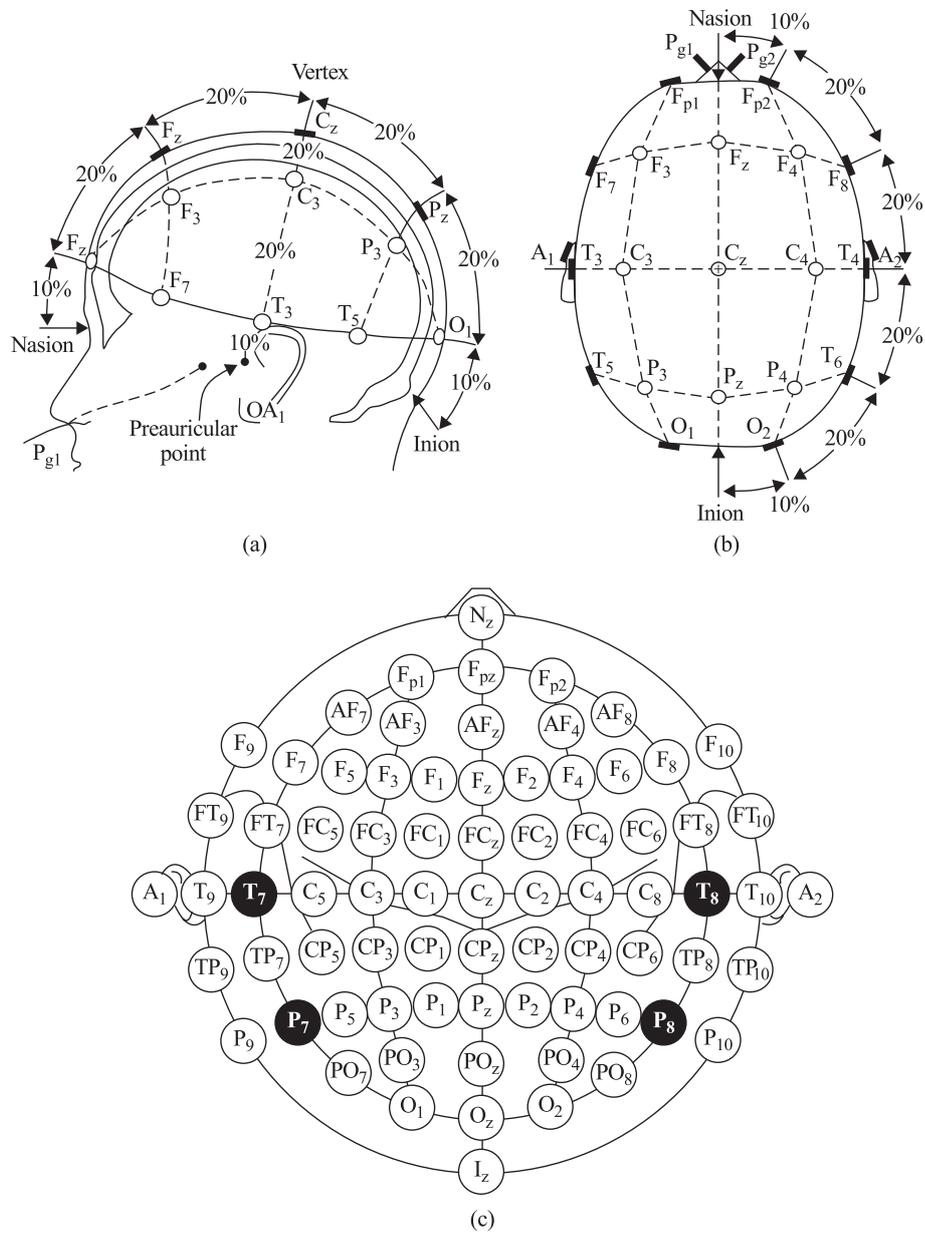
Figure 7.1: A scheme of 10-20 electrode settings for 75 electrodes including the reference electrodes. Source: [131].

The International Federation of Societies for Electroencephalography and Clinical Neurophysiology recommends a conventional electrode setting (also called 10-20) for 21 electrodes plus two reference electrodes connected to the earlobes. The 10-20 system considers the nasion and the inion positions as reference points and then uses 10 % or 20 % of the distance between these two points as the distance between the electrodes (see Fig. 7.1). The odd-labelled electrodes are on the left side of the scalp, whereas the even-labelled ones are located on the right side. If a larger number of electrodes is to be set, in compliance with the above conventional system the rest of the electrodes have to be placed between the previously allocated electrodes, in equidistant positions from them (see Fig. 7.1).

7.2.2   Classification of brain rhythms

Commonly, clinicians classify EEG waveforms into five classes or "waves": *delta* $(0-4\,\text{Hz})$, *theta* $(4-8\,\text{Hz})$, *alpha* $(8-12\,\text{Hz})$, *beta* $(13-30\,\text{Hz})$, and *gamma* $(30-60\,\text{Hz})$. A plot of these waves is reported in Figure 7.2. *Delta* waves are associated with deep sleep and are present in waking state. *Theta* waves are linked to unconscious material, creative inspiration and deep meditation. *Alpha* waves indicate a relaxed awareness without any attention or concentration. *Beta* waves are linked to active attention, thinking, focus on the outside world and solving concrete problems. *Gamma* waves are linked to brain diseases and synchronization of the brain.

7.3   Pre-processing analysis

Pre-processing of an EEG signal is a crucial step for successive analysis. This step consists in a series of analysis carried out to remove the artefacts from the recordings. Typical artefacts are eyes movement, eyes blinking, neck muscles contraction, movements of the head, environmental noise and perturbations and variations in the electrical contact between scalp and electrodes [138]. The most simplest method to detect artefacts is a threshold method that is applied both in the time domain and in the frequency domain
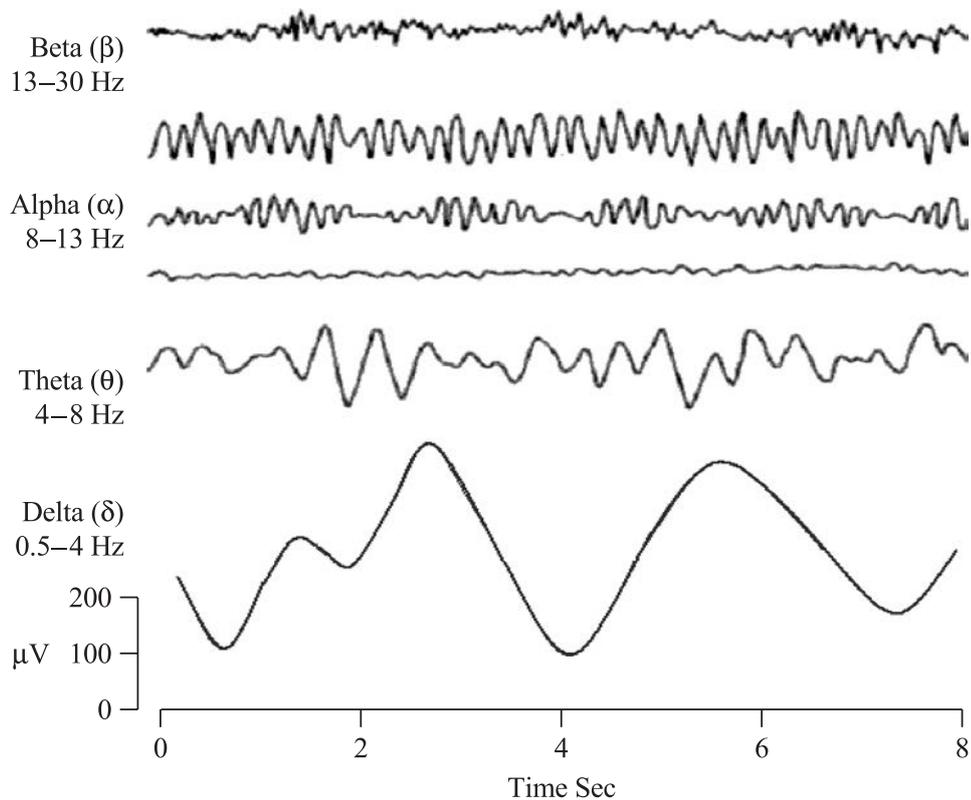
Figure 7.2: Typical dominant brain normal rhythms ("waves"), from high to low frequencies. Source: [131].

to detect bad electrodes and filter out environmental components, respectively. Another method to clean the signal is the parametric modelling [139], that is based on the assumption that the signal can be simulated by means of a mathematical model. Software tools to pre-process EEG signals are available. For example, EEGLAB [140] is an open-source MATLAB plug-in that collects a set of methods to remove artefacts.

One of the most used method to pre-processing EEG signals is the independent component analysis (ICA) [141, 140, 142, 138]. ICA assumes that a multi-dimensional EEG recording can be decomposed in independent components localized on the scalp. The localization gives us the possibility to filter out artefacts: for example, eyes movements can be removed by deleting the components located in the frontal part of the scalp. Because the location

plays a crucial role in separating and denoising signals, other methods to localize the sources of EEG recordings were developed; two examples are multiple signal classification [143, 144] and low resolution electromagnetic tomography algorithm [131].

In the last decade, techniques based on machine learning were used to carry out real time EEG analysis. Goals are real time decoding of the brain activity, development of brain-computer interfaces [145], and diagnosis [146].

7.4 Chaotic behaviour in brain

Nonlinear EEG analysis started in 1985 with Babloyantz *et al.* [3] and Rapp *et al.* [129]. In the analysis of EEG data, different chaotic measures are used in the scientific literature: fractal dimension, correlation dimension, Kolmogorov entropy, approximate entropy and maximum Lyapunov exponent (see Chapter 4).

In the first period of nonlinear EEG analysis, there was a premature enthusiasm for chaos underlying normal EEGs [147, 148, 149, 4]; it was only after some years that scientists realized that filtered noise can mimic low-dimensional chaos, so that those early claims were critically re-examined [150, 4]. Surrogate data testing methods were decisive to identify wrong interpretations. These methods rely of the generation of new data by randomly reshuffling the original ones. By applying the same analytical techniques to original data and to surrogate data and comparing the results, it is possible to assess whether the techniques can give information about the dynamics of the analysed systems.

Nonlinear approaches are a key ingredient for clinical assessments concerning brain functions, and particular epileptic brain states.

In the following, we report a summary of facts regarding EEG analysis and brain states.

7.4.1 Resting-state of healthy subjects

The analysis of this type of signal shows that there is no evidence of low-dimensional chaos in EEG [150]; however, at the same time, it is clear that EEG signals reflect a weak but significant nonlinear activity [151, 152, 153, 154].

7.4.2 Sleep

In 2014, Motamedi-Fakhr *et al.* [155] published a comprehensive review on signal processing techniques – both linear and nonlinear – applied to EEG recordings that were sampled in human sleep.

The first paper that studies the chaotic behaviour in EEG sleep recordings was published by Babloyantz *et al.* [3] in 1985. In 1994, Achermann *et al.* [156] analysed all night sleep recordings, finding evidence for a weak nonlinear structure. These results were confirmed by Fell *et al.* [157] two years later.

In 2002, Ferri *et al.* [158, 159] analysed sleep EEG signals of adults and young infants. They found that a nonlinear structure is present in the first class of subjects while it only sporadically occurs in the second one. This difference hints at the effect of ageing on the brain connections and underlines the crucial role of nonlinear analysis in understanding the brain functionality [4].

7.4.3 Anaesthesia, coma and vegetative state

Important results were obtained also for anaesthesia, coma and vegetative state. Nan and Jinghua [160] were the first to suggest a relation between changes in consciousness and the correlation dimension of EEG signals. Nonlinear methods can be used to measure and monitor the anaesthetic depth and coma [161]. Sarà and Pistoia [162, 163] studied the vegetative state and showed a complexity loss in EEG recordings of patients in vegetative state.

7.4.4 Epilepsy

Epilepsy is probably the most important application of nonlinear EEG analysis [4]. The first descriptions of nonlinear analysis applied to clinical epilepsy are the works by Iasemidis *et al.* [164, 165, 166]. They based the research on the assumption that a seizure represents a transition of the epileptic brain from a chaotic state to a more ordered one; therefore, dynamical properties are linked to the clinical state. By studying the maximum Lyapunov exponent for patients with temporal lobe epilepsy, they showed that the EEG activity becomes progressively less chaotic as the seizure approaches [164]. Since these pioneering studies, nonlinear methods have been used to quantify the changes in the brain dynamics before the onset of seizures, providing evidence to the hypothesis of a route to seizure. Lehnertz *et al.* [167] published a method to predict the seizures by using nonlinear analysis. Adeli *et al.* [168] proposed a method based on a wavelet-chaos methodology to detect epilepsy. In 2012, Acharya *et al.* [169, 170] described an automated method for the detection of epileptic subjects; the method relies on the assessment of entropy.

7.4.5 Normal cognition

Nonlinear EEG analysis is also applied to investigate the role of brain dynamics in cognitive processing. Several studies report an increase in the correlation dimension during cognitive tasks [171, 172, 173]; this phenomenon occurs in silent reading [174] and also in arithmetic tasks [173]. In 1998, Jeong *et al.* [175] showed that listening music induces changes in the brain dynamics: so, while the white noise increases brain complexity, the so-called $1/f$-noise type music decreases complexity. The fact that various cognitive tasks induce changes in the brain complexity and that the extent of these changes is linked to the task difficulty and the type of thinking involved raises the question whether EEG complexity might be related to intelligence [4].

7.4.6 Dementia

Jeong [176, 177] published an extensive review of nonlinear EEG analysis in dementia. In 1991, Pritchard *et al.* [178] showed that the increase in correlation dimension accompanying eye opening in non-demented subjects was diminished in Alzheimer patients; this behaviour was interpreted as a "lack of dynamical responsivity". Several studies dealing with correlations between nonlinear EEG measures and performance on neuropsychological tests [179] support the hypothesis of "complexity loss" that underlies cognitive dysfunctions in dementia. A chaotic dynamics is also present in Parkinson patients. Studies show that the maximum Lyapunov exponent measured on EEG recordings is higher in Parkinson patients than in Alzheimer patients [180, 181].

7.5 Analysis of an EEG signal

In this section, we apply the techniques developed in Chapters 4 and 5 to real EEG signals. The EEG recordings were sampled on healthy subjects in normal resting-state. The used sampling frequency is 5 kHz; for the following analysis we used the signal recorded from the electrode in position C_z (see Fig. 7.1).

We also analysed surrogate data sequences generated by randomly reshuffling the order of the EEG time series. Surrogate sequences have therefore the same average and the same amplitude of the original EEG signals. Fig. 7.3 shows the plateau value Λ_{pl}, obtained by using an embedding pair $(10, 4)$ (see Sec. 4.1.2), both for an EEG signal and its surrogate sequence. We can see that the plateau values for the EEG signal (circles) are larger than the respective values for the surrogate signal (crosses); these last are compatible with the behaviour predicted for a stochastic time series, as described in Sec. 4.1.2. Thus, we can conclude that there is a nonlinear dynamics in the EEG signal analysed.

We also carried out the finite-time MLE analysis presented in Chapter 5. Fig. 7.4 shows the diagram $\sigma_S(m, L)$ vs. $\Gamma_S(m, L)$. In this case the reference
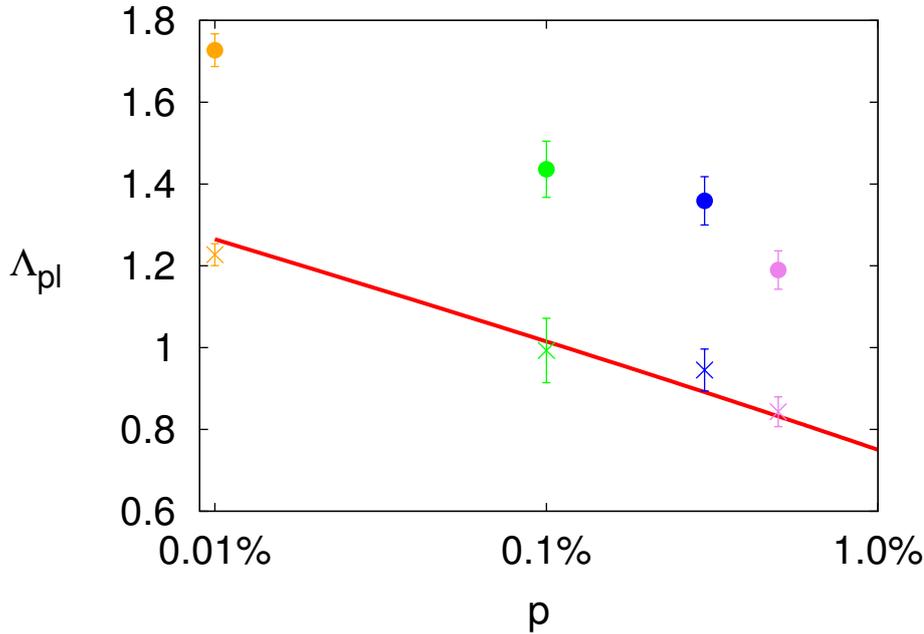
Figure 7.3: Λ_{pl} as a function of the percentile p: the circles are the plateau values Λ_{pl} obtained by analysing an EEG signal made of $3 \cdot 10^6$ samples. The crosses are the plateau values obtained by analysing the surrogate sequence. In both cases, the embedding pair is $(10, 4)$. The red line corresponds to $\Lambda_{pl}(10, p)$ given by Eq. (4.17).

maximum Lyapunov exponent is unknown; consequently, the $\Gamma_S(m, L)$ values cannot be normalized and no "good" values close to unity are given *a priori*. Fig. 7.4 shows a cluster for low $\Gamma_S(m, L)$ values. The sample mean and sample standard deviation of the Γ_S values of the points belonging to this cluster are 0.026 and 0.010, respectively. The last value corresponds to Σ_S. By using a level of significance equal to 0.01, we can conclude that Γ_S is compatible with 0. In other words, although we observe a nonlinearity in the analysed EEG signal by virtue of the Λ_{pl} analysis, we do not observe any low-dimensional chaotic dynamics, in agreement with the work by Theiler and Rapp [150] cited above (see Sec. 7.4.1).

Finally, we qualitatively compare our EEG signals with the signals used by Babloyantz *et al.* [3]. Fig. 7.5 shows a two-dimensional "phase portrait" – namely the two-dimensional plot of the vector $(x(t), x(t + \tau))$ – obtained by
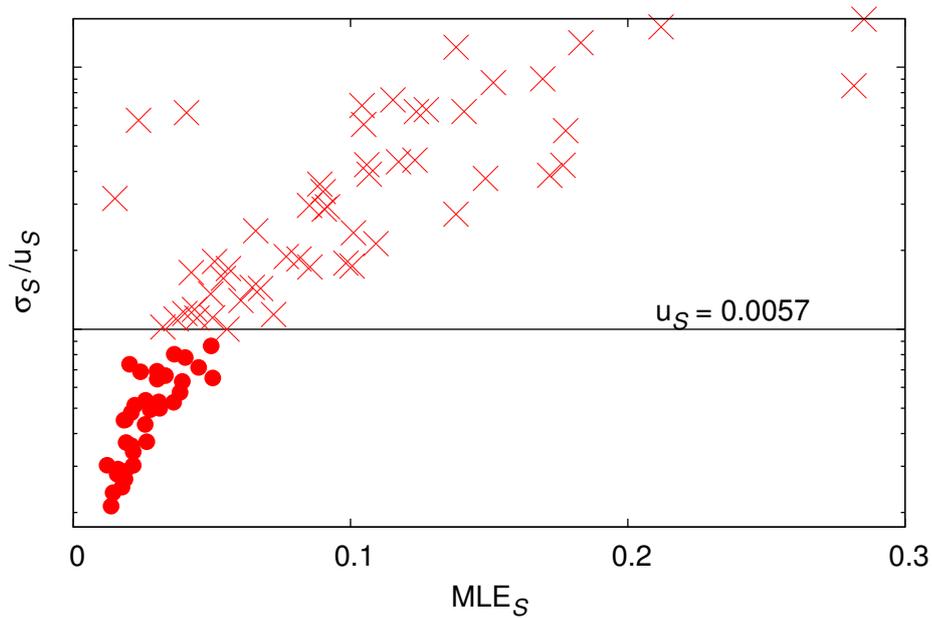
Figure 7.4: Diagram of the uncertainty $\sigma_S(m, L)$ (normalized to the respective u_S; see Sec. 5.3) vs. the MLE $\Gamma_S(m, L)$. Each point corresponds to an embedding pair. Black lines correspond to $\sigma = u_S$. Points below (above) the black lines are represented with dots (crosses). Ordinate axis are logarithmically scaled.

using the same delay shift τ as that one described in the work by Babloyantz *et al.*. In our case $\tau = 500\delta t = 0.1$ s, where $\delta t = 0.2$ ms is the sampling time; in the other work, $\tau = 10\delta t = 0.1$ s, where the declared sampling frequency is 100 Hz. Our phase portrait, Fig. 7.5, has a close resemblance to the phase portrait reported in Fig. 7.6(a) and obtained by Babloyantz *et al.* [3] under similar conditions.
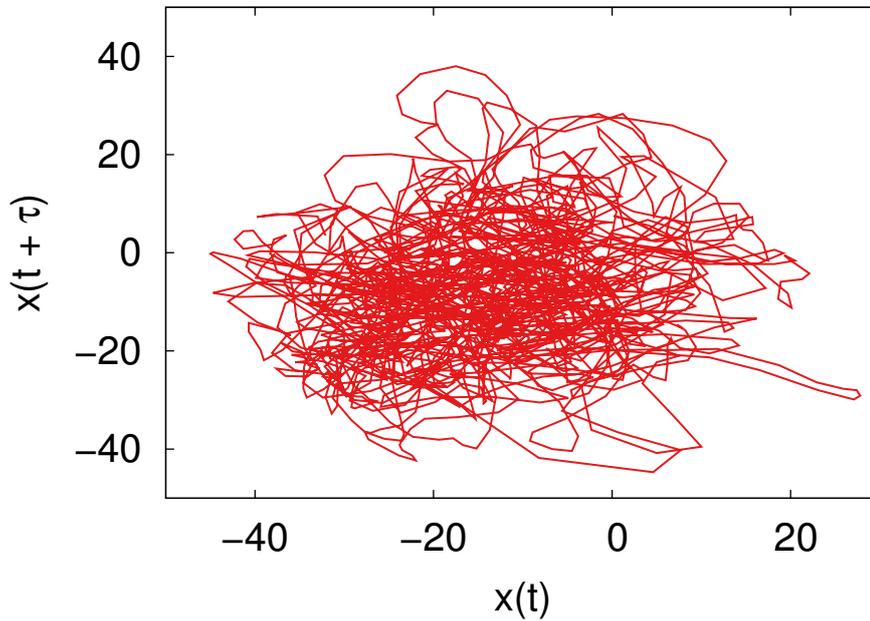
Figure 7.5: Two-dimensional phase portrait derived from an EEG signal by using a delay shift of 0.1 s.
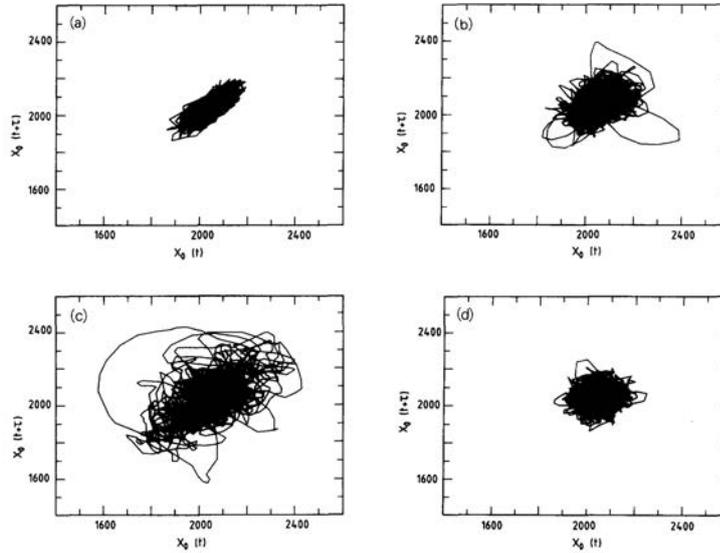
Figure 7.6: Two-dimensional phase portraits derived from EEGs of (a) an awake subject, (b) a sleep stage, (c) another sleep stage, (d) a REM sleep. The value of the delay shift is 0.1 s. Source: [3].

Chapter 8

Conclusion

In this PhD thesis, upon a discussion of the state of the art of the chaotic time series analysis based on the Takens embedding theorem, we have presented our three major results in this research field:

- a protocol based on the divergent rate method to calculate, from a time series, the finite-time maximum Lyapunov exponent and its uncertainty; other than previous methods, our approach is completely automatic;

- a statistical analysis of the finite-time maximum Lyapunov exponent, and its link to the existing theory [18];

- a new method based on a Kolmogorov–Smirnov approach to assess the optimal embedding pair.

In the work we have emphasized, and possibly better understood, the role of two measures: the uncertainty Σ_S with which the maximum Lyapunov exponent Γ_S can be estimated out of a time series by means of the divergence rate method (see Sec. 5.3), and the "plateau" value Λ_{pl} of the divergence exponent (see Sec. 4.1.2). These two measures turn out to be invariants of a dynamical system. Thus, they can be used to analyse and characterize its behaviour. The first measure is of help in identifying optimal embedding choices, or in excluding bad ones. The second measure allows to identify a purely stochastic behaviour behind a time series.

As stated in the Introduction, this work was motivated by the attempt to answer the question on whether noise within a neural system – and in particular the brain – has a functional role. This issue might be tackled by considering electroencephalographic (EEG) recordings: besides its robustness, a major motivation for the choice of this technique is its noninvasiveness and the easiness of data collection. Although most of this work was devoted to synthetic time series, in its last phase, as described in Chapter 7, we have started applying the developed methods to analyse real EEG recordings. The analytical techniques presented in this study, along with traditional ones like correlation dimension, maximum Lyapunov exponent and Kolmogorov entropy, could open new perspectives and insights in the investigation of the brain dynamics and, possibly, help tackling unsolved functional and diagnostic issues in disorders and diseases like autism, epilepsy, obsessive-compulsive disorder as well as vegetative state.

This line of action coincides with that one indicated one decade ago by Stam [4] who asserted that the future of nonlinear EEG analysis depends on the progress in three main directions: development of better tools for nonlinear time series analysis, a better theoretical understanding of the dynamics of normal and pathological brain states, and clinical application of nonlinear analysis to problems like seizure anticipation and detection/diagnosis of psychiatric and neurological disorders. We believe our work is a contribution to the quest of understanding the intimate operation of complex neural systems.

Another perspective of the work presented in this thesis regards the issue of disentangling noise from chaos. A first step consists in studying the dependence of measures like Γ_S, Σ_S, Λ_{pl} on dynamical or additive noise, i.e. a noise added at the level of the differential equations that describe a chaotic system or, respectively, in the integral ones. Besides making up a very interesting topic in the field of nonlinear and complex systems, the development of a reliable approach to efficiently disentangle deterministic components from stochastic ones could be of great interest in the understanding of neural dynamics.

A final, more applicative perspective regards the implementation of nonlinear techniques and algorithms in embedded systems with high speed and performance. In this way, real time analysis could be carried out, thus removing the gap between theory and its application. For example, the real time implementation of the new method based on a Kolmogorov–Smirnov approach to calculate the optimal embedding pair could make up an important tool to perform real time analysis of experimental time series.

Bibliography

[1] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70(1):223–287, 1998.

[2] D. Tabarelli, A. Vilardi, C. Begliomini, F. Pavani, M. Turatto, and L. Ricci. Statistically robust evidence of stochastic resonance in human auditory perceptual system. *The European Physical Journal B*, 69(1):155–159, 2009.

[3] A. Babloyantz, J. M. Salazar, and C. Nicolis. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111(3):152–156, 1985.

[4] C. J. Stam. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301, 2005.

[5] J. B. Gao. Recognizing randomness in a time series. *Physica D: Nonlinear Phenomena*, 106(1-2):49–56, 1997.

[6] F. Takens. Detection strange attractor in turbulence. *Lecture Notes in Mathematica*, 898:366–381, 1981.

[7] J. Stark, D. S. Broomhead, M. E. Davies, and J. P. Huke. Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis: Theory, Methods & Applications*, 30(8):5303–5314, 1997.

[8] J. B. Gao and Z. M. Zheng. Local exponential divergence plot and optimal embedding of a chaotic time-series. *Physics Letters A*, 181(2):153–158, 1993.

[9] J. B. Gao and Z. Zheng. Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time-series. *Physical Review E*, 49(5, Part A):3807–3814, 1994.

[10] J. B. Gao and Z. Zheng. Direct dynamical test for deterministic chaos. *Europhysics Letters*, 25(7):485–490, 1994.

[11] W Liebert, K Pawelzik, and H. G Schuster. Optimal Embeddings of Chaotic Attractors from Topological Considerations. *Europhysics Letters*, 14(6):521–526, 1991.

[12] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring. The analysis of observed chaotic data in physical systems. *Reviews of modern physics*, 65(4):1331–1392, 1993.

[13] H. D. I. Abarbanel. *Analysis of observed chaotic data*, volume 39. 1996.

[14] C. J. Cellucci, A. M. Albano, and P. E. Rapp. Comparative study of embedding methods. *Physical Review E*, 67(6):66210, 2003.

[15] G. Benettin, L. Galgani, A. Giorgilli, and J. M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica*, 15:9–20, 1980.

[16] G. Benettin, L. Galgani, A. Giorgilli, and J. M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 2: Numerical application. *Meccanica*, 15:21–30, 1980.

[17] M. Franchi and L. Ricci. Statistical properties of the maximum Lyapunov exponent calculated via the divergence rate method. *Physical Review E*, 90(6):062920, 2014.

[18] P. Grassberger, R. Badii, and A. Politi. Scaling laws for invariant measures on hyperbolic and nonhyperbolic atractors. *Journal of Statistical Physics*, 51(1-2):135–178, 1988.

[19] M. Gardner. Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123., 1970.

[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 3rd edition, 2007.

[21] C. Li and G. Chen. Estimating the Lyapunov exponents of discrete systems. *Chaos*, 14:343–346, 2004.

[22] B. Aulbach and B. Kieninger. On three definitions of chaos. *Nonlinear Dyn. Syst. Theory*, 1(1):23–37, 2001.

[23] T. Y. Li and J. A. Yorke. Period three implies chaos. *The American Mathematical Monthly*, 82(10):985–992, 1975.

[24] L. S. Block and W. A. Coppel. *Dynamics in one dimension*. Springer, 1992.

[25] R. L. Devaney. *An introduction to chaotic dynamical systems*, volume 13046. Addison-Wesley Reading, 1989.

[26] Y. Hirata and K. Aihara. Devaney's chaos on recurrence plots. *Physical Review E*, 82(3):036209, 2010.

[27] Ch. Skokos. The Lyapunov Characteristic Exponents and Their Computation. *Dynamics of Small Solar System Bodies and Exoplanets*, 790:63–135, 2010.

[28] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey. On Devaney's definition of chaos. *American Mathematical Monthly*, pages 332–334, 1992.

[29] E. Lorenz. Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas? *American Association for the Advancement of Science, Washington, DC*, 1972.

[30] E. Ott. *Chaos in Dynamical Systems - 2nd Edition*. 2002.

[31] M. Hénon. A two-dimensional mapping with a strange attractor. *Communications in Mathematical Physics*, 50(1):69–77, 1976.

[32] P. Cvitanović, G. Gunaratne, and I. Procaccia. Topological and metric properties of Hénon-type strange attractors. *Physical Review A*, 38(3):1503–1520, 1988.

[33] M. Michelitsch and O. E. Rössler. A new feature in Hénon's map. *Computers & Graphics*, 13(2):263–265, 1989.

[34] M. Benedicks and L. Carleson. The Dynamics of the Henon Map. *Annals of Mathematics*, 133(1):pp. 73–169, 1991.

[35] W. F. H. Al-Shameri. Dynamical properties of the H{é}non mapping. *Int. J. Math. Anal.*, 6:2419–2430, 2012.

[36] Robert M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467, 1976.

[37] P. F. Verhulst. La Loi d'Accroissement de la Population. *Nouv. Mem. Acad. Roy. Soc. Belle-lettr. Bruxelles*, 18:1, 1845.

[38] S. Phatak and S. Rao. Logistic map: A possible random-number generator. *Physical Review E*, 51(4):3670–3678, 1995.

[39] M. Andrecut. Logistic Map as a Random Number Generator. *International Journal of Modern Physics B*, 12(09):921–930, 1998.

[40] N. K. Pareek, V. Patidar, and K. K. Sud. Discrete chaotic cryptography using external key. *Physics Letters A*, 309(1-2):75–82, 2003.

[41] N. K. Pareek, V. Patidar, and K. K. Sud. Image encryption using chaotic logistic map. *Image and Vision Computing*, 24(9):926–934, 2006.

[42] K. Ikeda. Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system. *Optics Communications*, 30(2):257–261, 1979.

[43] K. Ikeda, H. Daido, and O. Akimoto. Optical Turbulence: Chaotic Behavior of Transmitted Light from a Ring Cavity. *Physical Review Letters*, 45(9):709–712, 1980.

[44] A. P. Kuznetsov, L. V. Turukina, and E. Mosekilde. Dynamical systems of different classes as models of the kicked nonlinear oscillator. *International Journal of Bifurcation and Chaos*, 11(4):1065, 2001.

[45] A. P. Kuznetsov, A. V. Savin, and D. V. Savin. On some properties of nearly conservative dynamics of Ikeda map and its relation with the conservative case. *Physica A: Statistical Mechanics and its Applications*, 387(7):1464–1474, 2008.

[46] H. D. I. Abarbanel, R. Brown, and M. B. Kennel. Local Lyapunov exponents computed from observed data. *Journal of Nonlinear Science*, 2(3):343–365, 1992.

[47] E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.

[48] H. Haken. Analogy between higher instabilities in fluids and lasers. *Physics Letters A*, 53(1):77–78, 1975.

[49] H. Haken. Pattern Formation and Chaos in Synergetic Systems. *Physica Scripta*, 1985(T9):111, 1985.

[50] E. Knobloch. Chaos in the segmented disc dynamo. *Physics Letters A*, 82(9):439–440, 1981.

[51] I. M. Moroz. The Extended Malkus–Robbins Dynamo as a Perturbed Lorenz System. *Nonlinear Dynamics*, 41(1):191–210, 2005.

[52] M. Gorman, P. J. Widmann, and K. A. Robbins. Nonlinear dynamics of a convection loop: A quantitative comparison of experiment with theory. *Physica D: Nonlinear Phenomena*, 19(2):255–267, 1986.

[53] N. Hemati. Strange attractors in brushless DC motors. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 41(1):40–45, 1994.

[54] K. Cuomo and A. Oppenheim. Circuit implementation of synchronized chaos with applications to communications. *Physical Review Letters*, 71(1):65–68, 1993.

[55] D. Poland. Cooperative catalysis and chemical chaos: a chemical model for the Lorenz equations. *Physica D: Nonlinear Phenomena*, 65(1-2):86–99, 1993.

[56] S. I. Tzenov. Strange Attractors Characterizing the Osmotic Instability. page 6, 2014.

[57] O. E Rössler. An equation for continuous chaos. *Physics Letters A*, 57(5):397–398, 1976.

[58] M. C. Mackey and L. Glass. Oscillation and chaos in physiological control systems. *Science*, 197(4300):287–289, 1977.

[59] A. Namajunas, K. Pyragas, and A. Tamasevicius. An electronic analog of the Mackey-Glass system. *Physics Letters A*, 201(1):42–46, 1995.

[60] C. Hayashi. *Nonlinear oscillations in physical systems*. Princeton University Press, 2014.

[61] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 1983.

[62] N. Q. Hu and X. S. Wen. The application of Duffing oscillator in characteristic signal detection of early fault. *Journal of Sound and Vibration*, 268(5):917–931, 2003.

[63] C. Wang and S. S. Ge. Synchronization of two uncertain chaotic systems via adaptive backstepping. *International Journal of Bifurcation and Chaos*, 11(06):1743–1751, 2001.

[64] S. Bowong. Stability analysis for the synchronization of chaotic systems with different order: application to secure communications. *Physics Letters A*, 326(1-2):102–113, 2004.

[65] H. Fotsin, S. Bowong, and J. Daafouz. Adaptive synchronization of two chaotic systems consisting of modified Van der Pol–Duffing and Chua oscillators. *Chaos, Solitons & Fractals*, 26(1):215–229, 2005.

[66] L. J. Sheu, H. K. Chen, J. H. Chen, and L. M. Tam. Chaotic dynamics of the fractionally damped Duffing equation. *Chaos, Solitons & Fractals*, 32(4):1459–1468, 2007.

[67] L. O. Chua. Chua's circuit: ten years later. *IEICE Transactions on Fundamentals*, E77-A(11):1811–1822, 1994.

[68] C. C. Hwang, H. Y. Chow, and Y. K. Wang. A new feedback control of a modified Chua's circuit system. *Physica D: Nonlinear Phenomena*, 92(1-2):95–100, 1996.

[69] M. T. Yassen. Adaptive control and synchronization of a modified Chua's circuit system. *Applied Mathematics and Computation*, 135(1):113–128, 2003.

[70] T. Wu and M. S. Chen. Chaos control of the modified Chua's circuit system. *Physica D: Nonlinear Phenomena*, 164(1-2):53–58, 2002.

[71] G.P. Jiang, W. K.S. Tang, and G. Chen. A simple global synchronization criterion for coupled chaotic systems. *Chaos, Solitons & Fractals*, 15(5):925–935, 2003.

[72] J. Zhang, C. Li, H. Zhang, and J. Yu. Chaos synchronization using single variable feedback based on backstepping method. *Chaos, Solitons & Fractals*, 21(5):1183–1193, 2004.

[73] E. P. dos Santos, M. S. Baptista, and I. L. Caldas. Dealing with final state sensitivity for synchronous communication. *Physica A: Statistical Mechanics and its Applications*, 308(1-4):101–112, 2002.

[74] Z. Li, K. Li, C. Wen, and Y. C. Soh. A new chaotic secure communication system. *IEEE Transactions on Communications*, 51(8):1306–1312, 2003.

[75] K. Li, Y.C. Soh, and Z.G. Li. Chaotic cryptosystem with high sensitivity to parameter mismatch. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 50(4):579–583, 2003.

[76] H. A. Albuquerque, R. M. Rubinger, and P. C. Rech. Theoretical and experimental time series analysis of an inductorless Chua's circuit. *Physica D: Nonlinear Phenomena*, 233(1):66–72, 2007.

[77] A. Zeni, J. A. C. Gallas, A. Fioretti, F. Papoff, B. Zambon, and E. Arimondo. Lyapunov exponents and return maps for a model of a laser with saturable absorber. *Physics Letters A*, 172(4):247–255, 1993.

[78] H. F. Liu, Z. H. Dai, W. F. Li, X. Gong, and Z. H. Yu. Noise robust estimates of the largest Lyapunov exponent. *Physics Letters A*, 341(1-4):119–127, 2005.

[79] Y. Hirata, T. Yamada, J. Takahashi, and H. Suzuki. Real-time multi-step predictors from data streams. *Physics Letters A*, 376(45):3092–3097, 2012.

[80] J. B. Gao, J. Hu, X. Mao, and W. W. Tung. Detecting low-dimensional chaos by the "noise titration" technique: Possible problems and remedies. *Chaos, Solitons & Fractals*, 45(3):213–223, 2012.

[81] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, and F. Rossi. GNU Scientific Library Reference Manual - Third Edition (v1.12). *Network Theory Ltd*, 83, 2009.

[82] P. J. Prince and J. R. Dormand. High order embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 7(1):67–75, 1981.

[83] E. Hairer, S. P. Nø rsett, and G. Wanner. *Solving Ordinary Differential Equations I Nonstiff problems*. Springer, Berlin, second edition, 2000.

[84] J. R. Dormand and P. J. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, 1980.

[85] S. A. Sarra and C. Meador. On the numerical solution of chaotic dynamical systems using extend precision floating point arithmetic and very high order numerical methods. *Nonlinear Analysis: Modelling and Control*, 16(3):340–352, 2011.

[86] H. Whitney. Differentiable Manifolds. *Annals of Mathematics*, 37(3):645–680, 1936.

[87] L. Noakes. The Takens Embedding Theorem. *International Journal of Bifurcation and Chaos*, 01(04):867–872, 1991.

[88] T. Sauer, J. A. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65(3):579–616, 1991.

[89] T. Schreiber. Interdisciplinary application of nonlinear time series methods. *Physics Reports*, 308(1):1–64, 1999.

[90] R. Hegger, H. Kantz, and T. Schreiber. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos*, 9(2):413–435, 1999.

[91] M. R. Muldoon, D. S. Broomhead, J. P. Huke, and R. Hegger. Delay embedding in the presence of dynamical noise. *Dynamics and Stability of Systems*, 13(2):175–186, 1998.

[92] E. Monroig, K. Aihara, and Y. Fujino. Modeling dynamics from only output data. *Physical Review E*, 79(5):056208, 2009.

[93] J. Stark. Delay Embeddings for Forced Systems. I. Deterministic Forcing. *Journal of Nonlinear Science*, 9(3):255–332, 1999.

[94] J. Stark, D. S. Broomhead, M. E. Davies, and J. P. Huke. Delay Embeddings for Forced Systems. II. Stochastic Forcing. *Journal of Nonlinear Science*, 13(6):519–577, 2003.

[95] Th. Buzug and G. Pfister. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Physica D: Nonlinear Phenomena*, 58(1-4):127–137, 1992.

[96] M. T. Rosenstein, J. J. Collins, C. J. De Luca, and C. J. D. Luca. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1-2):117–134, 1993.

[97] H. Kantz. A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics Letters A*, 185(1):77–87, 1994.

[98] M. Franchi and L. Ricci. Appropriateness of dynamical systems for the comparison of different embedding methods via calculation of the maximum Lyapunov exponent. *Journal of Physics: Conference Series*, 490(1):012094, 2014.

[99] A. M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3):531–534, 1992.

[100] M. Hénon and C. Heiles. The applicability of the third integral of motion: Some numerical experiments, 1964.

[101] J. Ford and G. Lunsford. Stochastic Behavior of Resonant Nearly Linear Oscillator Systems in the Limit of Zero Nonlinear Coupling. *Physical Review A*, 1(1):59–70, 1970.

[102] S. Stoddard and J. Ford. Numerical Experiments on the Stochastic Behavior of a Lennard-Jones Gas System. *Physical Review A*, 8(3):1504–1512, 1973.

[103] H. Kantz, G. Radons, and H. Yang. The problem of spurious Lyapunov exponents in time series analysis and its solution by covariant Lyapunov vectors. *Journal of Physics A: Mathematical and Theoretical*, 46(25):254009, 2013.

[104] I. Shimada and T. Nagashima. A Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems, 1979.

[105] J. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617–656, 1985.

[106] K. Ramasubramanian and M. S. Sriram. A comparative study of computation of Lyapunov spectra with different algorithms. *Physica D: Nonlinear Phenomena*, 139(1-2):72–86, 2000.

[107] J. Theiler. Spurious dimension from correlation algorithms applied to limited time-series data. *Physical Review A*, 34(3):2427–2432, 1986.

[108] J. Theiler. Estimating fractal dimension. *Journal of the Optical Society of America A*, 7(6):1055, 1990.

[109] J. D. Farmer. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D: Nonlinear Phenomena*, 4(3):366–393, 1982.

[110] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*, volume 27. 1949.

[111] P. Grassberger and I. Procaccia. Characterization of Strange Attractors. *Physical Review Letters*, 50(5):346–349, 1983.

[112] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208, 1983.

[113] J. L. Kaplan and J. A. Yorke. Chaotic behavior of multidimensional difference equations. In *Functional differential equations and approximation of fixed points*, pages 204–227. Springer, 1979.

[114] M. Sano and Y. Sawada. Measurement of the Lyapunov Spectrum from a Chaotic Time Series. *Physical Review Letters*, 55(10):1082–1085, 1985.

[115] A. N. Kolmogorov. New Metric Invariant of Transitive Dynamical Systems and Endomorphisms of Lebesgue Spaces. In *Doklady of Russian Academy of Sciences*, volume 119, pages 861–864, 1958.

[116] H. G. Schuster and W. Just. *Deterministic chaos: an introduction*. John Wiley & Sons, 2006.

[117] P. W. Milonni, M. L. Shih, and J. R. Ackerhalt. *Chaos in laser-matter interactions*, volume 6. World Scientific, 1987.

[118] Y. B. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Mathematical Surveys*, 32(4):55–114, 1977.

[119] D. Ruelle. *Chaotic evolution and strange attractors*, volume 1. Cambridge University Press, 1989.

[120] A. Prasad and R. Ramaswamy. Characteristic distributions of finite-time Lyapunov exponents. *Physical Review E*, 60(3):2761–2766, 1999.

[121] S. S. Shapiro and M. B. Wilk. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591–611, 1965.

[122] K. Kanno and A. Uchida. Finite-time Lyapunov exponents in time-delayed nonlinear dynamical systems. *Physical Review E*, 89(3):032918, 2014.

[123] Human brain project. `www.humanbrainproject.eu`.

[124] BRAIN Initiative. `www.braininitiative.nih.gov`.

[125] R. Caton. The Electric Currents of the Brain. *Brit. Med. J.*, 2:278, 1875.

[126] B. Hjorth. EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3):306–310, 1970.

[127] N. Kannathal, U. R. Acharya, C. M. Lim, and P. K. Sadasivan. Characterization of EEG–a comparative study. *Computer methods and programs in biomedicine*, 80(1):17–23, 2005.

[128] O. Faust and M. G. Bairy. Nonlinear analysis of physiological signals: a review. *Journal of Mechanics in Medicine and Biology*, 12(04):1240015, 2012.

[129] P. E. Rapp, I. D. Zimmerman, A. M. Albano, G. C. Deguzman, and N. N. Greenbaun. Dynamics of spontaneous neural activity in the simian motor cortex: The dimension of chaotic neurons. *Physics Letters A*, 110(6):335–338, 1985.

[130] M. Teplan. Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11, 2002.

[131] S. Sanei and J. A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.

[132] P. Faure and H. Korn. Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *Life Sciences*, 324(9):773–793, 2001.

[133] K. Lehnertz and C. E. Elger. Can Epileptic Seizures be Predicted? Evidence from Nonlinear Time Series Analysis of Brain Electrical Activity. *Physical Review Letters*, 80(22):5019–5022, 1998.

[134] J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, and F. J. Varela. Epileptic seizures can be anticipated by non-linear analysis. *Nature Medicine*, 4(10):1173–6, 1998.

[135] R. Kulisek, Z. Hrncir, M. Hrdlicka, L. Faladova, K. Sterbova, P. Krsek, E. Vymlatilova, M. Palus, A. Zumrová, and V. Komárek. Nonlinear analysis of the sleep EEG in children with pervasive developmental disorder. *Neuro endocrinology letters*, 29(4):512–7, 2008.

[136] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson. EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine*, 9(1):18, 2011.

[137] S. Bhat, U. R. Acharya, H. Adeli, G. M. Bairy, and A. Adeli. Automated diagnosis of autism: in search of a mathematical marker. *Reviews in the neurosciences*, 25(6):851–61, 2014.

[138] A. Delorme, T. Sejnowski, and S. Makeig. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34(4):1443–9, 2007.

[139] J. Pardey, S. Roberts, and L. Tarassenko. A review of parametric modelling techniques for EEG analysis. *Medical Engineering & Physics*, 18(1):2–11, 1996.

[140] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[141] A. J. Bell and T. J. Sejnowski. An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[142] L. Spyrou, M. Jing, S. Sanei, and A. Sumich. Separation and localisation of P300 sources and their subcomponents using constrained blind source separation. *Eurasip Journal on Advances in Signal Processing*, 2007, 2007.

[143] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[144] J. C. Mosher and R. M. Leahy. Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE Transactions on Biomedical Engineering*, 45(11):1342–54, 1998.

[145] K. R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and Benjamin Blankertz. Machine learning for real-time single-trial EEG-analysis: from brain-computer interfacing to mental state monitoring. *Journal of Neuroscience Methods*, 167(1):82–90, 2008.

[146] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki. A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *NeuroImage*, 49(1):668–93, 2010.

[147] A. Babloyantz and A. Destexhe. Strange Attractors in the Human Cortex. In *Temporal Disorder in Human Oscillatory Systems*, volume 36 of *Springer Series in Synergetics*, pages 48–56. Springer Berlin Heidelberg, 1987.

[148] W. S. Pritchard and D. W. Duke. Dimensional Analysis of No-Task Human EEG Using the Grassberger-Procaccia Method. *Psychophysiology*, 29(2):182–192, 1992.

[149] P. E. Rapp, T. R. Bashore, J. M. Martinerie, A. M. Albano, I. D. Zimmerman, and Alistair I. Mees. Dynamics of brain electrical activity. *Brain Topography*, 2(1-2):99–118, 1989.

[150] J. Theiler and P. E. Rapp. Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, 98(3):213–222, 1996.

[151] W. S. Pritchard, D. W. Duke, and K. K. Krieble. Dimensional analysis of resting human EEG II: Surrogate-data testing indicates nonlinearity but not low-dimensional chaos. *Psychophysiology*, 32(5):486–491, 1995.

[152] G. L. Gebber, S. Zhong, C. Lewis, and S. M. Barman. Human brain alpha rhythm: nonlinear oscillation or filtered noise? *Brain Research*, 818(2):556–560, 1999.

[153] D. Maurice, R. Cerf, and M. Toussaint. Automated detection of low-dimensional EEG alpha-episodes. An example of application to psychopharmacological data. *Computer methods and programs in biomedicine*, 68(2):93–108, 2002.

[154] T. Gautama, D. Mandic, and M. Van Hulle. Indications of nonlinear structures in brain electrical activity. *Physical Review E*, 67(4):046204, 2003.

[155] S. Motamedi-Fakhr, Mo. Moshrefi-Torbati, M. Hill, C. M. Hill, and P. R. White. Signal processing techniques applied to human sleep EEG

signals—A review. *Biomedical Signal Processing and Control*, 10:21–33, 2014.

[156] P. Achermann, R. Hartmann, A. Gunzinger, W. Guggenbüh, and A. A. Borbély. Correlation Dimension of the Human Sleep Electroencephalogram: Cyclic Changes in the Course of the Night. *European Journal of Neuroscience*, 6(3):497–500, 1994.

[157] J. Fell, J. Röschke, and C. Schäffner. Surrogate data analysis of sleep electroencephalograms reveals evidence for nonlinearity. *Biological Cybernetics*, 75(1):85–92, 1996.

[158] R. Ferri, L. Parrino, A. Smerieri, M. G. Terzano, M. Elia, Sebastiano A. Musumeci, Salvatore Pettinato, and Cornelis J. Stam. Nonlinear EEG measures during sleep: effects of the different sleep stages and cyclic alternating pattern. *International Journal of Psychophysiology*, 43(3):273–286, 2002.

[159] R. Ferri, R. Chiaramonti, M. Elia, S. A. Musumeci, A. Ragazzoni, and Cornelis J Stam. Nonlinear EEG analysis during sleep in premature and full-term newborns. *Clinical Neurophysiology*, 114(7):1176–1180, 2003.

[160] X. Nan and X. Jinghua. The fractal dimension of EEG as a physical measure of conscious human brain activities. *Bulletin of Mathematical Biology*, 50(5):559–565, 1988.

[161] Z. Liang, Y. Wang, X. Sun, D. Li, L. J. Voss, Jamie W Sleigh, Satoshi Hagihira, and Xiaoli Li. EEG entropy measures in anesthesia. *Frontiers in computational neuroscience*, 9:16, 2015.

[162] M. Sarà and F. Pistoia. Complexity Loss in Physiological Time Series of Patients in a Vegetative State. *Nonlinear Dynamics Psychology and Life Sciences*, 14(1):1–13, 2010.

[163] M. Sarà, F. Pistoia, P. Pasqualetti, F. Sebastiano, P. Onorati, and Paolo M Rossini. Functional isolation within the cerebral cortex in the veget-

ative state: a nonlinear method to predict clinical outcomes. *Neurore-habilitation and neural repair*, 25(1):35–42, 2011.

[164] L. D. Iasemidis, J. Chris Sackellares, H. P. Zaveri, and W. J. Williams. Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures. *Brain Topography*, 2(3):187–201, 1990.

[165] L. D. Iasemidis and J. C. Sackellares. The evolution with time of the spatial distribution of the largest Lyapunov exponent on the human epileptic cortex. *Measuring chaos in the human brain*, pages 49–82, 1991.

[166] L. D. Iasemidis, D. S. Shiau, J. C. Sackellares, P. M. Pardalos, and A. Prasad. Dynamical resetting of the human brain at epileptic seizures: application of nonlinear dynamics and global optimization techniques. *IEEE Transactions on Biomedical Engineering*, 51(3):493–506, 2004.

[167] K. Lehnertz, F. Mormann, T. Kreuz, R. G. Andrzejak, C. Rieke, P. David, and C.E. Elger. Seizure prediction by nonlinear EEG analysis. *IEEE Engineering in Medicine and Biology Magazine*, 22(1):57–63, 2003.

[168] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr. A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy. *IEEE Transactions on Biomedical Engineering*, 54(2):205–11, 2007.

[169] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K. H. Ng, and J. S. Suri. Automated diagnosis of epileptic EEG using entropies. *Biomedical Signal Processing and Control*, 7(4):401–408, 2012.

[170] U. R. Acharya, S. Vinitha Sree, G. Swapna, R. J. Martis, and J. S. Suri. Automated EEG analysis of epilepsy: A review. *Knowledge-Based Systems*, 45:147–165, 2013.

[171] A. Meyer-Lindenberg, U. Bauer, S. Krieger, S. Lis, K. Vehmeyer, Georg Schüler, and Bernd Gallhofer. The Topography of Non-Linear Cortical

Dynamics at Rest, in Mental Calculation and Moving Shape Perception. *Brain Topography*, 10(4):291–299, 1998.

[172] S. Micheloyannis, N. Flitzanis, E. Papanikolaou, M. Bourkas, D. Terzakis, S Arvanitis, and C J Stam. Usefulness of non-linear EEG analysis. *Acta neurologica Scandinavica*, 97:13–19, 1998.

[173] S. Micheloyannis, E. Papanikolaou, E. Bizas, C. J. Stam, and P. G. Simos. Ongoing electroencephalographic signal study of simple arithmetic using linear and non-linear measures. *International Journal of Psychophysiology*, 44:231–238, 2002.

[174] C. Tomberg. Focal enhancement of chaotic strange attractor dimension in the left semantic (Wernicke) human cortex during reading without concomitant change in vigilance level. *Neuroscience Letters*, 263:177–180, 1999.

[175] J. Jeong, M. K. Joung, and S. Y. Kim. Quantification of emotion by nonlinear analysis of the chaotic dynamics of electroencephalograms during perception of 1/f music. *Biological Cybernetics*, 78(3):217–225, 1998.

[176] J. Jeong. Nonlinear dynamics of EEG in Alzheimer's disease. *Drug Development Research*, 56(2):57–66, 2002.

[177] J. Jeong. EEG dynamics in patients with Alzheimer's disease. *Clinical Neurophysiology*, 115(7):1490–505, 2004.

[178] W. S. Pritchard, D. W. Duke, and Kerry L. Coburn. Altered EEG Dynamical Responsivity Associated with Normal Aging and Probable Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 2(2):102–105, 1991.

[179] B. Jelles, J. H. van Birgelen, J. P. J. Slaets, R. E. M. Hekster, E. J. Jonkman, and C.J. Stam. Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls. *Clinical Neurophysiology*, 110(7):1159–1167, 1999.

[180] K. J. Stam, D. L. J. Tavy, B. Jelles, H. A. M. Achtereekte, J. P. J. Slaets, and R. W. M. Keunen. Non-linear dynamical analysis of multichannel EEG: Clinical applications in dementia and Parkinson's disease. *Brain Topography*, 7(2):141–150, 1994.

[181] C. J. Stam, B. Jelles, H. A. M. Achtereekte, S. A. R. B. Rombouts, J. P. J. Slaets, and R. W. M. Keunen. Investigation of EEG non-linearity in dementia and Parkinson's disease. *Electroencephalography and Clinical Neurophysiology*, 95(5):309–317, 1995.

List of Publications

- L. Ricci, L. M. Martini, M. Franchi and A. Bertoldi, "A current-carrying coil design with improved liquid cooling arrangement", *Review of Scientific Instruments*, **84**:065115, 2013.

- M. Franchi and L. Ricci, "Appropriateness of dynamical systems for the comparison of different embedding methods via calculation of the maximum Lyapunov exponent." *Journal of Physics: Conference Series*, **490**(1):012094, 2014.

- M. Franchi and L. Ricci, "Statistical properties of the maximum Lyapunov exponent calculated via the divergence rate method." *Physical Review E*, **90**(6):062920, 2014.