Ph. D. Thesis submitted to the

Dipartimento di Fisica – Università degli Studi di Trento

in fulfilment of the requirements for the

Degree of Philosophiæ Doctor in Physics

Under the Supervision of

Prof. Leonardo Ricci

A new approach

to optimal embedding of time series

Alessio Perinelli

XXXIII cycle

UNIVERSITÀ DEGLI STUDI DI TRENTO

Contents

Abstract

The analysis of signals stemming from a physical system is crucial for the experimental investigation of the underlying dynamics that drives the system itself. The field of *time series analysis* comprises a wide variety of techniques developed with the purpose of characterizing signals and, ultimately, of providing insights on the phenomena that govern the temporal evolution of the generating system. A renowned example in this field is given by spectral analysis: the use of Fourier or Laplace transforms to bring time-domain signals into the more convenient frequency space allows to disclose the key features of *linear* systems. A more complex scenario turns up when *nonlinearity* intervenes within a system's dynamics. Nonlinear coupling between a system's degrees of freedom brings about interesting dynamical regimes, such as self-sustained periodic (though anharmonic) oscillations ("limit cycles"), or quasi-periodic evolutions that exhibit sharp spectral lines while lacking strict periodicity ("limit tori"). Among the consequences of nonlinearity, the onset of *chaos* is definitely the most fascinating one. Chaos is a dynamical regime characterized by unpredictability and lack of periodicity, despite being generated by deterministic laws. Signals generated by chaotic dynamical systems appear as irregular: the corresponding spectra are broad and flat, prediction of future values is challenging, and evolutions within the systems' state spaces converge to *strange* attractor sets with noninteger dimensionality. Because of these properties, chaotic signals can be mistakenly classified as noise if linear techniques such as spectral analysis are used. The identification of chaos and its characterization require the assessment of dynamical invariants that quantify the complex features of a chaotic system's evolution. For example, Lyapunov exponents provide a marker of unpredictability; the estimation of attractor dimensions, on the other hand, highlights the unconventional geometry of a chaotic system's state space.

Nonlinear time series analysis techniques act directly within the state space of the system under investigation. However, experimentally, full access to a system's state space is not always available. Often, only a scalar signal stemming from the dynamical system can be recorded, thus providing, upon sampling, a scalar *sequence*. Nevertheless, by virtue of a fundamental theorem by Takens, it is possible to reconstruct a proxy of the original state space evolution out of a single, scalar sequence. This recon-

struction is carried out by means of the so-called *embedding procedure*: m-dimensional vectors are built by picking successive elements of the scalar sequence delayed by a lag L. On the other hand, besides posing some necessary conditions on the integer embedding parameters m and L, Takens' theorem does not provide any clue on how to choose them correctly. Although many optimal embedding criteria were proposed, a general answer to the problem is still lacking. As a matter of fact, conventional methods for optimal embedding are flawed by several drawbacks, the most relevant being the need for a subjective evaluation of the outcomes of applied algorithms.

Tackling the issue of optimally selecting embedding parameters makes up the core topic of this thesis work. In particular, I will discuss a novel approach that was pursued by our research group and that led to the development of a new method for the identification of suitable embedding parameters. Rather than most conventional approaches, which seek a single optimal value for m and L to embed an input sequence, our approach provides a *set* of embedding choices that are equivalently suitable to reconstruct the dynamics. The suitability of each embedding choice m, L is assessed by relying on statistical testing, thus providing a criterion that does not require a subjective evaluation of outcomes. The starting point of our method are embedding-dependent correlation integrals, i.e. cumulative distributions of embedding vector distances, built out of an input scalar sequence. In the case of Gaussian white noise, an analytical expression for correlation integrals is available, and, by exploiting this expression, a *gauge transformation* of distances is introduced to provide a more convenient representation of correlation integrals. Under this new gauge, it is possible to test—in a computationally undemanding way—whether an input sequence is compatible with Gaussian white noise and, subsequently, whether the sequence is compatible with the hypothesis of an underlying chaotic system. These two statistical tests allow ruling out embedding choices that are unsuitable to reconstruct the dynamics. The estimation of correlation dimension, carried out by means of a newly devised estimator, makes up the third stage of the method: sets of embedding choices that provide uniform estimates of this dynamical invariant are deemed to be suitable to embed the sequence.

The method was successfully applied to synthetic and experimental sequences, providing new insight into the longstanding issue of optimal embedding. For example, the relevance of the *embedding window* $(m-1)L$, i.e. the time span covered by each embedding vector, is naturally highlighted by our approach. In addition, our method provides some information on the adequacy of the sampling period used to record the input sequence.

The method correctly distinguishes a chaotic sequence from surrogate ones generated out of it and having the same power spectrum. The technique of surrogate generation, which I also addressed during my Ph. D. work to develop new dedicated algorithms and to analyze brain signals, allows to estimate significance levels in situations where standard analytical algorithms are unapplicable. The novel embedding approach being able to tell apart an original sequence from surrogate ones shows its capability

to distinguish signals beyond their spectral—or autocorrelation—similarities.

One of the possible applications of the new approach concerns another longstanding issue, namely that of distinguishing noise from chaos. To this purpose, complementary information is provided by analyzing the asymptotic (long-time) behaviour of the so-called *time-dependent divergence exponent*. This embedding-dependent metric is commonly used to estimate—by processing its short-time linearly growing region—the maximum Lyapunov exponent out of a scalar sequence. However, insights on the kind of source generating the sequence can be extracted from the—usually overlooked—asymptotic behaviour of the divergence exponent. Moreover, in the case of chaotic sources, this analysis also provides a precise estimate of the system's correlation dimension. Besides describing the results concerning the discrimination of chaotic systems from noise sources, I will also discuss the possibility of using the related correlation dimension estimates to improve the third stage of the method introduced above for the identification of suitable embedding parameters.

The discovery of chaos as a possible dynamical regime for nonlinear systems led to the search of chaotic behaviour in experimental recordings. In some fields, this search gave plenty of positive results: for example, chaotic dynamics was successfully identified and tamed in electronic circuits and laser-based optical setups. These two families of experimental chaotic systems eventually became versatile tools to study chaos and its possible applications. On the other hand, chaotic behaviour is also looked for in climate science, biology, neuroscience, and even economics. In these fields, nonlinearity is widespread: many smaller units interact nonlinearly, yielding a collective motion that can be described by means of few, nonlinearly coupled effective degrees of freedom. The corresponding recorded signals exhibit, in many cases, an irregular and complex evolution. A possible underlying chaotic evolution—as opposed to a stochastic one—would be of interest both to reveal the presence of determinism and to predict the system's future states. While some claims concerning the existence of chaos in these fields have been made, most results are debated or inconclusive. Nonstationarity, low signal-to-noise ratio, external perturbations and poor reproducibility are just few among the issues that hinder the search of chaos in natural systems. In the final part of this work, I will briefly discuss the problem of chasing chaos in experimental recordings by considering two example sequences, the first one generated by an electronic circuit and the second one corresponding to recordings of brain activity.

The present thesis is organized as follows. The core concepts of time series analysis, including the key features of chaotic dynamics, are presented in Chapter 1. A brief review of the search for chaos in experimental systems is also provided; the difficulties concerning this quest in some research fields are also highlighted. Chapter 2 describes the embedding procedure and the issue of optimally choosing the related parameters. Thereupon, existing methods to carry out the embedding choice are reviewed and their limitations are pointed out. In addition, two embedding-dependent nonlinear techniques that are ordinarily used to characterize chaos, namely the estimation of correlation dimension by means

of correlation integrals and the assessment of maximum Lyapunov exponent, are presented. The new approach for the identification of suitable embedding parameters, which makes up the core topic of the present thesis work, is the subject of Chapter 3 and 4. While Chapter 3 contains the theoretical outline of the approach, as well as its implementation details, Chapter 4 discusses the application of the approach to benchmark synthetic and experimental sequences, thus illustrating its perks and its limitations. The study of the asymptotic behaviour of the time-dependent divergent exponent is presented in Chapter 5. The alternative estimator of correlation dimension, which relies on this asymptotic metric, is discussed as a possible improvement to the approach described in Chapters 3, 4. The search for chaos out of experimental data is discussed in Chapter 6 by means of two examples of real-world recordings. Concluding remarks are finally drawn in Chapter 7.

1. Nonlinear dynamics, noise and chaos in theory and experiments

Natural and artificial dynamical systems exhibit a wide variety of dynamical regimes, from trivial stable fixed points to periodic and aperiodic evolutions. The characterization of dynamical regimes and their identification in experimental measurements are fundamental problems in Physics and make up the core business of *time series analysis*. Because the concept of *dynamical system* is quite general—anything that *moves* can be considered a dynamical system [1]—analytical techniques devoted to the study of dynamical systems were eventually extended to other fields, from biology and physiology to economics. Particular interest is elicited by *chaotic* dynamical regimes, which exhibit—similarly to stochastic processes—unpredictability and aperiodicity, despite being driven by deterministic laws.
This Chapter provides the general framework of this thesis by introducing the essential concepts of time series analysis.

Mathematically, the instantaneous state of a dynamical system is described by a vector $\mathbf{s}$ within a *state space* $\mathcal{S}$ (typically, $\mathcal{S} \subseteq \mathbb{R}^M$). The state vector evolves in time according to an evolution operator $\mathcal{E}_t$:

$$\mathbf{s}(t + t_0) = \mathcal{E}_t \left[\mathbf{s}(t_0)\right] \ .$$

When dealing with theoretical systems, the state space $\mathcal{S}$ is well-defined and the evolution operator $\mathcal{E}_t$ is given: both these objects are either known from first-principles or provided as the results of some modeling. Most often, $\mathcal{E}_t$ is defined by a set of differential equations[1], which can be solved (analytically or by numerical integration) to find the system's evolution $\mathbf{s}(t)$ out of the initial conditions $\mathbf{s}(0)$. An ideal pendulum can be regarded as the prototypical dynamical system: $\mathbf{s}$ is defined by the angle and the velocity of the pendulum, $\mathcal{S} = \mathbb{R}^2$, while $\mathcal{E}_t$ is a second-order differential equation stemming from Newton's law.
The scenario is way more challenging if we consider experimentally observed systems. First, the state

[1] Alternatively, discrete maps (difference equations), partial differential equations and differential equations with delays are other possible "rules" that tell how the state of the system changes in time.

space is not always fully accessible; in some cases, it is even hard to specify what are the degrees of freedom making up $\mathcal{S}$. Second, the evolution operator $\mathcal{E}_t$ is rarely known: most likely, collecting information on $\mathcal{E}_t$ is the very purpose of experimental measurements. Among the countless examples of this scenario, it is worth mentioning possibly one of most complex dynamical system in nature, namely the human brain: both $\mathcal{S}$ and $\mathcal{E}_t$ are definitely unknown.
Experimentally, observing a system consists in recording a *signal* $y(t)$ out of it. The signal is, in general, the result of a "measurement function" $\mathcal{M}$ acting on the system's state:

$$y(t) = \mathcal{M}\left[\mathbf{s}(t)\right] .$$

The observed signal—say, a voltage produced by a transducer—is eventually sampled and digitized, yielding a finite *time series* (or *sequence*) $\{y_n\}$, $n \in [1, \ell]$.
Often, we do not really need to determine $\mathcal{S}$ or $\mathcal{E}_t$, but we are instead interested in classifying and characterizing the system's dynamics in terms of few, relevant quantities, or *dynamical invariants*. This task is the core business of *time series analysis*. Quoting literally Bradley & Kantz [2]:

> *"Time series analysis is essentially data compression: we compute a few characteristic numbers from a large sample of data."*

The "characteristic numbers" of interest might be, for example, the oscillation frequencies of the system. The assessment of dynamical invariants eventually leads to the classification of the system's evolution, which in turn sheds some light on the underlying mechanism that drives the system, i.e. on the nature of $\mathcal{E}_t$. An example of the concepts introduced so far is summarized in Fig. 1.1. In this example, a "synthetic"

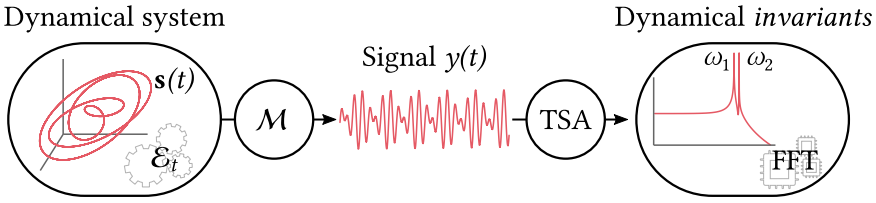


FIGURE 1.1 – *A dynamical system's state* $\mathbf{s}(t)$ *evolves according to* $\mathcal{E}_t$*. The experimentally accessible quantity is a scalar signal* $y(t)$ *that is the outcome of a measurement function* $\mathcal{M}$ *acting on* $\mathbf{s}$*. Time series analysis (TSA) techniques allow to assess dynamical invariants that characterize the system. In this example, by evaluating the FFT of the signal, two sharp spectral lines at* ω_1*,* ω_2 *are identified, thus characterizing the system's periodic nature.*

dynamical system is observed through a measurement function $\mathcal{M}$; the produced scalar signal is of the kind $\sin(\omega_1 t) \cdot \sin(\omega_2 t)$. The two characteristic frequencies ω_1, ω_2 are revealed by applying one of the most known time series analysis technique, namely Fast Fourier Transform (FFT), to the scalar signal.

Despite $\mathcal{E}_t$ remains unknown, the spectrum estimated by means of the FFT suggests us that the system's evolution is periodic and presumably linear.

A key question concerning the classification of a dynamical system is whether its evolution is *deterministic*, implying that future states are univocally determined by the present state, or *stochastic*, i.e. random. Signals belonging to the latter class, which are referred to as *noise*, are ubiquitous in nature [3]. Noise is usually characterized in terms of its power spectrum, or equivalently, by virtue of Wiener-Khinchin theorem, in terms of its autocorrelation function [4, 5]. Typically, stochastic signals exhibit a broad (possibly flat) spectrum, corresponding to short autocorrelation times. The prototypical example of noise is *Gaussian white noise* (GWN), which consists of a sequence of uncorrelated random numbers each distributed according to the same Gaussian distribution: $y_n \sim \mathcal{N}(\mu, \sigma^2)$. Due to the lack of temporal correlation, the autocorrelation function of GWN is a described by a Dirac delta function, and its power spectral density is frequency-independent. The presence of noise in physical systems is unavoidable due to two fundamental sources, namely the fluctuation-dissipation theorem and the uncertainty principle of quantum mechanics [3]. Consequently, although generated by a deterministic dynamic, experimentally recorded signals are *always* contaminated by noise, both affecting the system's evolution (*dynamical noise*) as well as added during measurement of the signal (*observational noise*). A crucial task in time series analysis consists in estimating the amount of noise contained within a sequence and, possibly, to separate the noisy contributions from the deterministic signal.

As far as determinism is concerned, we can distinguish between *linear* systems, for which the superposition principle holds, i.e.

$$\mathcal{E}_t\left[c_1\mathbf{s}_1 + c_2\mathbf{s}_2\right] = c_1\mathcal{E}_t\left[\mathbf{s}_1\right] + c_2\mathcal{E}_t\left[\mathbf{s}_2\right] ,$$

and *nonlinear* systems for which the last equality is not satisfied. Nonlinearity enriches the possible dynamical regimes of a system and, for certain ranges of the system's parameters, can lead to *chaos* [6].

Chaos in theory

Chaos is a dynamical regime characterized by an aperiodic and unpredictable behaviour despite the underlying equations being deterministic [1]. For example, a simple three-dimensional system of differential equations such as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y , \tag{1.1a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = z , \tag{1.1b}$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = y^2 - x - Az , \tag{1.1c}$$

leads to a chaotic evolution if A is set within $2.0168 < A < 2.0577$ [7]. The system[2] is characterized by only one nonlinearity, given by the term y^2 in the last equation, which "pushes up" z at each point along the trajectory except for $y = 0$. Eventually, joined with the first two equations, this term would make the system diverge. On the other hand, the $-Az$ term provides a damping-like action on y, which is tuned through the parameter A, and together with the term $-x$ manages to keep the system within a bounded region of the phase space.

Qualitatively, the "irregularity" of chaotic evolution can be appreciated by looking at plots of the state space trajectory or of the corresponding time series. Figure 1.2 shows an example of evolution

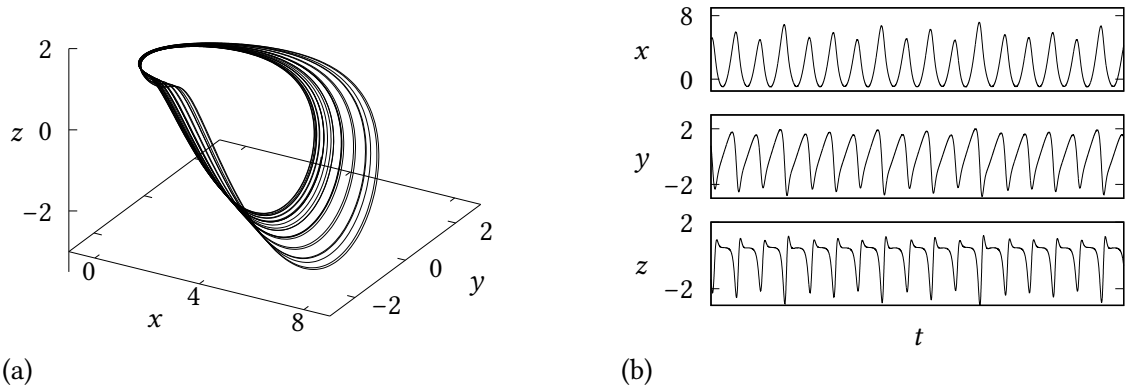


FIGURE 1.2 – *State space evolution (a) and corresponding time series (b) of a solution—obtained through numerical integration—to Eqs. (1.1). The parameter A is set to* 2.04*; initial conditions are randomly assigned (the initial transient is not shown).*

governed by Eqs. (1.1): the system's state is identified by $\mathbf{s} = (x, y, z)$. The evolution shown in the figure, which is obtained by setting $A = 2.04$ and by numerically integrating the system of differential equations, is purely deterministic. Nevertheless, the state space trajectory exhibits a complex structure, and the corresponding time series show seemingly random fluctuations in amplitude.
These characteristics are not just qualitative features of chaotic evolution: they correspond to mathematical properties of the set of points within the state space that make up the evolution trajectory, i.e. the system's *attractor* [5]. Following the definition by Devaney [8], three conditions are necessary and sufficient to define a system as *chaotic*: (i) sensitivity to initial conditions; (ii) topological transitivity; (iii) density of periodic orbits.

(i) Sensitivity to initial conditions is what makes chaos hard to predict, despite its evolution being deterministic. Given two initial conditions $\mathbf{s}_A(0)$, $\mathbf{s}_B(0)$ that are arbitrarily close within the state space ($\|\mathbf{s}_A(0) - \mathbf{s}_B(0)\| < \varepsilon$), the two trajectories evolving out of these initial conditions diverge

[2] The system of Eqs. (1.1) can also be recast into a single third-order differential equation (physically, a *jerk* equation): $\dddot{x} + A\ddot{x} - \dot{x}^2 + x = 0$.

exponentially in time, i.e.

$$\|\mathbf{s}_A(t) - \mathbf{s}_B(t)\| \propto e^{\lambda t},$$

where λ is called *maximum Lyapunov exponent* (MLE). In other words, two arbitrarily close—but different—initial states eventually evolve into completely different trajectories. This property is also known as "butterfly effect" because of the celebrated talk by Lorenz [9] entitled *"Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?"*, where the author stated

> *"The question which really interests us is [...] whether, for example, two particular weather situations differing by as little as the immediate influence of a single butterfly will generally after sufficient time evolve into two situations differing by as much as the presence of a tornado."*

The practical consequence of this property is that, regardless of how precisely we can initially measure a system's state, our prediction of the system's future will eventually fail even if we know the *exact* dynamical laws that rule the evolution. A diagram depicting sensitivity to initial conditions is shown in Fig. 1.3, where two trajectories obtained by numerically integrating Eq. (1.1) ($A = 2.04$) and projected on the x-z plane are displayed. The two evolutions begin from two initial states displaced—along the x axis—by a relative amount equal to 10^{-3}. In time, the trajectories diverge and their distance increases, as shown in the magnification.

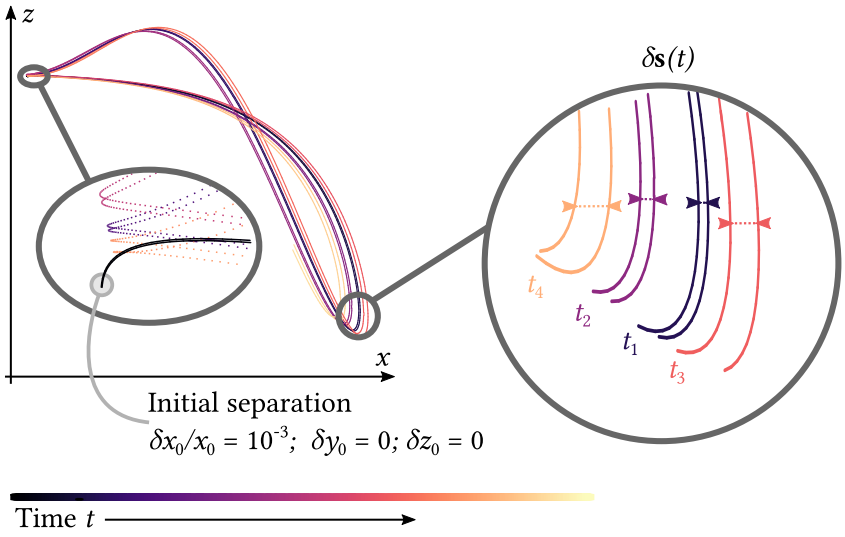


FIGURE 1.3 – *Chaotic systems exhibit sensitivity to initial conditions. Two evolutions of the system described by Eq. (1.1) (projected on the x-z plane) are initialized with two slightly different starting points. Line color corresponds to the time coordinate: the two trajectories are initially indistinguishable (dark colors), but progressively diverge as time progresses.*

(ii) Topological transitivity is the property according to which a chaotic trajectory eventually connects any region of the state space with any other. In other words, the state space of a chaotic system cannot be decomposed into disjoint subsets.

(iii) Within the state space, we can find a periodic orbit arbitrarily close to any given point: consequently, periodic orbits make up a dense set.

While Devaney's definition provides a formal characterization of chaos, the assessment of (ii) and (iii) is unfeasible when dealing with time series. In practice, sensitivity to initial conditions is regarded as the most prominent hallmark of chaos [1]. Moreover, the mathematical properties outlined above have observable consequences on the features of chaotic time series. First of all, the evolution of a chaotic system within the state space converges to a *strange attractor*, i.e. a set of state space points that has a noninteger dimensionality and exhibits self-similarity (i.e. it is a *fractal*) [1]. A sketch showing this property is displayed in Fig. 1.4 for a trajectory corresponding to a solution (obtained by numerical integration) of Eq. (1.1) projected on the x-y plane. The region close to the origin is progressively magnified: the same qualitative features are visible at all magnification levels, as a consequence of the self-similarity of the attractor. Although strange attractors exist also in nonchaotic systems [10], the

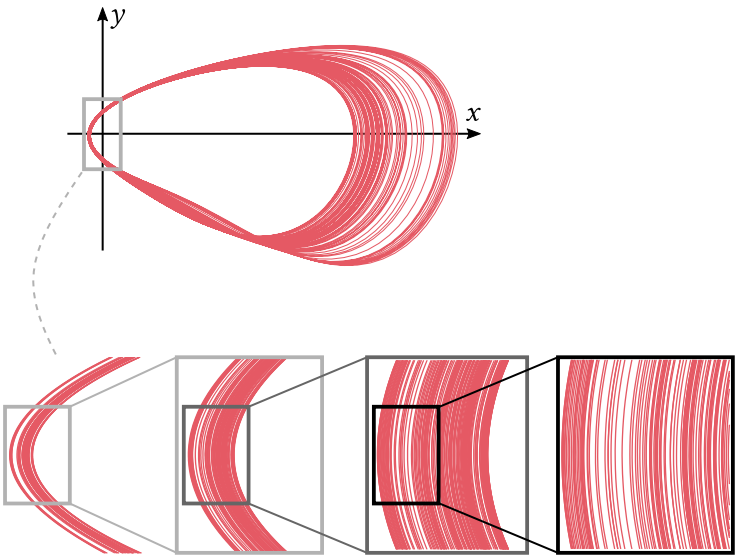


FIGURE 1.4 – *Chaotic attractors are called* strange *because they exhibit a self-similar structure. A trajectory generated by integrating Eq. (1.1) and projected on the x-y plane is progressively magnified to reveal self-similarity.*

assessment of an attractor's fractal structure by estimating the corresponding dimension is one of the most widespread techniques to identify chaos.

Besides self-similarity, another consequence of the three mathematical properties listed above regards the spectral features of chaotic systems: sequences generated by this class of systems are aperiodic and characterized by broad, noise-like Fourier spectra [11]. An example is shown in Fig. 1.5, where the Fourier transform of a signal generated by a chaotic system, namely the x-coordinate of a Lorenz sys-

tem (see Eq. (1.2) below), is compared to the same function estimated out of filtered GWN[3]. Despite the purely deterministic nature of the former signal, the corresponding Fourier amplitudes and phases are essentially indistinguishable from the same quantities computed out of the stochastic sequence. This

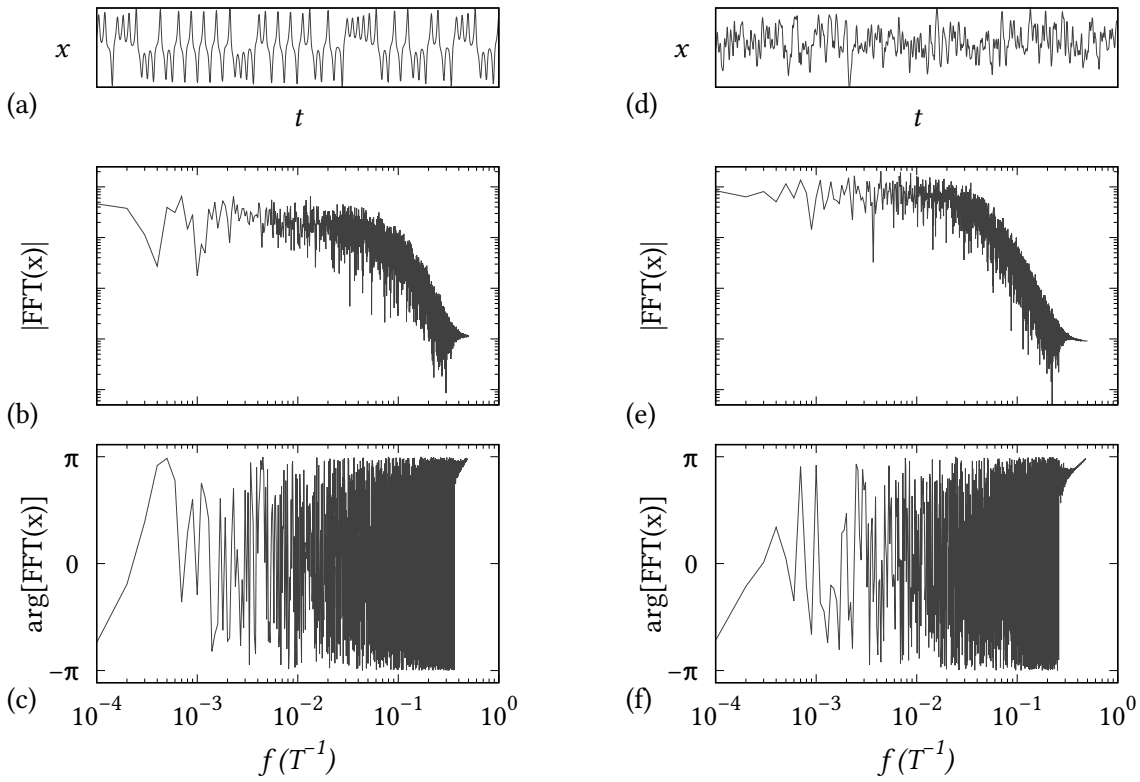


FIGURE 1.5 – *Chaotic sequences exhibit spectra that can be easily mistaken for those of stochastic processes. (a) Sequence corresponding to the x coordinate of a Lorenz system (see Eq. (1.2) below). By evaluating the FFT of the sequence, the Fourier amplitude (b) and the corresponding phase (c) are estimated. (d) Sequence corresponding to GWN processed by a fourth-order low-pass filter having cut-off frequency equal to 1/20 of the sampling frequency. The FFT amplitude and phase are shown in (e) and (f), respectively. In (b,c,e,f) frequencies are measured in units of inverse sampling period (T^{-1}).*

example provides an evidence of how linear techniques (such as the Fourier transform) can be fooled by chaotic sources. The complex nature of chaotic evolution cannot be fully grasped just by analyzing sequences in the frequency space, as we could do with linear systems: techniques that act directly onto the system's state space are called for.

One might wonder, at this point, how can we develop "state space techniques" if the state space itself is inaccessible, as discussed above, and the only available quantity is a scalar sequence. This apparent

[3]The sequence is generated by first drawing random numbers distributed according to a standard normal distribution and, thereupon, by piping the sequence of random numbers into a fourth-order digital low-pass filter with cut-off frequency equal to 1/20 of the sampling frequency.

contradiction is solved thanks to the so-called *"embedding procedure"*, which is thoroughly discussed in Chapter 2 and whose correct implementation makes up the central topic of this thesis work.
Provided that the state space evolution of a system is accessible through experimental recordings, the identification and characterization of chaos mainly relies on assessing sensitivity to initial conditions and on estimating the geometric properties of the system's attractor. The former feature is assessed by computing the MLE λ, as described in Sec. 2.5. Similarly, the estimation of the dimensionality of a system's attractor is a crucial step to identify chaotic dynamical regimes: a discussion on dimensions and their calculation is the topic of Sec. 2.4.

Chaotic behaviour was assessed in many theoretical continuous systems described by a set of ordinary differential equations (*flows*), as well as in discrete systems described by difference equations (*maps*). A collection of chaotic systems is provided in Ref. [12].
In the following, for the sake of simplicity, we only consider flows. Because chaotic evolution is deterministic and aperiodic, no intersections are allowed within the trajectory and the evolution cannot return twice on the same point within the state space. Consequently, the minimum number of independent dimensions that are required to sustain a chaotic flow is 3, corresponding to a set of three first-order differential equations such as Eqs. (1.1), or, equivalently, to a third-order differential equation.

In this work, three theoretical dynamical systems are used as references to provide synthetic sequences generated by a chaotic dynamic. The first system considered here is the well-known Lorenz system [13], which is a model of atmospheric convection representing forced dissipative hydrodynamic flows. The set of equations describing the system is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma(y - x)\,, \tag{1.2a}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = x(r - z) - y\,, \tag{1.2b}$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = xy - bz\,. \tag{1.2c}$$

Standard settings for the parameters are $\sigma = 10$, $r = 28$, $b = 8/3$ [14]. An example of state space evolution, as well as of the corresponding time series, is shown in Fig. 1.6.
The Ueda oscillator [15], which stems from Duffing's equation, is described by two differential equations with a periodic forcing:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y\,, \tag{1.3a}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = -x^3 - ky + A\sin(t)\,, \tag{1.3b}$$

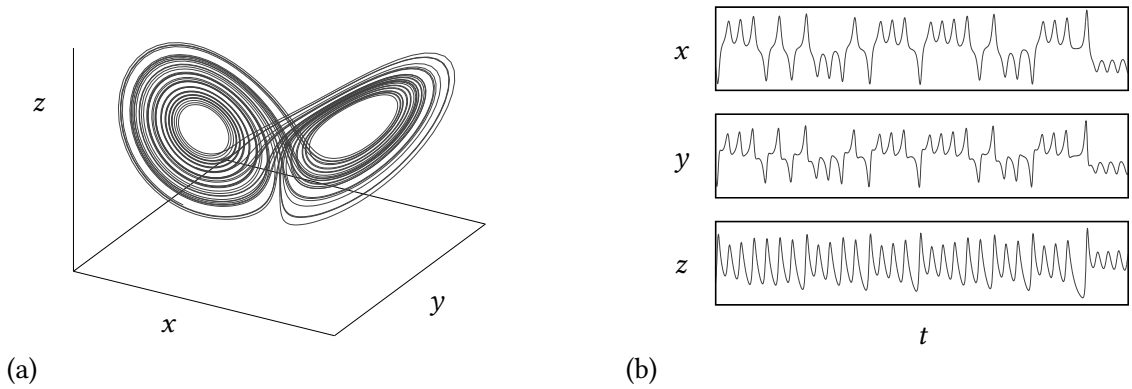


FIGURE 1.6 – *State space evolution (a) and corresponding time series (b) for the benchmark Lorenz system used in this work.*

where the parameters are typically set to $A = 7.5$, $k = 0.05$ [14]. Figure 1.7 shows an example of state space evolution for the Ueda oscillator, as well as of the corresponding time series; the z coordinate is provided by the periodic forcing $\sin(t)$ (the Ueda oscillator is non-autonomous).

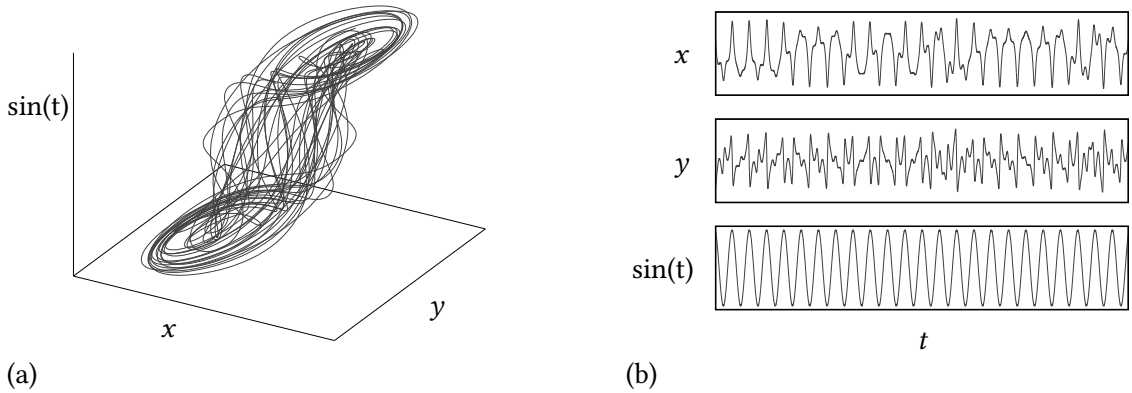


FIGURE 1.7 – *State space evolution (a) and corresponding time series (b) for the benchmark Ueda oscillator used in this work.*

The third dynamical system considered in this work is the Rössler system [16], introduced as a "simpler" version of the Lorenz system (a single nonlinearity is present in the equations) and defined by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -y - z\,, \tag{1.4a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = x + ay\,, \tag{1.4b}$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = b + z(x - c)\,. \tag{1.4c}$$

In the following, two different settings for the system's parameters are considered: the settings $a = b = 0.2$,

$c = 5.7$ are referred to as "standard", while setting $a = 0.343$, $b = 1.82$, $c = 9.75$ yields the so-called "funnel"[4] attractor [17]. Figure 1.8 shows examples of state space evolution and of the corresponding time series for these two settings of the Rössler system.

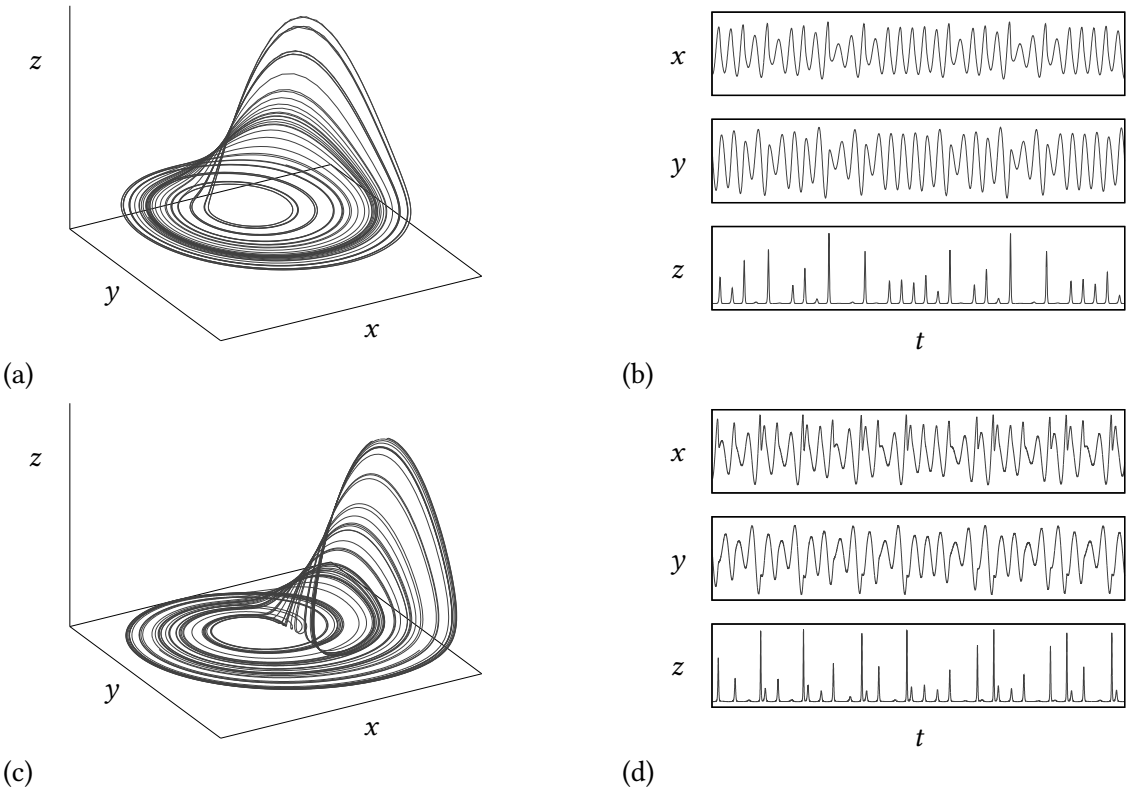


FIGURE 1.8 – *State space evolution (a,c) and corresponding time series (b,d) for the benchmark Rössler systems used in this work. (a,b) Rössler system with "standard" parameter settings; (c,d) Rössler system with "funnel" parameter settings.*

Throughout this work, benchmark synthetic sequences are generated out of these theoretical systems by numerical integration. Initial conditions are randomly assigned, and integration is carried out by means of a Runge-Kutta Prince-Dormand (8,9) algorithm [18]. System parameters, as well as the integration time step δt, are specified every time a synthetic sequence is used. The sampling period T is always taken equal to δt.

[4] This name is due to the peculiar shape of the system's trajectories, which are folded over to give a shape similar to that of a *cornucopia*.

Chaos in experiments

The discovery of theoretical systems admitting chaotic evolutions was eventually followed by the search of chaos in natural systems. The first observations of chaos in experiments were carried out in chemical, mechanical, optical and electronic systems.

As early as 1981, chaotic behaviour was observed in the Belousov-Zhabotinskyi reaction, a nonlinear chemical oscillator far from thermodynamic equilibrium [19, 20]. Since this first observations, chemical chaos was thoroughly investigated [21, 22, 23]: for example, reactions governed by the same equations as the Lorenz system were identified [24]. Chemical chaos is also responsible for complex spatial patterns, with possible applications to surface-catalyzed reactions [22].

The presence of chaos was also investigated in industrial systems to improve control and diagnostics. Attractors corresponding to chaotic dynamics were reconstructed from recordings of electric arc furnace discharges with the aim of improving real-time control of these systems [25]. Mechanical vibrations in metal cutting [26, 27] and drilling [28] were analyzed by means of nonlinear time series analysis techniques: a possible application concerns the real-time monitoring of industrial machinery [29]. Atomic force microscopy microcantilevers were also shown to exhibit—under certain conditions—chaotic oscillations, which introduce a source of "deterministic error" in measurements [30].

Optics is one of the research fields in which chaos has been mostly studied. One of the first works about chaos in optical systems was published in 1979 by Ikeda [31], who discussed a model of "instability" of transmitted light by a ring cavity—later renamed as "optical turbulence" [32]—that eventually leads to chaotic behaviour. These theoretical results were confirmed by several experiment using lasers with time-delayed feedback [33, 34], as well as in pulsed [35], frequency-modulated [36] and gain-modulated [37] lasers. Moreover, some laser systems were shown to be governed by the same equations as those of a Lorenz system [38], an equivalence that was experimentally proved [39, 40].

The sequence corresponding to the intensity of a far-infrared (81.5 µm wavelength) NH_3 ring laser studied in Ref. [40] and publicly available [41] is shown in Fig. 1.9. Besides showing an excerpt of the intensity sequence $\{y_n\}$, Fig. 1.9 also shows an attempted reconstruction of the system's attractor performed by means of the so-called "embedding procedure", which is thoroughly described in Chapter 2: points in the three-dimensional space of Fig. 1.9(b) correspond to vectors (y_n, y_{n+1}, y_{n+2}). The far-infrared laser intensity sequence is used as a reference experimental recording throughout this thesis work.

The interest in chaotic lasers grew in the last decades and particular attention was devoted to the possible application of chaotic lasers as (pseudo-)random generators [42, 43, 44], e.g. for data encryption. A comprehensive review on laser chaos—including the more recent cryptographic applications—can be found in Ref. [45].

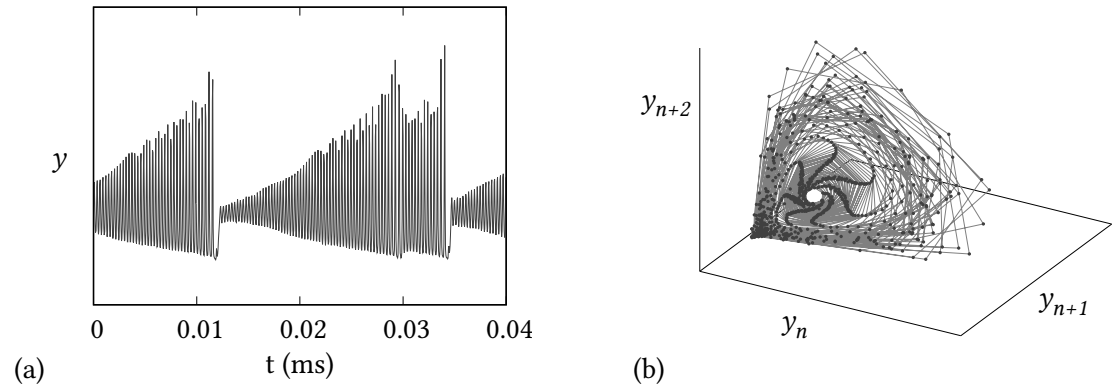


FIGURE 1.9 – *(a) Excerpt of an intensity sequence stemming from a far-infrared laser running in chaotic regime [40, 41]. (b) Tentative reconstruction—via the* embedding procedure *described in Chapter 2—of the corresponding attractor.*

Electronic circuits provide excellent platforms for the study of chaos. As for optical systems, the first observations of chaos in electronics date back to the early Eighties [46]. The prototypical chaotic circuit is Chua's circuit [47], whose study and implementation prompted a huge number of publications [48]. The high degree of customization of electronic circuits makes them an efficient tool to generate and optimize different chaotic regimes [49, 50, 51, 52], as well as to simulate theoretical models [53]. Moreover, the inexpensiveness and scalability of electronic setups allows to investigate chaotic behaviour emerging from collective phenomena in networks of oscillators [54, 55] or even in fractal-like topologies [56].

The list of experimental findings concerning chaos in real systems presented so far is limited to *controlled* systems. Electronic circuits, lasers, mechanical devices and chemical reactions are dynamical systems that can be designed, optimized and stabilized with a high degree of control. Therefore, these systems make up excellent test benches, as well as versatile chaos generators. The search for chaos in natural systems that elude the experimentalist's control is definitely more challenging. Consequently, the corresponding results in terms of chaos indentification are often questionable.

As a first example, the irregular fluctuations in the number of sunspots were claimed to be generated by a chaotic dynamics [57, 58] and to exhibit a self-similar structure [59]. However, sunspot number was also shown to be well described by a relaxation oscillator model with stochastic perturbations [60], thus contradicting the claims of chaoticity. Moreover, an analysis of "light curves" of stars (i.e. sequences of recorded brightness) showed that the corresponding dynamics is *strange*, but *nonchaotic* [61].

Similar considerations apply in geophysics. The appealing idea of an underlying deterministic source—and thus of predictability, albeit limited—prompted the use of nonlinear time series analysis techniques in this field [62]: chaos was claimed to be present in the oscillations of sea levels [63, 64], in seismic

activity [65, 66, 67, 68] and in volcanic eruptions [69]. These claims were criticized by pointing out that techniques for the identification of chaos were misused [70]; for example, some works showed that seismic phenomena can be described by stochastic processes [71, 72]. Nonlinear analytical techniques are also used to study atmospheric data [73, 74], mostly aimimg at forecasting. Because the Lorenz system is a model of atmospheric convection [13], researchers expect the weather to show—at least under some conditions and for some spatial scales—a chaotic dynamical regime. Chaos detection in weather data was positively assessed, for example, by processing pressure time series [75] and wind recordings [76, 77]. Again, some doubts were raised concerning the correct application of nonlinear analytical techniques to climate data, as well as on the interpretation of the corresponding results [78, 79, 80]. In this context, a "climatic attractor" was claimed to exist and to have dimension 3.1 [81], although the reliability of this assessment was criticized due to it being based on as few as 500 data points [82]. Similarly, a finite dimension of about 7 was estimated out of a time series of about 4000 values of vertical wind velocity [83]. However, as pointed out by Procaccia [84], the amount of data to reliably estimate the dimension of such a system would be of order 10^7: high dimensions assessed out of short sequences should be considered unreliable. In many cases, for example for sequences of daily temperature, stochastic modeling is sufficient to extract system features [85].

The difficulties in chaos detection outlined so far are common also in the field of physiology. Mackey and Glass first introduced [86] a model for physiological control—based on a differential equation with delay—that can produce chaotic evolutions, thus prompting the search for chaos in physiological recordings [87].

The discovery of chaotic regimes in the dynamics of heart pacemaker cells [88, 89] opened the way to nonlinear analysis into cardiology [90]. Some works claimed that electrocardiographic recordings indeed provide evidence of a chaotic heart [91, 92]. However, other studies showed that heartbeat modulations can be reproduced by means of a stochastic model [93]: a general consensus was established on the fact that heartbeats are not chaotic, but are rather generated by a stochastic process that exhibits "multifractality" [94, 95, 96]. A review on nonlinear analysis of the human heart is provided in Ref. [97].

The human brain produces electromagnetic signals that are routinely recorded through electroencephalography (EEG) [98] or magnetoencephalography (MEG) [99]. Brain recordings show complex, irregular evolutions [100], thus stimulating the search for possible chaotic regimes in neurobiology [101]. Chaos was first assessed in sleep recordings of human brain EEG by Babloyantz *et al.* [102, 103]. Particular interest was raised by epileptic seizures, especially with regards to the possibility of their prediction via nonlinear techniques [104, 105]. Both Babloyantz & Destexhe [106] and Lehnertz & Elger [107] claimed that the human brain attractor drifts from higher to lower dimensionality at the onset of an epileptic seizure. Other examples concern the existence of chaos in the olfactory system [108], while listening to music [109], or simply in resting-state [110, 111]. The possible role of chaos in neural

systems, related models and experimental evidence are reviewed in Refs. [112, 113]. The search for chaos in the human brain was also encouraged by the experimental evidence of chaotic regimes in the activity of single neurons [114, 115] and in numerical simulations of neuron populations [116]. However, the experimental evidence mentioned so far were critically re-examined and sometimes rejected: as pointed out by some studies, the limitations of the algorithms applied to detect chaos were often overlooked [117, 118, 119]. Consequently, the efforts in nonlinear analysis of human brain signals were redirected [119] towards the development of new nonlinear metrics for the quantitative characterization of complexity, rather than to specifically assess deterministic chaos [111, 117, 120].

The above list of experimental investigations looking for chaos in natural systems is definitely not exhaustive. Nevertheless, it allows to draw some conclusions concerning the possibility of detecting chaotic regimes in the real world, as well as, more in general, on nonlinear analysis of experimental recordings. First of all, chaos is definitely achievable in laboratory systems (above all, optics and electronics) for which we can *tune* parameters with sufficient precision. Noise, in this case, does not spoil the chaotic evolution: as shown in Chapter 6 for an electronic circuit, chaos is attainable and detectable by using cheap electronic components and basic equipment. Similar considerations also apply to sufficiently small biological systems, such as single neurons. On the contrary, a definitive assessment of chaos (or the lack of it) in systems comprising a huge number of interacting "units", such as earth climate or the human brain, is far from trivial. The difficulties concerning the analysis of these systems, which are exemplified in the literature discussion above, can be ascribed to several causes. First of all, chaotic behaviour is expected in this class of systems as a consequence of the huge number of interacting units—being them air molecules or neurons—"organizing" their motion into few effective degrees of freedom due to nonlinear coupling [1]. However, this organization does not necessarily exist at all spatial scales within the system, and the physical quantities that we can access might not probe the system with sufficient depth. For instance, the dynamics of small groups of neurons might be chaotic, but EEG and MEG recordings can only measure the fields resulting from a superposition of the activity of many thousands of neurons: therefore, unless some synchronization occurs, these recordings might be essentially stochastic. Another crucial problem to deal with is the nonstationary of this kind of systems: both a human brain and the weather definitely drift between different dynamical regimes, possibly confounding analysis algorithms. Other technical issues exist: for example, the signal-to-noise ratio of electromagnetic fields recorded out of a human brain is low and the interesting dynamics might be washed out by noise [121, 99, 100]. Finally, insufficiently long or too coarsely sampled sequences can lead to misleading results [1, 117]: as discussed by Kantz & Olbrich [122], the length of scalar sequences required to detect determinism increases exponentially with the system's dimension. While "simple" systems (e.g. a chaotic circuit) have low dimensionality ($2 \div 3$), a human brain's attractor is possibly characterized by a higher dimension, and the available recordings migh just be too short to provide

conclusive evidence on its deterministic nature.

The issue of detecting chaos in experimental recordings is discussed in Chapter 6 by applying the method described in Chapter 3 to an electronic circuit and to MEG recordings of a human brain in resting-state.

2. Chaos characterization and the longstanding issue of optimal embedding

The characterization of a nonlinear system essentially consists in condensing its dynamics into a handful of *invariant quantities*. This operation relies on the possibility to analyze the state space evolution of the system itself. If the differential (or difference) equations that rule the dynamics are known, the corresponding state space evolution is readily accessible. On the other hand, observing the whole state space is almost never possible in experimental conditions. In most cases, only few (or even just a single one) observable quantities can be recorded out of a dynamical system. Nevertheless, it is possible in principle to reconstruct the full dynamics out of a single measured signal, i.e. a scalar sequence. This quite remarkable possibility, which stems from Takens' theorem, is implemented by means of the so-called *time delay embedding*—or simply *embedding*—procedure. The correct application of the embedding procedure to experimental sequences, however, is not straightforward at all, due to the presence of noise, the finite sampling rate and the finite length of recordings. Despite the considerable efforts to find a criterion for *optimal embedding*, a definitive method is still lacking.
In the first part of this chapter the embedding procedure is described and, thereupon, the issues concerning its correct implementation are discussed. The second part of the chapter deals with two techniques that rely on embedding and allow to estimate fundamental quantities that characterize nonlinear systems, namely correlation dimension and maximum Lyapunov exponent. These embedding-dependent techniques provide the building blocks of the methods discussed in Chapters 3–5.

2.1 State space reconstruction: the *embedding* procedure

As discussed in Chapter 1, the instantaneous state of a dynamical system is described, in general, by an M-dimensional vector $\mathbf{s} = (s_1, s_2, \ldots, s_M)$ within an M-dimensional *state space*. The system's evolution,

i.e. the trajectory $\mathbf{s}(t)$ within this space, is provided by a system of differential equations:

$$\frac{\mathrm{d}\mathbf{s}(t)}{\mathrm{d}t} = \mathbf{F}\left[\mathbf{s}(t)\right] .$$

In most experimental conditions, observing the system's state space evolution $\mathbf{s}(t)$ is impracticable, and experimentally available data are limited to a single, scalar quantity $y(t)$ that is a function of $\mathbf{s}(t)$. This limitation emerges immediately if we think about complex dynamical systems such as the Sun, financial markets, or the human brain. The observable $y(t)$ is usually sampled with a sampling period T, so that data come in the form of a scalar sequence $\{y_n\}$: $y_n = y(nT)$.

Despite the apparent scantiness of the information contained within $\{y_n\}$ with respect to $\mathbf{s}(t)$, many properties of the system can be estimated by reconstructing a *proxy* state space evolution out of the available scalar sequence. This reconstruction is implemented by the so-called *time delay embedding* procedure, which is henceforth simply referred to as *"embedding"*. Embedding consists in building a sequence of m-dimensional vectors $\mathbf{Y}_n$ by picking m time-delayed samples of the sequence y_n:

$$\mathbf{Y}_n = \left(y_n,\ y_{n+L},\ y_{n+2L},\ \ldots,\ y_{n+(m-1)L}\right) . \tag{2.1}$$

The parameter L, which is an integer number, is the *lag* (in terms of sequence index) at which samples of the scalar sequence are picked. A "causal" version of Eq. (2.1), in which samples y_{n-kL} are picked in place of y_{n+kL} ($k \in [0, m-1]$), is by all means equivalent. A diagram of the embedding procedure is shown in Fig. 2.1.

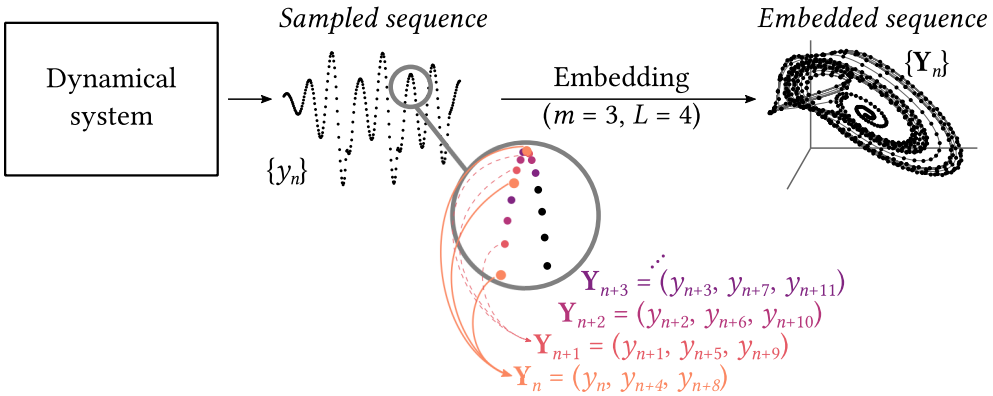


FIGURE 2.1 – *The state space evolution of a dynamical system is reconstructed out of a scalar sequence y_n by means of the embedding procedure ($m = 3$, $L = 4$).*

The time delay embedding procedure was first devised by Packard *et al.* [123]. Soon after, Takens [124] and Mañé [125] provided the proofs that the procedure—if done correctly—indeed recon-

structs a state space evolution that is topologically identical to the full original dynamics [2]. In other words, according to Takens' theorem, a "good" embedding provides a smooth one-to-one map from the original state space evolution to the reconstructed one [1], thus allowing to estimate many properties of dynamical systems that are invariant under such mapping. Takens' theorem poses two conditions on the embedding parameters m (dimension) and L (lag) for an embedding to be successful. The first condition prescribes that $m > 2D$, where D is the dimension of the manifold corresponding to the system's evolution. This condition can be relaxed to $m > D$ when estimating correlation dimension [126]. The second one requires L not to be a multiple of the period of any system's orbit [127, 2].

Before carrying on with the discussion on embedding, it is worth mentioning that, although time delay embedding is by far the most common reconstruction method [11, 2], other embedding models exist. For example, a method proposed by Broomhead & King [128] (and a similar one by Landa & Rosenblum [129]) is based on principal component analysis. Another reconstruction method consists in building vectors by picking the derivatives of $y(t)$ (estimated from y_n) of successively higher order [130]. In general, any vector that is reconstructed from scalar data is related to time delay embedding by a linear transformation [122]. In the case of reconstruction relying on derivatives, this relationship is revealed by the first-order estimator of derivatives, namely (T is the sampling period):

$$\dot{y}(t) \approx \frac{y_{n+1} - y_n}{T},$$

which is indeed a linear combination of delayed coordinates.

2.2 The issue of optimal embedding

The embedding procedure is quite a powerful tool. In principle, it allows us to study an arbitrarily complex dynamical system by simply measuring any observable produced by it. Takens' theorem, as well as later reformulations such as the one by Sauer *et al.* [131], are existence proofs and do not give any clue on how to choose m and L besides the minimal conditions stated above. We might then conclude that, in agreement with the theorem, *any* valid choice of m and L is good to reconstruct the dynamics. However, this is true only for noiseless, finely sampled and infinitely long sequences [1, 2]. Real datasets put forward several issues regarding the choice of embedding dimension and lag: the presence of noise, which blurs states, and the finite sampling period make the selection of m and L critical [130].

As far as m is concerned, the condition $m > 2D$ implies knowledge of D, which is usually unknown and whose estimation depends (circularly) on the sequence being embedded. Because any $m > 2D$ satisfies the theorem's condition, one might be tempted to choose a "very large" $m \gg 1$. However, assuming the deterministic evolution of the system to be correctly embedded with a dimension $m_0 < m$,

the remaining $m - m_0$ dimensions would then be "populated" by noise. Of the m-dimensional reconstructed space, only a subspace of dimension m_0 would be informative: the remaining dimensions would be redundant, thus increasing the computational cost of analysis algorithms, or even leading them astray [11, 5]. A sketch of this issue is provided in Fig. 2.2. As an example, a sinusoidal signal

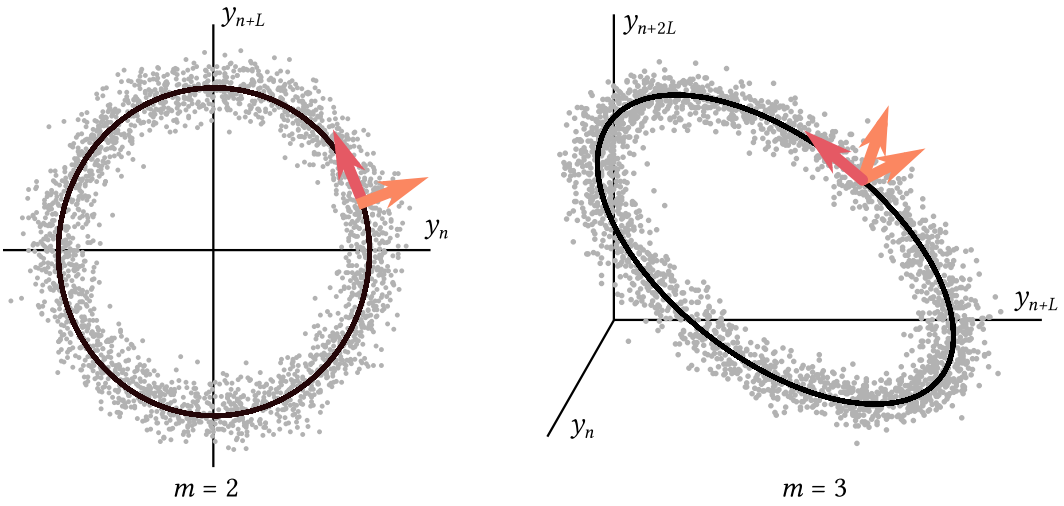


FIGURE 2.2 – *Embedding of a sinusoidal sequence contaminated by observational GWN (L is chosen as 1/4 of the period of the sinusoid). Dots correspond to embedding vectors; the black, solid line shows the reconstructed trajectory in the noiseless case. The direction along the noiseless trajectory is highlighted by means of a red arrow, while orange arrows identify the $m - 1$ directions orthogonal to it.*

contaminated by observational GWN[1] is considered. In this simplified example, we already know that $m = 2$ is enough to reconstruct the state space evolution of the system, which is one-dimensional. The sequence is embedded with $m = 2$ and $m = 3$ (the lag is chosen as 1/4 of the period of the sinusoid). Both choices of m yield a trajectory that faithfully reproduces the deterministic evolution. However, the directions orthogonal to the deterministic noiseless trajectory are "populated" by noise and do not provide extra information: the larger m, the more are these redundant, "noisy dimensions".
An additional problem, which might be relevant when dealing with short sequences, is that the larger is the embedding dimension m, the fewer independent embedding vectors are available.
For these reasons, the smallest m that leads to the topologically correct reconstruction is desired [2].

Regarding the lag L, its selection is not trivial, too. On the one hand, if L is too short, the elements of the embedding vectors $\mathbf{Y}_n$ are strongly correlated with each other: the embedded evolution described by $\mathbf{Y}_n$ is clustered around the diagonal of the m-dimensional space [5]. If the noise level is higher than the "thickness" of this evolution, the reconstructed trajectory is essentially indistinguishable from the diagonal itself. On the other hand, a too large L leads to the opposite situation: the elements of $\mathbf{Y}_n$ are

[1] Here, the sinusoidal signal has unit amplitude, while the noise standard deviation is 0.1.

completely uncorrelated with each other, and the trajectory is folded over on itself [1, 2]. A faithful reconstruction of state space trajectories thus requires a trade-off between these two extreme cases. An example of the outcomes of selecting too small or too large L for state space reconstruction is shown in Fig. 2.3. In this example, a sequence corresponding to the x-coordinate of a Rössler system[2]

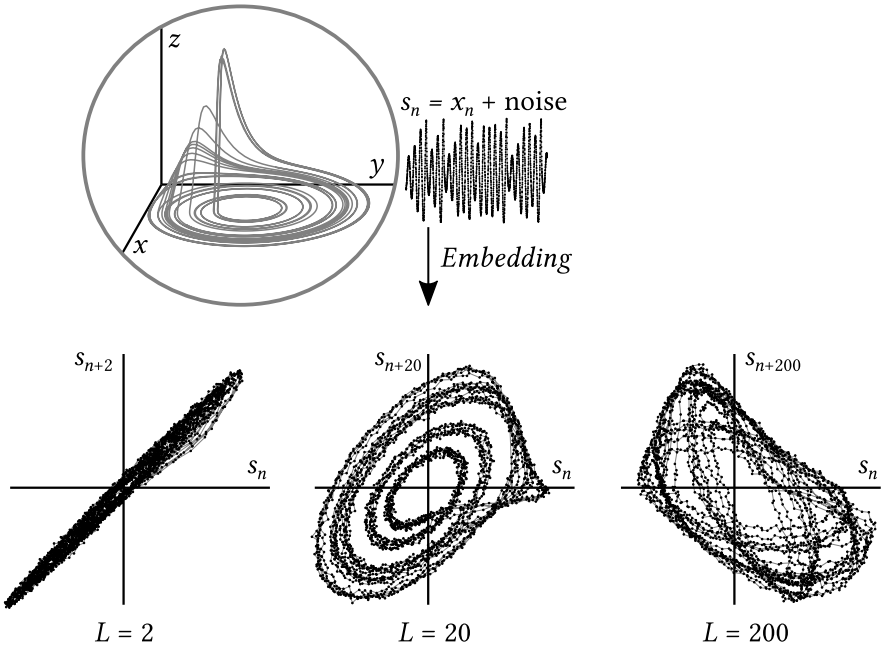


Figure 2.3 – *Sequence generated by integrating a Rössler system corrupted by observational GWN. The sequence is embedded with L = 2, 20, 200. For the sake of this example, m = 2.*

contaminated by observational GWN[3] is embedded with $m = 2$. The choices of too small ($L = 2$) or too large ($L = 200$) lag yield a "bad" reconstructed trajectory: in the former case, the trajectory is squeezed onto the diagonal of the state space and noise blurs the reconstructed evolution; in the latter case, the system's evolution is lost because the components of embedding vectors are uncorrelated. Faithful reconstruction is achieved by intermediate values of L.

2.3 State of the art

The issues outlined in the previous Section prompted the development of many *optimal embedding* methods, i.e. algorithms aimed at choosing the correct m and L to faithfully reconstruct state spaces out of experimental sequences. So far, no criterion to make this choice is considered conclusive [132, 133]. A general solution to the problem is hindered for two reasons. First, it was shown that intrinsic

[2]See Chapter 1, Eq. (1.4) for the system's equations. Here the system parameters are set to $a = 0.15$, $b = 0.2$, $c = 10$, and the integration time step is $\delta t = 0.05$.

[3]Here, the signal (peak-to-peak) amplitude is ~ 32, and the noise standard deviation is 0.2.

statistical fluctuations—which are due to the finiteness of sequences—hamper this search [133]. Second, the optimal embedding criterion depends on both "the time series and the applied measure" [132]: the m and L provided by a given approach might be optimal to estimate some of the quantities that characterize the underlying system, but not others.

Conventionally, optimal embedding is conceived as two separate problems: first, the lag L is optimized; subsequently, the optimal dimension m is determined [2].

In traditional approaches, the optimization of L is carried out by relying on some statistic that measures the degree of independence of L-separated points in the sequence. As discussed in Sec. 2.2, the desired value of L is the one that selects sequence elements that are neither too strongly correlated nor completely uncorrelated with each other. The most straightforward implementation of this criterion consists in computing the autocorrelation function $A(L)$ of the sequence:

$$A(L) = \sum_{n=L+1}^{\ell} \frac{(y_n - \overline{y}) \cdot (y_{n-L} - \overline{y})}{\sqrt{\sum_{n=L+1}^{\ell} (y_n - \overline{y})^2} \sqrt{\sum_{n=L+1}^{\ell} (y_{n-L} - \overline{y})^2}} ,$$

where ℓ is the sequence length and $\overline{y}$ denotes the average over all values of y within the sum. Thereupon, one selects L as the lag at which the $A(L)$ function first passes through zero [11, 2], thus minimizing *linear* dependence between elements of the embedding vectors. Because the systems we are dealing with are nonlinear, one might want to minimize *nonlinear* dependence: to this purpose, the commonly accepted statistic to use is the average mutual information, defined as

$$I(L) = \sum_{n=1}^{\ell-L} P(y_n,\ y_{n+L}) \log \left[\frac{P(y_n,\ y_{n+L})}{P(y_n)\ P(y_{n+L})} \right] ,$$

where $P(y_n,\ y_{n+L})$ is the joint probability of observing y_n, followed by y_{n+L} and $P(y_n)$, $P(y_{n+L})$ are the respective marginal probabilities. In practical implementations, joint and marginal probabilities are estimated by building 2-dimensional histograms of $(y_n\,, y_{n+L})$ or 1-dimensional histograms of y_n, respectively (efficient algorithms to carry out this estimation exist [134]). If the sequence is stationary and $\ell \gg L$, $P(y_n)$ and $P(y_{n+L})$ are the same. The average mutual information $I(L)$ is a measure of how much we learn on y_{n+L} by observing y_n: $I(L)$ is thus a generalization of autocorrelation that takes into account any kind of dependence (and not only linear ones) [134]. The optimal embedding lag for a sequence under investigation is finally selected as the value of L for which $I(L)$ exhibits the first minimum, i.e. the smallest L that minimizes nonlinear dependence between sequence elements [134, 135, 11].

Once the lag L is determined, the dimension m has to be determined. Methods to carry out this selection can be classified into two families.

A first group of methods relies on the estimation, as a function of m, of invariant quantities that char-

acterize the system under analysis. The idea is that, once a sufficiently large m is selected, a further increase of the embedding dimension does not yield significant changes in the estimated quantity. Typically, in this "saturation of system invariants" approach, the invariant quantities considered in the estimation are the correlation dimension or the maximum Lyapunov exponent (see Sec. 2.4 and 2.5, respectively).

A second family of methods relies on the concept of *false nearest neighbours* (FNN), which can be summarized as follows. Suppose that, in an m-dimensional space, two embedding vectors $\mathbf{Y}_i^m$, $\mathbf{Y}_j^m$ are neighbours, i.e. closer than some distance threshold. If the embedding dimension is increased to $m+1$ and the two embedding vectors $\mathbf{Y}_i^{m+1}$, $\mathbf{Y}_j^{m+1}$ are not neighbours in the $m+1$-dimensional space, the vectors $\mathbf{Y}_i^m$, $\mathbf{Y}_j^m$ are deemed to be *false* neighbours: their proximity is due to the reconstruction of the state space evolution being faulty in the m-dimensional space. A sketch of this idea is shown in Fig. 2.4. The

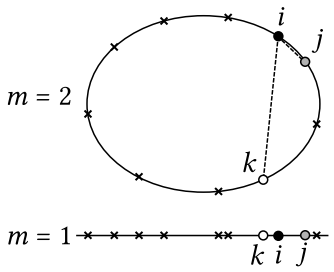


FIGURE 2.4 – *For $m = 1$, the points j, k marked by a gray and a white circle, respectively, are both neighbours of the point i, marked by a black circle. Increasing the dimension to $m = 2$ reveals that point k is a* false *neighbour of i, while points j and i are still close despite the increased dimension of the space. Diagram adapted from Ref. [136].*

points i and j are genuine neighbours in $m = 1$: when m is increased to 2, the corresponding points in the two-dimensional space are still neighbours. On the contrary, the points i and k are false neighbours in $m = 1$: their neighbourhood relationship does not hold anymore in the higher-dimensional space $m = 2$.

The existence of false nearest neighbours is taken as a marker of the inadequateness of a given dimension m to reconstruct the system's evolution: m is typically increased until the number of false nearest neighbours drops below some threshold. The natural way to evaluate the number of false nearest neighbours is to compute Euclidean distances $\|\mathbf{Y}_i - \mathbf{Y}_j\|_m$, $\|\mathbf{Y}_i - \mathbf{Y}_j\|_{m+1}$ in the m- and $m+1$-dimensional embedding space [137]. Two embedding vectors are deemed to be false neighbours in the m-dimensional space if

$$\frac{\sqrt{\|\mathbf{Y}_i - \mathbf{Y}_j\|_{m+1}^2 - \|\mathbf{Y}_i - \mathbf{Y}_j\|_m^2}}{\|\mathbf{Y}_i - \mathbf{Y}_j\|_m} > R_{\text{th}}\,,$$

where R_{th} is a suitably selected threshold. The evaluation of the last expression is carried out by spanning the whole sequence ($i \in [1, \ell - L]$): for each embedding vector $\mathbf{Y}_i$, the nearest neighbour $\mathbf{Y}_j$ is considered [137]. The fraction of false nearest neighbours (i.e. the fraction of i values for which the above expression holds) is computed as a function of the embedding dimension. The value of m above which the fraction of false nearest neighbours vanishes (or, more realistically, falls below an "acceptable

value" [132]), is deemed to be the optimal embedding dimension.

The use of average mutual information to optimize L combined with FNN to determine m is possibly the most common way to assess optimal embedding parameters out of a recorded sequence [5, 2]. The popularity of these two methods is reflected by the number of citations of the papers where they were introduced ($\sim$ 1800 for average mutual information [134] and $\sim$ 1300 for FNN [137], respectively), as well as by their implementation in the TISEAN package [138]—a widely used software package for time series analysis—as quantitative tools dedicated to the choice of embedding parameters. Despite being so popular, the combined use of these two methods does not settle the issue of optimal embedding [5, 139]. For example, as criticized by Garcia & Almeida [140], using mutual information to assess the optimal L has several drawbacks: first of all, the estimated mutual information depends on the kind of binning used to build histograms. Moreover, the approach described above maximizes independence between y_n and y_{n+L}, which does not necessarily imply the independence between y_n and y_{n+2L} or any other successive element of the sequence. In addition, a clear minimum of mutual information might not be present, thus hindering an objective determination of the lag L that maximizes independence [11]. Identifying false neighbours by relying on Euclidean distances is not necessarily the best choice: other metrics might be, for example, more robust against noise [132]. Finally, one might want to optimize *both* m and L with respect to the *same* criterion. The method of Gao and Zheng, the method of characteristic lengths and the method of Schuster aim at determining both embedding parameters by means of a single FNN-like argument, although they differ from the latter in terms of the metric implemented to quantify the amount of false neighbours [132]. These methods are summarized below.

The method of Gao and Zheng [141, 142] is based on the idea that trajectories that evolve out of two neighbouring points separate more quickly if the points are false neighbours with respect to genuine neighbours. This criterion is implemented by first selecting a set $\{(i,j)\}$ of embedding vector pairs $\mathbf{Y}_i, \mathbf{Y}_j$ such that their Euclidean distance $\|\mathbf{Y}_i - \mathbf{Y}_j\|_m$ is smaller than some threshold (i.e. they are neighbours in the m-dimensional space). Thereupon, the following measure is evaluated:

$$\Lambda(\delta k, m, L) = \frac{1}{N_{\text{pairs}}} \sum_{\{(i,j)\}} \log\left(\frac{\|\mathbf{Y}_{i+\delta k} - \mathbf{Y}_{j+\delta k}\|_m}{\|\mathbf{Y}_i - \mathbf{Y}_j\|_m}\right),$$

where N_{pairs} is the size of the set of embedding vector pairs, and $\delta k \in \mathbb{N}^+$. Because the last expression is an estimator of how much neighbouring points separate after a time step δk, $\Lambda(\delta k, m, L)$ is expected to be minimum when optimal embedding is achieved, i.e. when the number of false neighbours is minimized. One thus evaluates $\Lambda(\delta k, m, L)$ for increasing m and L until a minimum is found. This approach has a significant drawback: the choice of δk is arbitrary and is thus based on subjective assessments.

The method of characteristic length follows the same argument as the method of Gao and Zheng,

though bypassing the issue of choosing a time step δk to observe the divergence of trajectories [132]. A characteristic size is first estimated by computing the average distance $\mathcal{J}(m, L) = \langle \|\mathbf{Y}_i - \mathbf{Y}_j\|_m \rangle$ over a randomly selected set of embedding vector pairs. Given two neighbours $\mathbf{Y}_k$, $\mathbf{Y}_h$, the time required by the two points to separate is estimated as the integer δk such that

$$\|\mathbf{Y}_{k+\delta k} - \mathbf{Y}_{h+\delta k}\|_m > \alpha \mathcal{J}(m, L)\,,$$

where $0 < \alpha < 1$. If $\mathbf{Y}_k$, $\mathbf{Y}_h$ are false neighbours, the corresponding trajectories are expected to separate faster than in the case of genuine neighbours, and the corresponding δk is smaller. Taking the average $K = \langle \delta k \rangle$ over a set of neighbouring embedding vectors provides a (relative) measure of the number of false neighbours: the larger is K, the fewer false neighbours are present. The criterion thus consists in identifying the optimal m and L as those for which K is maximized. The same drawback as for the method of Gao and Zheng is present: the assessment of separation is based on a subjectively selected parameter (α).

As the method of Gao and Zheng and the method of characteristic length, the method of Schuster [143] is based on a metric $W(m, L)$ that quantifies the amount of false neighbours. The metric $W(m, L)$ is a measure of how much the set of nearest neighbours of $\mathbf{Y}_i$ changes when the embedding dimension is increased to $m + 1$ (see Ref. [132] for the definition of $W(m, L)$). In this case, $W(m, L)$ is proportional the number of false neighbours. Consequently, one expects $W(m, L)$ to decrease for increasingly large m and L, up to when a further increase of the embedding parameters does not produce significant changes in $W(m, L)$ (the metric $W(m, L)$ exhibits a "limiting behaviour" [132]). The evaluation of this limiting behaviour, however, again relies on a subjective assessment.

A comparative study of the methods described above was published by Cellucci, Albano & Rapp in 2003 [132]. As pointed out by the authors, all the methods mentioned so far require the user to set some parameters, such as R_{th} in the FNN approach, δk in the method of Gao and Zheng, α in the method of characteristic length. Moreover, these methods also require to carry out the evaluation of "limiting behaviours" to assess whether variations in m, L do not produce "significant changes" in the metric of interest. This evaluation is unavoidably subjective [132, 139]. Various attempts to improve and possibly overcome these methods have been published, yielding a plethora of sophisticated techniques for optimal embedding.

Nonuniform embedding, namely an embedding implemented with different lags $L_1 \ldots L_{m-1}$ within the same vector, drew considerable attention [144, 145, 140, 139]: such a reconstruction strategy is expected to perform better when multiple time scales characterize the dynamics of the underlying system [2].

Several works criticized and attempted to improve the FNN method. For example, Hegger & Kantz

[146] evaluated the effects of noise in the selection of neighbouring embedding vectors and thereupon concluded that correlated noise is easily mistaken as chaos by the FNN method. Garcia & Almeida [140] extended the FNN method to nonuniform lags $L_1 \dots L_{m-1}$, claiming that an optimization of lags based on FNN performs better than the conventional mutual information approach. In other cases, the measure used to evaluate the amount of false neighbours was modified, for example by evaluating false nearest "strands", namely trajectory segments (rather than single points) [147, 148], or by replacing—in determining false neighbours—the evaluation of Euclidean distance with the evaluation of the derivative in the direction of the nearest neighbour [149].

Other studies proposed alternative metrics to replace the assessment of false neighbours. Regardless of the specific metric, these approaches boil down to the maximization (or minimization) of a given cost function [150]. For example, Garland & Bradley [151] investigated—with the purpose of time series forecasting—a measure of "shared information" between an embedding vector and its future evolution. In a more recent work [152], minimization of the local curvature of reconstructed trajectories was proposed as a way to assess the optimal lag L (while ignoring the dimension m).

Another key aspect of optimal embedding concerns the so-called *embedding window*, namely the product $(m-1)L$. The embedding window, which (multiplied by the sampling period) corresponds to the time span covered by each embedding vector [153], is deemed to be a relevant parameter in the framework of optimal embedding [130, 2]. Nevertheless, its role is overlooked in most works where optimal embedding criteria are investigated [139].

Finally, some works claimed that optimal embedding parameters obtained by using neural networks performed better than ordinary methods [154, 155]. However, these investigations tested the optimal embedding parameters provided by neural networks only in the context of time series prediction, which is the same task used to train the networks themselves. No clues are given about the performance of these optimal embedding methods when estimating dynamical invariants.

All these improved optimal embedding criteria do not solve the twofold problem introduced before. On the one hand, selection and trimming of the parameters on which these methods rely is arbitrary; on the other hand, the evaluation of "limiting behaviours" is subjective. A significant step towards the solution of this crucial issue was made by Pecora *et al.* [139], who developed an optimal embedding method that relies exclusively on statistical testing: in this way, the selection of heuristic parameters is replaced by the choice of significance threshold, i.e. probabilities, which have a clearer and objective interpretation. The algorithm evaluates a "continuity statistic" [156] to test functional dependence between the sequence elements contained in each embedding vector. The core idea is, again, that optimal attractor reconstruction is the one for which vectors have independent coordinates [139]. Therefore, the dimension and (possibly nonuniform) lags that provide—according to the "continuity" statistical test—functionally independent elements are deemed to be optimal.

Another approach that is based on statistical testing—and thus avoids subjective interpretations—was developed by Carroll & Byers [157]. Given m and L, their method statistically tests whether the embedded sequence is isotropic in the m-dimensional space. Because noise fills isotropically any space when embedded, anisotropy is considered to be a marker of a correct reconstruction of the attractor.
The strategy of assessing probabilities by carrying out statistical tests, rather than relying on heuristic parameters, is a core property of the method discussed in Chapter 3 for the identification of suitable embedding pairs.

2.4 Correlation dimension

As introduced in Chapter 1, one of the fundamental characteristics of chaotic systems is the self-similarity of trajectories within the state space. This property is quantified by estimating the *dimension* of the system's attractor. While linear systems have integer dimension, a noninteger (*fractal*) dimension is a marker of chaoticity [5]. Noise, on the other hand, is characterized by an *infinite* dimension. Because the dimension of a system provides significant information on its dynamics, the estimation of dimensions out of scalar sequences is a key technique in the analysis of nonlinear systems.

The most intuitive way to estimate the dimension of a set consists in assessing how the (hyper) volume V occupied by the set scales as a function of length l:

$$V \propto l^D .$$

The estimation of the dimension D of a set of points[4] within an m-dimensional space is carried out by partitioning the space in hypercubes of size ξ and by counting, as a function of ξ, the fraction η of these hypercubes containing at least one point of the set. For $l \to 0$, the curve $\eta(l)$ allows to estimate the dimension D given that $\eta \propto l^D$. This method to estimate D, which is referred to as *box-counting dimension*, is computationally demanding and sensitive to data length [158, 2].
Given a set of points $\{\mathbf{X}_n\}$ within an m-dimensional space partitioned in hypercubes of size ξ, and given p_i the fraction of points of the set that fall within the i-th hypercube, the idea of dimension can be generalized according to the following expression [1]:

$$D_q = \lim_{\xi \to 0} \frac{1}{q-1} \frac{\log\left(\sum_i p_i^q\right)}{\log(\xi)} .$$

For a uniform fractal (all p_i are equal), D_q is independent of q, and we can choose any q to estimate the dimension. For a nonuniform fractal, D_q decreases with q: computing D_q for $q > 0$ provides an

[4] Formally, the dimension of a set of points is zero. What is actually sought is an estimate of the dimension of the "ideal" underlying set, *represented* by the set of available points that are unavoidably finite in number and in precision [1, 11].

underestimated dimension (the less uniform is the fractal, the larger is the underestimation) [1].
The case $q = 0$ corresponds to the intuitive definition introduced above. The value for $q \to 1$ is called *information dimension* because of the numerator becoming $\sum_i p_i \log(p_i)$, in analogy with information theory. Setting $q = 2$ yields, at the numerator, $\log\left(\sum_i p_i^2\right)$: because p_i^2 is the probability of finding *two* points within the i-th hypercube, D_2 is referred to as *correlation dimension* [11]. D_2 is essentially a measure of pairwise proximity of points: this interpretation prompted the introduction by Grassberger & Procaccia [159, 160] of the so-called *correlation integral* defined as

$$C(r) \equiv \lim_{\ell \to \infty} \frac{1}{\ell^2} \sum_{i \neq j}^{\ell} \theta\left(r - \|\mathbf{X}_i - \mathbf{X}_j\|\right) , \tag{2.2}$$

where ℓ is the total number of points $\mathbf{X}_k$ within the space, and $\theta(x)$ is the Heaviside theta. The key observation by Grassberger & Procaccia was that the correlation integral, for small distances ($r \to 0$), grows as

$$C(r) \propto r^{D_2} . \tag{2.3}$$

By virtue of this last property, it is possible to estimate dimensions out of finite sets of points in a computationally efficient way. In practice, D_2—and thus the correlation integral $C(r)$—has to be estimated out of a scalar sequence $\{y_n\}$. Clearly, the set of points that describe the state space trajectory of the system has to be reconstructed by embedding the scalar sequence. A sample correlation integral is assessed out of an embedded sequence $\{\mathbf{Y}_n\}$ according to

$$\widehat{C}_{m,L}(r) = \frac{1}{N_{\text{pairs}}} \sum_{i,j} \theta\left(r - \|\mathbf{Y}_i - \mathbf{Y}_j\|\right) , \tag{2.4}$$

where the subscript m, L highlights the dependence of the sample correlation integral on the embedding parameters. Sample correlation integrals are computed by considering a number N_{pairs} of pairs i, j of vectors. In principle, one might consider all possible pairs (with $i \neq j$) within the embedded sequence. However, this approach leads to a computationally demanding evaluation, as the number of possible pairs grows quadratically with the sequence length. Alternatively, one proceeds with a Monte Carlo evaluation by fixing the number N_{pairs} and by extracting random pairs of vectors out of the embedded sequence [161].
If a system having dimension D is embedded with $m < D$, the full dimensionality of the system cannot be unfolded and the embedded sequence simply fills all the available space. Consequently, for any $m < D$, correlation integrals exhibit a scaling $\propto r^m$, rather than $\propto r^D$. On the other hand, for $m > D$ and regardless of m, the correct scaling $\propto r^D$ holds [1]. This phenomenon is sometimes exploited to determine the minimum embedding dimension required to reconstruct a system's state space [126].

Figure 2.5 shows some examples of sample correlation integrals in the case of two different input sequences and for different embedding dimensions m (L is set to 4). The correlation integrals in Fig. 2.5(a)

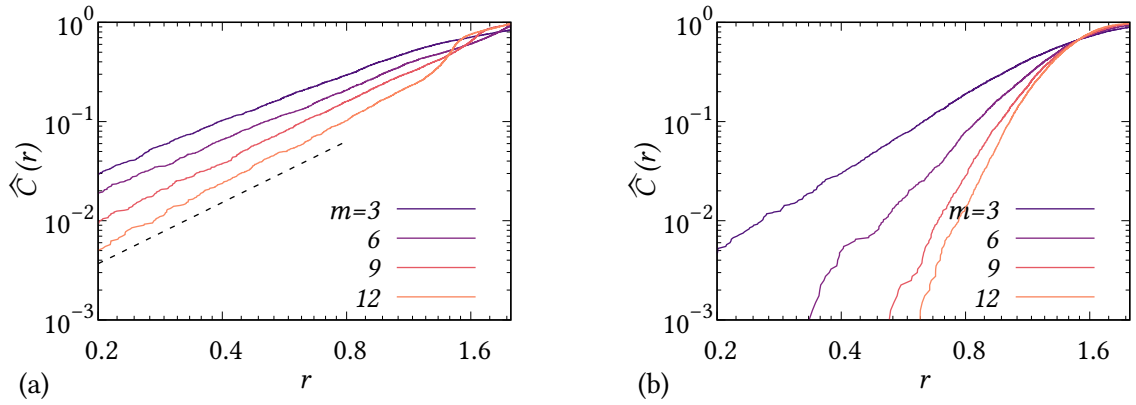


FIGURE 2.5 – *Examples of sample correlation integrals for different embedding dimensions m. The lag L is set to* 4. *(a) The input sequence corresponds to the x-coordinate of a Lorenz system. The black, dashed line highlights the Grassberger-Procaccia power law scaling* $C(r) \propto r^{D_2}$. *(b) The input sequence is GWN.*

are built out of a sequence corresponding to the x-coordinate of a Lorenz system[5]. The $\propto r^{D_2}$ scaling is highlighted by means of a black, dashed line, where $D_2 = 2.05$ is the accepted value of the correlation dimension for the Lorenz system used in this example [14]: increasing m does not change the small-distance scaling of the correlation integral. On the other hand, the correlation integrals in Fig. 2.5(b) are built out of a GWN sequence. For this kind of signal, the Grassberger-Procaccia scaling is not present: increasing m provides an increasingly larger slope of the correlation integral in log-log scale, hinting at a very large (and possibly infinite) dimensionality of the underlying system. Indeed, since noise is characterized by an infinite dimension, correlation integrals corresponding to embedded noise sequences grow as $\propto r^m$: in other words, noise *fills* all the available dimensions of the reconstructed state space, regardless of m.

The estimation of a system's dimension out of a scalar sequence through Eq. (2.4) is among the most popular techniques in nonlinear time series analysis: the two seminal works by Grassberger & Procaccia [159, 160] count more than 3000 citations each. Reliable estimations of D_2, however, are not straightforward [1]: the finite size of sequences and the presence of noise are the main sources of uncertainty and bias in the estimation of the logarithmic slope of correlation integrals [127]. Significant efforts were made to investigate issues related to edge effects [162, 163], to an oscillating behavior of $\widehat{C}(r)$ [164], and to statistical errors produced by the finite number of points used to build $\widehat{C}(r)$ [165, 14]. A source of error regarding the estimation of $C(r)$ are temporal correlations within the input sequence. If we include, within the sum of Eq. (2.4), distances between embedding vectors that are close in time,

[5] See Chapter 1, Eq. (1.2) for the system's equations. Here, the system parameters are set to $\sigma = 10$, $r = 28$, $b = 8/3$ and the integration time step is $\delta t = 0.03$.

the resulting D_2 turns out to be smaller than the true value [166]. Ignoring this issue leads to spurious assessment of finite correlation dimensions even in the case of stochastic sequences, such as noise with $1/f^{\alpha}$ spectrum [167, 168]. This problem is avoided by including in Eq. (2.4) only vector pairs whose time separation is larger than some minimum delay c_0, i.e. by prescribing that $|i-j| \geqslant c_0$. The minimum delay c_0 is typically set equal to the first zero of the autocorrelation function [166], or its first minimum [169].

Sample correlation integrals are the building blocks of the method discussed in Chapter 3.

2.5 Maximum Lyapunov exponent

Besides their geometric properties—quantified by the dimension of the attractor—state space trajectories of chaotic systems exhibit sensitivity to initial conditions or, more quantitatively, an exponential divergence of initially close trajectories (see also Chapter 1). The rate of this divergence is determined by the so-called *Lyapunov exponents*, which are invariants that characterize each specific system [11]. There are as many Lyapunov exponents as the number of state space dimensions: each exponent λ_i corresponds to one of the independent directions along which fiducial volumes within the state space contract ($\lambda_i < 0$) or expand ($\lambda_i > 0$) in an infinitesimal time interval. The maximum Lyapunov exponent (MLE) λ_1 is the most relevant one, as it is the exponent that dominates the divergence of nearby trajectories [5, 1]. The sign of λ_1 provides crucial information on the kind of system under investigation: a negative λ_1 implies that the system is stable and eventually reaches a stable fixed point; if $\lambda_1 = 0$, the system is again stable, but on a limit cycle; finally, a positive λ_1 is the marker of chaos. As far as noise is concerned, its short time behavior is diffusive and distances grow as $\sqrt{t}$: the corresponding diverging rate of expansion suggests an infinite MLE [5].

If the differential equations ruling a system's evolution are known, one can apply the so-called *standard method* [170, 171] and compute numerically *all* Lyapunov exponents. In experimental conditions, if the equations of motions are unknown, this computation is not possible. However, the MLE can be estimated out of scalar sequences by first reconstructing the state space evolution of the system through the embedding procedure. Thereupon, the estimation of the MLE relies on the so-called *divergence rate method* [172, 173, 174]. The divergence rate method is based on the evaluation of the time-dependent divergence exponent $\widehat{\Lambda}_{m,L}(k)$, which quantifies the average separation of nearby trajectories:

$$\widehat{\Lambda}_{m,L}(k) \equiv \frac{1}{N_{\text{pairs}}} \sum_{i,j} \log\left(\frac{\|\mathbf{Y}_{i+k} - \mathbf{Y}_{j+k}\|}{\|\mathbf{Y}_i - \mathbf{Y}_j\|}\right), \tag{2.5}$$

where the sum runs over a set of N_{pairs} pairs i, j of neighbouring vectors, i.e. vectors that satisfy $\|\mathbf{Y}_i - \mathbf{Y}_j\| < r$, with r a "shell radius". The subscript $_{m,L}$ highlights the fact that the vectors $\mathbf{Y}_n$ are em-

bedding vectors built with embedding parameters m and L. The time-dependent divergence exponent measures the progressive separation, as a function of the time delay k, of the trajectories corresponding to initially close points $\mathbf{Y}_i, \mathbf{Y}_j$. The initial proximity of the vector pairs is tuned by changing the shell radius r, which is typically estimated as the distance corresponding to the a given percentile p of the sample distribution of all Euclidean distances [133, 175].

If the underlying system is chaotic, $\widehat{\Lambda}(k)$ is expected to grow linearly with k, as a result of the exponentially increasing numerator within the logarithm of Eq. (2.5). Suitably fitting this linear growth (see Ref. [133] for a robust protocol) provides the value of the MLE. The growth of $\widehat{\Lambda}(k)$ cannot continue for arbitrarily large values of k: eventually, the separation becomes comparable to the size of the attractor and $\widehat{\Lambda}(k)$ saturates.

As discussed in Sec. 2.4 for correlation integrals, the evaluation of $\widehat{\Lambda}(k)$ cannot include pairs of embedding vectors $\mathbf{Y}_i, \mathbf{Y}_j$ separated by a time interval $|i-j|$ that is smaller or comparable to the autocorrelation time of the sequence. Therefore, only those vector pairs whose time separation satisfies $|i-j| \geqslant c_0$ can be included in the sum of Eq. (2.5). As in the case of correlation integrals, the minimum delay c_0 is set equal to the first zero of the autocorrelation function [166] or, more conservatively, to its second zero multiplied by $m(L+1)$ [133].

Figure 2.6 shows some examples of time-dependent divergence exponent in the case of two different input sequences and for different values of the distance percentile p. The curves $\widehat{\Lambda}(k)$ in Fig. 2.6(a) are

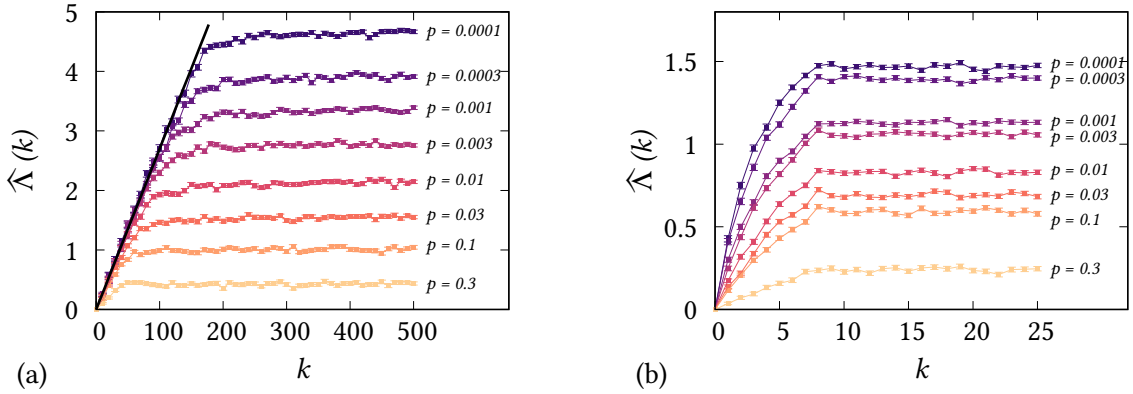


FIGURE 2.6 – *Examples of time-dependent divergence exponent for different distance percentile p. (a) The input sequence corresponds to the x-coordinate of a Lorenz system; the embedding parameters are set to* $m = 8$, $L = 6$. *The black, solid line highlights the region of linear scaling* $\widehat{\Lambda}(k) \propto k$. *(b) The input sequence is GWN; the embedding parameters are set to* $m = 8$, $L = 1$.

built out of a sequence corresponding to the x-coordinate of a Lorenz system[6] embedded with $m = 8$, $L = 6$. The region of linear scaling $\widehat{\Lambda}(k) \propto k$ is highlighted by means of a black, solid line. The slope

[6] See Chapter 1, Eq. (1.2) for the system's equations. Here, the system parameters are set to $\sigma = 10$, $r = 28$, $b = 8/3$ and the integration time step is $\delta t = 0.03$.

of the straight line is $0.9\,\delta t$, where 0.9 is the MLE of the system determined by means of the standard method and δt is the sampling period. Figure 2.6(b) shows, instead, the curves $\widehat{\Lambda}(k)$ in the case of a GWN sequence embedded with $m = 8$, $L = 1$. A linearly growing region is, in this case, lacking. $\widehat{\Lambda}(k)$ quickly saturates at $k = m$ (in this example, $m = 8$), which corresponds to the delay for which a vector $\mathbf{Y}_{i+k}$ and its precursor $\mathbf{Y}_i$ do not share any component and, therefore, are uncorrelated.

Besides the linearly growing $\widehat{\Lambda}(k)$ within the small k region, which is processed to estimate the MLE, the asymptotic "plateau-like" behavior attained for $k \gg 1$ is also interesting. In particular, the dependence of the asymptotic $\widehat{\Lambda}(k)$ on the distance percentile p provides a tool to distinguish finite-dimensional chaotic systems from stochastic sources and, for the former ones, it allows to precisely estimate correlation dimension. These results are the topic of Chapter 5.

3. Identification of suitable embedding regions

As discussed in Chapter 2, a general solution to the problem of optimally setting the embedding parameters m, L is lacking. Existing methods have several drawbacks, most notably the arbitrary choice of parameters, the subjective evaluation of "limiting behaviours", and the separation of the optimization of m and L into two distinct problems, thus overlooking the crucial role of the embedding window $(m-1)L$. Finally, a *single* optimal choice (m, L) does not necessarily exist: more realistically, a set of (m, L) can provide reconstructions of the underlying dynamics that are equivalently good.

In this chapter, I will discuss the development, carried out within our research group, of a new method for the identification of suitable embedding dimensions and lags [161]. We tackled the issue of optimal embedding by trying to overcome the drawbacks that characterize existing methods. To this purpose, our method does not provide a *single* choice of parameters, but rather a *set* of suitable embedding pairs (m, L) that are statistically equivalent to reconstruct the dynamics. In addition, our approach relies on statistical testing and does not require any subjective evaluation of results: the only parameters to be chosen are significance thresholds, i.e. probabilities, which have a straightforward interpretation.

A graphical outline of the method is shown in Fig. 3.1. Given an input scalar sequence, several operations are carried out for each element of an *embedding lattice*, i.e. a discrete set of choices of the embedding parameters $\{(m, L) \,|\, m \in [2, m_{\max}],\ L \in [1, L_{\max}],\}$. In the following, $m_{\max}$ and $L_{\max}$ are both set to 20. For each *embedding pair* or *embedding point* (m, L), sample correlation integrals are built as described in Sec. 3.1. The analytical expression of correlation integrals corresponding to Gaussian white noise (GWN) allows us to introduce a "gauge" transformation of distances. As a consequence, correlation integrals are remapped into a more convenient form, henceforth referred to as "correlation bridges". In addition to these preprocessing steps, the core of the method consists of three stages. The first two stages rely on statistical tests to rule out embedding pairs that are unsuitable to embed the sequence: first, correlation bridges are tested for compatibility with GWN sources (Sec. 3.2.1); second,

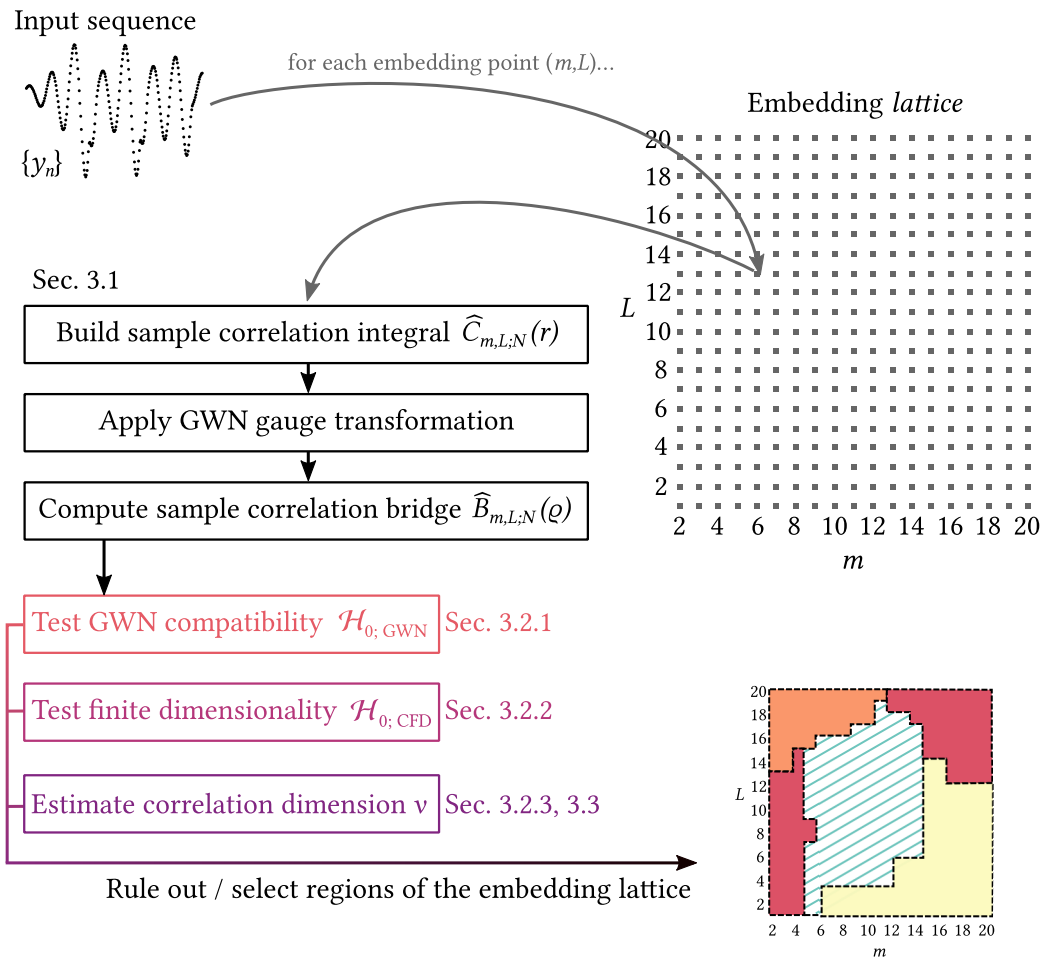


FIGURE 3.1 – *Graphical summary of the method described in this chapter. Given an input scalar sequence, the processing stages within the flowchart are carried out for each* embedding point (m, L) *within the* embedding lattice. *As a result, the procedure yields a* map *of the embedding lattice that provides information on which* (m, L) *are suitable or not to embed the input sequence.*

correlation bridges are tested for compatibility with a finite-dimensional source (Sec. 3.2.2). In the third stage, correlation dimension is estimated out of correlation bridges on those embedding pairs that survived the previous skimming stage (Sec. 3.2.3, 3.3). The result of these stages is a *map* of the embedding lattice in which some points are deemed to be unsuitable to embed the sequence, while a set of suitable embedding pairs is identified by looking at the uniformity of the correlation dimension estimator.

Suitable embedding choices are expected to comply with two requirements concerning the corresponding embedding window $W = (m - 1)L$. On the one hand, the embedding window has to be large enough to overcome *redundance*. On the other hand, W cannot be too large, so that *irrelevance* is avoided [153, 130]. By calling τ_R and τ_I the *redundance* and *irrelevance* times, respectively, the above

requirements can be expressed as $\tau_R/T \lesssim (m-1)L \lesssim \tau_I/T$, where T is the sampling period of the input sequence.
These constraints have a straightforward interpretation on the embedding lattice: a given value of W identifies a hyperbola. Therefore, the irrelevance and redundance times correspond to two hyperbolae within the embedding lattice, and the region bounded by them is where we can expect to find suitable embedding choices. As shown in Sec. 3.2.3, the regions of suitable (m, L) identified by our method are indeed in agreement with these constraints.

Throughout this chapter, the stages of the method are illustrated by applying the corresponding algorithms to three benchmark sequences, defined as follows:

- a **GWN sequence**, built by generating i.i.d. random numbers out of a standard normal distribution;

- a sequence corresponding to **filtered GWN**, generated by first drawing i.i.d. random numbers out of a standard normal distribution and, thereupon, by piping the resulting sequence into a first-order digital low-pass filter with cut-off frequency equal to 1/20 of the sampling frequency;

- a sequence corresponding to the x-coordinate of a Lorenz system, or **Lorenz sequence**. See Chapter 1, Eq. (1.2) for the system's equations; here the system's parameters are set to $\sigma = 10$, $r = 28$, $b = 8/3$, and the integration time step, which coincides with the sampling period, is $\delta t = 0.03$.

3.1 Correlation integrals and GWN gauge

The method discussed in this chapter takes in input a scalar sequence $\{y_n\}$ consisting of ℓ points. As a preprocessing step, the sequence is standardized by subtracting its sample mean $\overline{y}$ and dividing by its sample standard deviation s_y each element: $(y_n - \overline{y})/s_y \to y_n$. Thereupon, the sequence is embedded with parameters m, L, so that a sequence of vectors $\mathbf{Y}_n = (y_n, y_{n+L}, \ldots, y_{n+(m-1)L})$ having length $\ell - (m-1)L$ is provided.

A sample correlation integral corresponding to the sequence y_n and embedded with parameters m, L is built by means of a Monte Carlo integration as follows. A set $\mathcal{S}$ of N vector pairs $\mathbf{Y}_i$, $\mathbf{Y}_j$ are randomly selected without replacement. As discussed in Sec. 2.4, temporally correlated vector pairs have to be avoided, and thus the time indexes i, j are chosen such that $|i - j| \geqslant c_0$, where c_0 is set equal to the largest between the second zero of the autocorrelation function and the product $2m(L+1)$. The

distance $d_{i,j}$ between vectors $\mathbf{Y}_i$, $\mathbf{Y}_j$ is evaluated as their Euclidean distance divided by[1] $\sqrt{m}$:

$$d_{i,j} = \left[\frac{1}{m} \sum_{k=0}^{m-1} \left(y_{i+kL} - y_{j+kL} \right)^2 \right]^{\frac{1}{2}} . \tag{3.1}$$

The sample correlation integral $\widehat{C}_{m,L;N}(r)$ is given by the empirical cumulative distribution of the N sample distances:

$$\widehat{C}_{m,L;N}(r) = \frac{1}{N} \sum_{(i,j)\in\mathcal{S}} \theta \left(r - d_{i,j} \right) . \tag{3.2}$$

The parameter N sets the resolution of the sample correlation integral and, at the same time, poses a constraint to the sequence length: provided that the sequence autocorrelation time is much smaller than the sequence length, i.e. $\ell \gg c_0$, the number of available pairs ($\sim \ell^2/2$) has to satisfy $\ell^2/2 \gg N$ to provide a reliable integration. In the following, $\ell = 10^5$ and $N = 4000$; the maximum observed c_0 is of order 10^3.

Figure 3.2 shows four examples of correlation integrals $\widehat{C}_{m,L;N}(r)$ built out of a GWN sequence and a Lorenz sequence. Embedding pairs (m, L) are reported in the figure key.

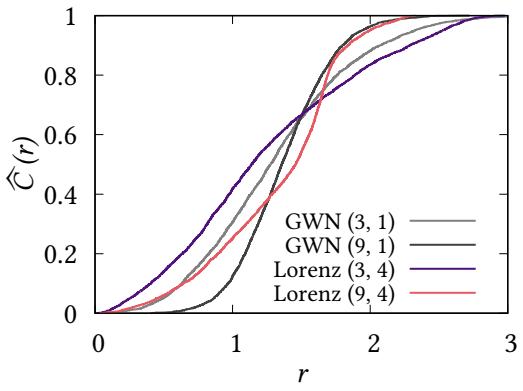


FIGURE 3.2 – *Example correlation integrals $\widehat{C}_{m,L;N}(r)$ for a GWN sequence and a sequence generated by the Lorenz system. Embedding dimensions and lags (m, L) are reported in the key.*

Sample correlation integrals $\widehat{C}_{m,L;N}(r)$ converge to the corresponding asymptotic limit[2] $C_{m,L}(r)$, henceforth referred to as "asymptotic correlation integral", when $N \to \infty$. The analytical assessment of $C_{m,L}(r)$ is impossible for most dynamical systems, even if the corresponding differential equations are known. On the other hand, GWN provides an exception: given a sequence of i.i.d. elements drawn from a normal distribution ($y_n \sim \mathcal{N}(\mu, \sigma^2)$) it is possible to write down $C_{m,L}(r)$ for any m, L. For a GWN sequence, $|i - j| \geqslant c_0 = 2m(L + 1) \geqslant 2$, since the second zero of the autocorrelation function corresponds to a delay equal to 2. Consequently, all elements entering into the sum of Eq. (3.1) are independent, and thus the differences $y_{i+kL} - y_{j+kL}$ is distributed according to a normal distribution with vanishing

[1] The normalization by $\sqrt{m}$ is introduced to provide an m-independent distance in the case $L = 0$.

[2] This result is guaranteed by the Glivenko-Cantelli theorem.

mean and variance equal to $2\sigma^2$. Because $d_{i,j}^2$ is the sum of the squares of m such differences, it follows—besides a scaling factor—a χ^2 distribution with m degrees of freedom, i.e. $d_{i,j}^2 \sim (2\sigma^2/m)\chi_m^2$. This last results holds if $L > 0$: if $L = 0$, regardless of m, we have instead $d_{i,j}^2 \sim 2\sigma^2\chi_1^2$. Asymptotic correlation integrals for GWN sequences are thus given by

$$C_{m;\mathrm{GNW}}(r) = \begin{cases} \dfrac{1}{\Gamma\left(\frac{m}{2}\right)}\gamma\left(\frac{m}{2}, \frac{mr^2}{4\sigma^2}\right), & \text{if } L > 0\,, \\ \\ \dfrac{1}{\Gamma\left(\frac{1}{2}\right)}\gamma\left(\frac{1}{2}, \frac{r^2}{4\sigma^2}\right), & \text{if } L = 0\,, \forall\ m\,. \end{cases} \tag{3.3}$$

The function $\gamma(s, x)$ is the *lower incomplete Gamma function*, defined as $\gamma(s, x) = \int_0^x t^{s-1}e^{-t}\mathrm{d}t$.

The analytical shape of GWN correlation integrals allows us to transform distances so that sample correlation integrals built out of GWN sequences are drawn from the same, embedding-independent distribution. This "gauge" transformation consists in mapping the distances $d_{i,j}$ into transformed distances $\delta_{i,j}$ by applying the analytic expression of Eq. (3.3):

$$\delta_{i,j} = C_{m;\mathrm{GNW}}\left(d_{i,j}\right) = \begin{cases} \dfrac{1}{\Gamma\left(\frac{m}{2}\right)}\gamma\left(\frac{m}{2}, \frac{md_{i,j}^2}{4}\right), & \text{if } L > 0\,, \\ \\ \dfrac{1}{\Gamma\left(\frac{1}{2}\right)}\gamma\left(\frac{1}{2}, \frac{d_{i,j}^2}{4}\right), & \text{if } L = 0\,, \forall\ m\,, \end{cases} \tag{3.4}$$

where σ is set to 1 in agreement with the standardization step described above. In this new gauge, sample correlation integrals $\widehat{C}'_{m,L;N}(\varrho)$ are evaluated by transforming "ordinary" sample distances $d_{i,j}$ into $\delta_{i,j}$; Eq. (3.2) becomes

$$\widehat{C}'_{m,L;N}(\varrho) = \frac{1}{N}\sum_{(i,j)\in\mathcal{S}}\theta\left(\varrho - \delta_{i,j}\right)\,.$$

Because $C_{m;\mathrm{GNW}}$ is a strictly monotonous function, we have that

$$\varrho = C_{m;\mathrm{GNW}}(r) \quad \Longrightarrow \quad \widehat{C}_{m,L;N}(r) = \widehat{C}'_{m,L;N}(\varrho)\,.$$

This last equality is the reason why we called Eq. (3.4) a "gauge transformation": the metric through which distances are measured is changed, but the information contained within $\widehat{C}_{m,L;N}(r)$ and $\widehat{C}'_{m,L;N}(\varrho)$ is the same.

Because $C_{m;\mathrm{GNW}}$ is a cumulative distribution function—and thus bounded between 0 and 1—the gauge transformation defined by Eq. (3.4) changes the d domain given by $\mathbb{R}^{0+}$ into the δ domain given by $[0, 1)$. Therefore, the gauge-transformed correlation integrals $\widehat{C}'_{m,L;N}(\varrho)$ are fully contained within

the unitary square $[0,1] \times [0,1]$. The correlation integrals of Fig. 3.2, upon application of the gauge transformation, are mapped into the correlation integrals shown in Fig. 3.3.

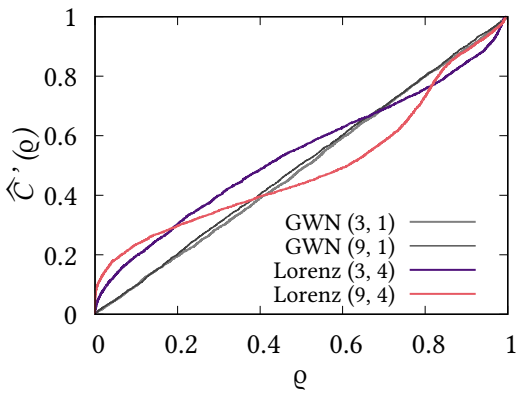


FIGURE 3.3 – *Four correlation integrals shown in Fig. 3.2 after the application of the gauge transformation of distances. Embedding dimensions and lags (m, L) are reported in the key.*

In the new gauge, correlation integrals $\widehat{C}'_{m,L;N}(\varrho)$ corresponding to a GWN sequence turn out to have the same shape (besides statistical fluctuations due to the finiteness of N). This result is due to the probability integral transform theorem [176]: if a random variable z follows a cumulative distribution function $F_z(z)$, then the mapping $\zeta = F_z(z)$ yields a new random variable ζ distributed according to a uniform distribution within $[0,1]$. In other words, $F_\zeta(\zeta) = \xi(\zeta) = \zeta$, with $\zeta \in [0,1]$. Therefore, in the new gauge, the asymptotic correlation integrals corresponding to GWN turn out to be given by $C'_{m,L}(\varrho) = \xi(\varrho)$, regardless of the embedding parameters m, L. Sample correlation integrals of GWN sequences are expected to be statistically compatible with $\xi(\varrho)$. Statistical compatibility between cumulative distributions is assessed by means of the Kolmogorov-Smirnov (K-S) test, which relies on the K-S statistic $\widehat{K}_{m,L;N}$ defined as

$$\widehat{K}_{m,L;N} = \sup_{\rho \in [0,1]} \left| \widehat{C}'_{m,L;N}(\varrho) - \xi(\varrho) \right| = \sup_{\rho \in [0,1]} \left| \widehat{C}'_{m,L;N}(\varrho) - \varrho \right| . \tag{3.5}$$

In the null hypothesis of the sample $\widehat{C}'_{m,L;N}(\varrho)$ being drawn from the parent asymptotic cumulative distribution $\xi(\varrho)$, the K-S statistic is distributed as [177]

$$F(K) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp\left(-2k^2 N' K^2\right) , \tag{3.6}$$

where the parameter N' is the effective number of samples: $N' = (\sqrt{N} + 0.12 + 0.11/\sqrt{N})^2$. Mean μ_K and standard deviation σ_K of $F(K)$ can be evaluated by taking its first derivative $\mathrm{d}F(K)/\mathrm{d}K$ and integrating

over $[0,\infty)$:

$$\begin{aligned}
\mu_K = \mathrm{E}[K] &= \int_0^\infty \frac{\mathrm{d}F(K)}{\mathrm{d}K} K \mathrm{d}K = 8N' \sum_{k=1}^{\infty} (-1)^{k-1} k^2 \int_0^\infty K^2 e^{-2k^2 N' K^2} \mathrm{d}K = \\
&= \sqrt{\frac{\pi}{2N'}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = \sqrt{\frac{\pi}{2N'}} \log 2\,, \qquad (3.7a) \\
\mathrm{E}[K^2] &= \int_0^\infty \frac{\mathrm{d}F(K)}{\mathrm{d}K} K^2 \mathrm{d}K = 8N' \sum_{k=1}^{\infty} (-1)^{k-1} k^2 \int_0^\infty K^3 e^{-2k^2 N' K^2} \mathrm{d}K = \\
&= 8N' \sum_{k=1}^{\infty} \frac{(-1)^{k-1} k^2}{2} \int_0^\infty \kappa e^{-2k^2 N' \kappa} \mathrm{d}\kappa = \\
&= \frac{1}{N'} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k^2} = \frac{1}{N'} \frac{\pi^2}{12}\,, \\
\sigma_K &= \sqrt{\mathrm{E}[K^2] - \mathrm{E}[K]^2} = \sqrt{\frac{\pi}{2N'}} \sqrt{\frac{\pi}{6} - \log^2 2}\,. \qquad (3.7b)
\end{aligned}$$

Both the mean and the standard deviation of $F(K)$ tend to zero when $N' \to \infty$, as expected.

3.2 Skimming and selecting suitable embedding pairs

3.2.1 Correlation bridges and the null hypothesis of a GWN source

The K-S statistic of Eq. (3.5) suggests the possibility of *adjusting* correlation integrals in a more convenient form: we define a "sample correlation bridge" $\widehat{B}_{m,L;N}(\varrho)$ as a gauge-transformed correlation integral $\widehat{C}'_{m,L;N}(\varrho)$ to which we subtract the reference line $\xi(\varrho)$:

$$\widehat{B}_{m,L;N}(\varrho) \equiv \widehat{C}'_{m,L;N}(\varrho) - \xi(\varrho) = \widehat{C}'_{m,L;N}(\varrho) - \varrho\,.$$

As in the case of correlation integrals, sample correlation bridges converge to the corresponding asymptotic correlation bridges for $N \to \infty$. A correlation bridge is a function defined in the interval $[0,1]$, taking up values between $[-1,1]$, and tied to zero at both ends, i.e. $\widehat{B}_{m,L;N}(\varrho) = 0$ for $\varrho = 0, 1$. Asymptotic correlation bridges for GWN are identically zero, while their sample counterparts turn out to be Brownian bridges [178]. The K-S statistic defined in Eq. (3.5) becomes

$$\widehat{K}_{m,L;N} = \sup_{\rho \in [0,1]} \left| \widehat{B}'_{m,L;N}(\varrho) \right|\,. \qquad (3.8)$$

Given a significance level α, the corresponding critical K-S statistic $K_{\alpha;N}$ is evaluated by relying on Eq. (3.6) and inverting $F(K_{\alpha;N}) = 1 - \alpha$. By virtue of Eq. (3.8), a sample correlation bridge built out of a

GWN sequence is fully contained within $\pm K_{\alpha;N}$ with probability $1-\alpha$. Figure 3.4 shows the correlation bridges corresponding to the correlation integrals of Fig. 3.3. The plots of Fig. 3.4 also show two shaded areas bounded by $\pm K_{\alpha;N}$: the inner one corresponds to the significance level $\alpha = 0.05$, while the outer one corresponds to $\alpha = 10^{-9}$. The sample correlation bridges corresponding to a GWN sequence turn

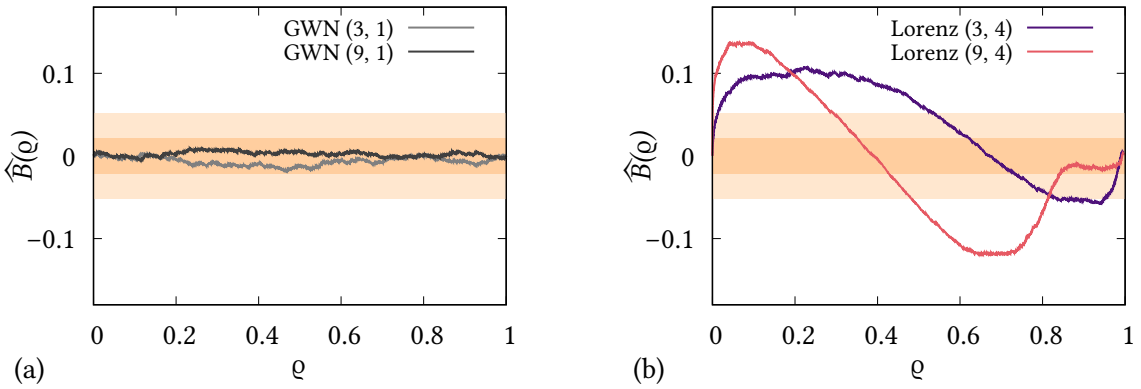


FIGURE 3.4 – *Correlation bridges $\widehat{B}_{m,L}(\varrho)$ for a GWN sequence (a) and a sequence generated by the Lorenz system (b). Each correlation bridge is obtained by subtracting the identity line $\xi(\varrho) = \varrho$ from the corresponding correlation integral of Fig. 3.3. Shaded areas correspond to the regions bounded by $K_{\alpha;N}$ for $\alpha = 0.05$ (darker, inner region) and $\alpha = 10^{-9}$ (lighter, outer region).*

out to be fully contained within the $\alpha = 0.05$ region; on the contrary, the two sample correlation bridges built out of a Lorenz sequence overcome the boundaries of the $\alpha = 10^{-9}$ region.

The K-S statistic $\widehat{K}_{m,L;N}$ estimates the compatibility of a given input sequence embedded with parameters m, L with GWN. For this reason, $\widehat{K}_{m,L;N}$ is henceforth referred as the sample "gap from noise".

The first stage of the method discussed in this chapter relies on the gap from noise—and the related statistical test—to rule out unsuitable embedding pairs (m, L). Under the assumption that a given input sequence is generated by a chaotic source, the idea is that, given an embedding pair (m, L), if the corresponding correlation bridge is compatible with a GWN source then the embedding pair (m, L) is unsuitable to reconstruct the sequence's underlying dynamics. As discussed at the beginning of this chapter, this evaluation is carried out on the embedding lattice, i.e. on the set of embedding pairs $\{(m, L) \,|\, m \in [2, 20],\ L \in [1, 20],\}$. A sample correlation bridge is evaluated for each point on the lattice and the corresponding gap from noise is computed. In principle, Eq. (3.6) allows to assess a p value corresponding to the value of $\widehat{K}_{m,L;N}$. However, to avoid the evaluation of the sum in Eq. (3.6), as well as to improve the robustness of the assessment, the computation of $\widehat{K}_{m,L;N}$ is repeated M times and, for each point on the lattice, the average gap from noise $\langle \widehat{K} \rangle_{m,L;N}$ is calculated, yielding a *map* of average gap from noise on the lattice: $\{\langle \widehat{K} \rangle_{m,L;N} \,|\, m \in [2, 20],\ L \in [1, 20],\}$. In the following, $M = 100$.

The null hypothesis $\mathcal{H}_{0;\mathrm{GWN}}$ to be tested is that the input sequence is GWN. Under $\mathcal{H}_{0;\mathrm{GWN}}$, provided

that $M \gg 1$, the averages $\langle \widehat{K} \rangle_{m,L;N}$ are approximately distributed according to a normal distribution with mean μ_K and variance σ_K^2/M by virtue of the central limit theorem, where μ_K, σ_K are given by Eq. (3.7). Therefore, for each point on the embedding lattice, $\mathcal{H}_{0;\mathrm{GWN}}$ is tested by computing a p value out of a $\mathcal{N}\left(\mu_K, \frac{\sigma_K^2}{M}\right)$ distribution, finally yielding a map of p values.
Figure 3.5 shows the maps of p values evaluated for the three benchmark sequences, namely a GWN sequence, a filtered noise sequence, and a Lorenz sequence. In the case of GWN, p values are all above

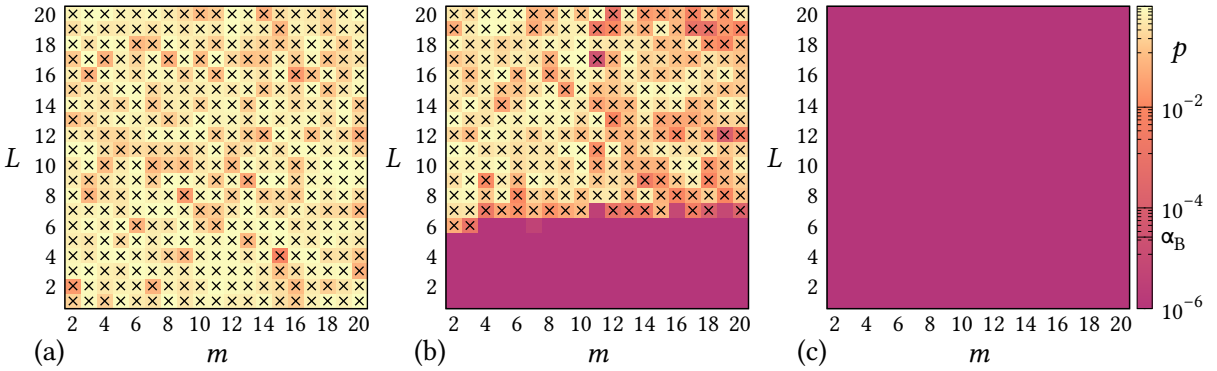


FIGURE 3.5 – *Maps of p values evaluated on the embedding lattice under $\mathcal{H}_{0;GWN}$ for a GWN sequence (a), a filtered noise sequence (b), and a Lorenz sequence (c). Black crosses (×) mark points for which the p value overcomes the $\alpha_B = 0.01/380$ threshold.*

the Bonferroni-corrected 1% significance threshold $\alpha_B = 0.01/380$ (the number of lattice points is 380): the input sequence is statistically compatible with GWN for any choice of the embedding parameters. The presence of embedding pairs for which the input sequence is incompatible with GWN, i.e. the corresponding p value is below α_B such as in the case of filtered noise and Lorenz sequence, indicates a non-GWN source. In the case of filtered noise, pairs with $L \gtrsim 6$ are compatible with the GWN hypothesis. This compatibility is due to the loss of correlation between elements of the embedding vectors when L becomes comparable to the typical autocorrelation time of the sequence: when "sampled" too slowly, filtered noise looks like GWN. On the contrary, the Lorenz sequence turns out to be incompatible with GWN regardless of the embedding parameters: the corresponding p values are all below the significance threshold α_B.

As a result of this first skimming procedure, embedding pairs (m, L) for which there is compatibility with GWN are deemed to be unsuitable to embed the input sequence. These embedding lattice points, marked with a black cross in Fig. 3.5, are ruled out from subsequent evaluations.

3.2.2 Small-scale behavior of correlation bridges

The second stage of the method consists in testing the finite dimensionality of the dynamical system generating the input sequence. The corresponding null hypothesis $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ is that of a chaotic, finite dimensional system with correlation dimension ν generating the sequence. In order to test this hypothesis, we first consider the derivative of (asymptotic) correlation bridges:

$$\frac{\mathrm{d}B_{m,L}(\varrho)}{\mathrm{d}\varrho} = \frac{\mathrm{d}C'_{m,L}(\varrho)}{\mathrm{d}\varrho} - 1 = \frac{\partial C_{m,L}(r)}{\partial r}\frac{\mathrm{d}r(\varrho)}{\mathrm{d}\varrho} - 1\,, \tag{3.9}$$

where $r(\varrho)$ is the inverse of the function $\varrho(r) = C_{m;\mathrm{GNW}}(r)$. By using the definition of $C_{m;\mathrm{GNW}}(r)$, i.e. Eq. (3.3), and by applying the inverse function theorem, the derivative of $r(\varrho)$ can be written as

$$\frac{\mathrm{d}r(\varrho)}{\mathrm{d}\varrho} = \left(\frac{\mathrm{d}\varrho(r)}{\mathrm{d}r}\right)^{-1} = \frac{1}{2}\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{4\sigma^2}\right)^{-\frac{m}{2}}\exp\left(\frac{mr^2}{4\sigma^2}\right)r^{1-m}\,. \tag{3.10}$$

Under $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$, according to Grassberger & Procaccia [159, 160], the asymptotic correlation integral $C_{m,L}(r)$ grows, for $r \to 0$, as a power of the distance $C_{m,L}(r) = \beta_{m,L}r^{\nu}$; $\beta_{m,L}$ is an embedding-dependent constant (see also Sec. 2.4). The derivative of $C_{m,L}(r)$ is thus given by

$$\frac{\mathrm{d}C_{m,L}(r)}{\mathrm{d}r} = \nu\beta_{m,L}r^{\nu-1} = \frac{\nu}{r}C_{m,L}(r)\,. \tag{3.11}$$

This last expression is only valid as long as $m > \nu$. Combining Eq. (3.9), Eq. (3.10) and Eq. (3.11) finally yields

$$\frac{\mathrm{d}B_{m,L}(\varrho)}{\mathrm{d}\varrho} = \frac{\nu\beta_{m,L}}{2}\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{4\sigma^2}\right)^{-\frac{m}{2}}\exp\left(\frac{mr^2}{4\sigma^2}\right)r^{\nu-m} - 1\,. \tag{3.12}$$

The consequence of the last expression is that, under $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ and provided that $m > \nu$, the derivative of $B_{m,L}(\varrho)$ diverges for $\varrho \to 0$. A lack of this divergence implies that the system generating the sequence under investigation, reconstructed with embedding parameters m, L, is not finite dimensional, either because $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ does not hold, or because $\nu \geqslant m$. While in the former case embedding is useless, in the latter case it is unsuitable. Therefore, assessing the divergent behaviour of correlation bridges for $\varrho \to 0$ provides an additional tool to rule out embedding pairs that are unsuitable to reconstruct the dynamics.

The divergent behaviour, as discussed so far, is a feature of *asymptotic* correlation bridges. However, this divergence has to be assessed by relying on *sample* correlation bridges built out of finite sequences: a statistical approach has to be followed. Given an input sequence and an embedding pair (m, L), the number N is supposed to be large enough so that the first sample of the correlation integral, corresponding to a distance r_1, falls within the Grassberger-Procaccia power-law scaling region. The

corresponding gauge transformed distance ϱ_1 provides the first value of the correlation bridge according to $b_1 = \widehat{B}_{m,L;N}(\varrho_1) = 1/N - \varrho_1$. A number M of correlation bridges are independently built, yielding M values of b_1. If any of these values is negative or zero, a divergent behaviour is considered disproved. If all the M collected b_1 values are positive, the probability of a divergent behaviour to exist is computed by relying on Chebyshev's inequality[3], as follows. The sample mean and standard deviation of the M values b_1, denoted with $\bar{b}_1$, s_{b_1}, respectively, are considered. By virtue of Chebyshev's inequality, we have

$$P\left(b_1 \leqslant \bar{b}_1 - k s_{b_1}\right) \leqslant P\left(|b_1 - \bar{b}_1| \geqslant k s_{b_1}\right) \leqslant \frac{1}{k^2},$$

where $k > 0$. Because $\bar{b}_1$ is positive, and setting $k = \bar{b}_1 / s_{b_1}$, we finally obtain

$$P\left(b_1 \leqslant 0\right) \leqslant \left(\frac{s_{b_1}}{\bar{b}_1}\right)^2 . \tag{3.13}$$

The last expression makes up a direct way to compute a p value out of the M values of b_1. If the probability of having $b_1 \leqslant 0$ is higher than some significance threshold α, the pair (m, L) is deemed to be unsuitable to embed the input sequence, as the corresponding reconstructed dynamics is incompatible with a finite-dimensional system.

The divergent behaviour is a more stringent marker compared to the compatibility with GWN: indeed, embedding points for which $\mathcal{H}_{0;\mathrm{GWN}}$ is not rejected automatically provide a rejection of $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$. In other words, embedding points excluded through the criterion of Sec. 3.2.1 are necessarily excluded by Eq. (3.13) as well.

The evaluation of Eq. (3.13) on the embedding lattice for the three sequences considered in Fig. 3.5 is shown in Fig. 3.6. In the case of a GWN sequence, this second skimming stage does not provide additional information: all embedding points were excluded by the first criterion, and thus the whole embedding lattice is incompatible with $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$. As far as filtered noise is concerned, most of the embedding points that survived the first stage are deemed to be incompatible with $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$, i.e. the filtered noise sequence, embedded on those points, is correctly classified as not being finite-dimensional. Interestingly, in the case of the Lorenz sequence, almost all embedding points with $m = 2$ are ruled out because of their incompatibility with $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$, i.e. the reconstructed state space trajectory does not appear as being finite-dimensional. Indeed, the system *cannot* be embedded with $m = 2$: such reconstruction would give trajectories with crossings, which is forbidden by determinism, as correctly detected by the skimming stage discussed in this Section.

[3] This approach is the most conservative one, as Chebyshev's inequality provides an *upper* limit for the p value.

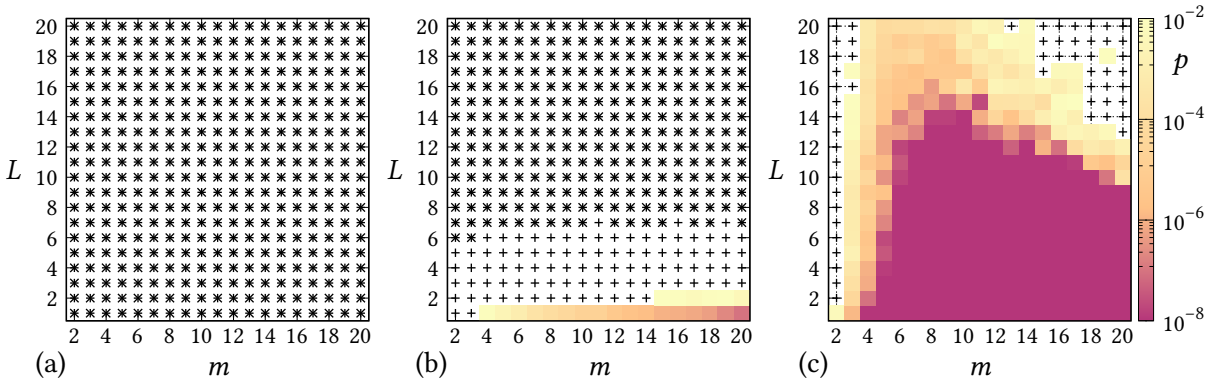


FIGURE 3.6 – *Maps of the probability evaluated according to Eq. (3.13) on the embedding lattice for a GWN sequence (a), a filtered noise sequence (b), and a Lorenz sequence (c). Black "+" signs mark points for which the probability overcomes the* 0.01 *significance threshold: these embedding points correspond to correlation bridges that are deemed to be incompatible with* $\mathcal{H}_{0;CFD}(\nu)$*. Black crosses (×) mark those points for which* $\mathcal{H}_{0;GWN}$ *holds, as in Fig. 3.5.*

3.2.3 Uniformity of the estimated ν within an embedding lattice

Correlation bridges can be processed to get an estimate of the correlation dimension $\widehat{\nu}$: this estimation is discussed in Sec. 3.3. The third stage of the method consists in using the estimated $\widehat{\nu}(m, L)$ to assess which embedding pairs are suitable to reconstruct the dynamics of the system generating an input sequence. This assessment relies on the idea that, if the null hypothesis $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ holds, and the embedding parameters satisfy the relation

$$\tau_R/T \lesssim (m-1)L \lesssim \tau_I/T\,,$$

i.e. the embedding window is smaller than the irrelevance time and larger than the redundance time of the input sequence, then the estimated correlation dimension $\widehat{\nu}$ is expected to be independent of the embedding pair (m, L). Therefore, while the statistical tests of Sec. 3.2.1 and 3.2.2 allowed us to *rule out* embedding pairs, here we can select *suitable* embedding pairs by identifying a region of the embedding lattice where $\widehat{\nu}$ is uniform. By virtue of Eq. (3.14), this region has to be approximately bounded by two hyperbolae, corresponding to the irrelevance and redundance times.

Figure 3.7(a) shows the map of the estimated $\widehat{\nu}$ for the same Lorenz sequence used in the previous sections. The estimated $\widehat{\nu}$ shown in the map corresponds to the average over $M = 100$ evaluations relying on Eq. (3.15). Embedding pairs that were ruled out by the criteria described in the previous sections are not included in the estimation. Figure 3.7(b) shows the histogram of embedding pairs corresponding to a given $\widehat{\nu}$ (the bin width is 0.2). A clear peak is present in the histogram at $\widehat{\nu} = 2.0 \pm 0.2$, in agreement with the accepted value of correlation dimension for the Lorenz system, namely 2.05 [14].

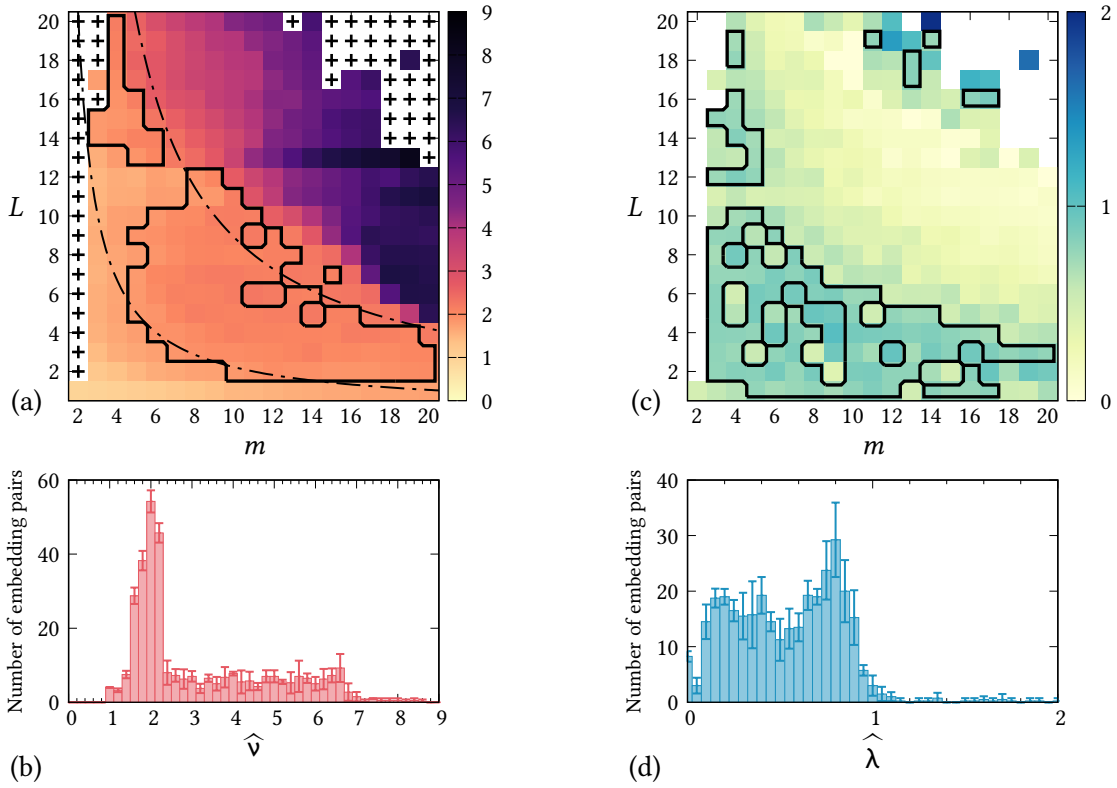


FIGURE 3.7 – *(a) Map of the estimated correlation dimension $\widehat{\nu}$ computed as the average of $M = 100$ evaluations of Eq. (3.15). Black "+" signs have the same meaning as in Fig. 3.6. The black, solid line bounds the plateau corresponding to embedding pairs having $\widehat{\nu} = 2.0 \pm 0.2$ making up the peak in the histogram shown in (b). The black, dash-dotted lines correspond to the hyperbolae that approximately enclose the plateau, described by $L = 20/(m-1)$ and $L = 80/(m-1)$. (b) Histogram of the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}$ falls within 0.2-wide bins. Bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the map in (a). (c) Map of the estimated MLE $\widehat{\lambda}$ evaluated by means of the divergence rate method on those embedding points for which $\mathcal{H}_{0;CFD}(\nu)$ holds. The black, solid line bounds the plateau corresponding to embedding pairs having $\widehat{\lambda} = 0.8 \pm 0.1$ making up the peak in the histogram shown in (d). (d) Histogram of the number of embedding pairs whose estimated MLE $\widehat{\lambda}$ falls within 0.05-wide bins. Bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the map in (c).*

The embedding pairs corresponding to the histogram peak make up a *plateau* region highlighted by means of a black contour within the map. This plateau is in agreement with the *uniformity* argument introduced above: its shape is indeed approximately bounded by two hyperbolae, as displayed by means of dash-dotted lines in Fig. 3.7(a). These hyperbolae correspond to $\tau_R/T \approx 20$ and $\tau_I/T \approx 80$, i.e. $\tau_R \approx 0.6$ and $\tau_I \approx 2.4$. The sufficient separation between the hyperbolae, as well as their position towards the middle of the map, suggest that the sampling period was correctly chosen. A too small sampling period (oversampling) would push the hyperbolae towards the top-right corner of the map. On the contrary, a

too large sampling period (undersampling) would compress the hyperbolae towards the axis and shrink the plateau. An example of how maps can provide information on the adequacy of the sampling period is discussed in Chapter 4, Sec. 4.1.

Figure 3.7 also shows another evaluation carried out, for the sake of comparison, on the embedding lattice. The maximum Lyapunov exponent (MLE) $\widehat{\lambda}$ is estimated—by means of the divergence rate method (see Sec. 2.5)—over all the embedding pairs (m, L) for which $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ holds, i.e. the same pairs over which $\widehat{\nu}$ was estimated. The $\widehat{\lambda}$ map and the corresponding histogram are shown in Fig. 3.7(c) and Fig. 3.7(d), respectively. The $\widehat{\lambda}$ histogram also exhibits a peak at around 0.8 ± 0.1, which is in agreement with the value 0.9057(8) obtained by means of the standard method (see Sec. 2.5). The set of embedding pairs corresponding to the histogram peak are highlighted by means of a black contour in the map of Fig. 3.7(c). The $\widehat{\lambda}$ plateau, despite sharing roughly half of the embedding points with the $\widehat{\nu}$ plateau, is less regular than the latter, thus reflecting the known variability of MLE estimated through the divergence rate method [133]. Comparing the two plateaus suggests that $\widehat{\nu}$ makes up a more reliable estimator for the selection of suitable embedding pairs.

3.3 Correlation dimension estimator in the GWN gauge

Besides the statistical tests discussed in Sec. 3.2.1 and Sec. 3.2.2, correlation bridges in the GWN gauge also allow to conveniently estimate the correlation dimension ν of the system generating the sequence under investigation, provided that $\mathcal{H}_{0;\mathrm{CFD}}(\nu)$ holds. Given an input sequence and an embedding pair (m, L), if the corresponding correlation bridge $\widehat{B}_{m,L;N}(\varrho)$ exhibits a diverging behaviour at the origin, then we can expect $\widehat{B}_{m,L;N}(\varrho)$ to grow up to a maximum and then decrease. The amplitude and position of the first maximum of $\widehat{B}_{m,L;N}(\varrho)$ provide an estimate $\widehat{\nu}$ of the correlation dimension ν. The estimator, which is derived in this Section, requires that the power-law scaling of $C_{m,L}(r)$ survives for sufficiently large values of ϱ.

Let us consider the asymptotic correlation bridge $B_{m,L}(\varrho)$ and its first derivative given by Eq. (3.12). Let ϱ_0 be the smallest value of ϱ for which this derivative vanishes, and r_0 the corresponding "original" distance. The inverse function theorem implies that, if ϱ_0 yields a vanishing derivative in Eq. (3.12), then r_0 satisfies the equation

$$\frac{\mathrm{d}\varrho(r)}{\mathrm{d}r} = \frac{\mathrm{d}C_{m,L}(r)}{\mathrm{d}r} \,. \tag{3.14}$$

We now consider the ratio between the first maximum of $B_{m,L}(\varrho)$ and ϱ_0:

$$\frac{B_{m,L}(\varrho_0)}{\varrho_0} = \frac{C_{m,L}(r_0)}{\varrho_0} - 1 \,.$$

Equation (3.11) provides a way to express $C_{m,L}(r_0)$:

$$C_{m,L}(r_0) = \frac{r_0}{\nu}\frac{\mathrm{d}C_{m,L}(r_0)}{\mathrm{d}r}\,.$$

By using this last expression, as well as Eq. (3.14), the ratio between the first maximum of $B_{m,L}(\varrho)$ and ϱ_0 can be written as

$$\frac{B_{m,L}(\varrho_0)}{\varrho_0} = \frac{1}{\nu}\frac{r_0}{\varrho_0}\frac{\mathrm{d}\varrho(r_0)}{\mathrm{d}r} - 1\,,$$

which is recast—by defining $x \equiv (mr^2)/(4\sigma^2)$, $x_0 \equiv (mr_0^2)/(4\sigma^2)$—into

$$\frac{B_{m,L}(\varrho_0)}{\varrho_0} = \frac{2}{\nu}\frac{x_0}{\varrho_0}\frac{\mathrm{d}\varrho(x_0)}{\mathrm{d}x} - 1\,.$$

The GWN gauge transformation for $m > 1$, namely $\varrho = \gamma(\frac{m}{2}, x)/\Gamma\left(\frac{m}{2}\right)$, as well as the corresponding derivative $\frac{\mathrm{d}\gamma(s,x)}{\mathrm{d}x} = x^{s-1}e^{-x}$, are inserted in the last expression, leading to

$$\frac{B_{m,L}(\varrho_0)}{\varrho_0} = \frac{2}{\nu\varrho_0}\frac{x_0^{m/2}e^{-x_0}}{\Gamma\left(\frac{m}{2}\right)} - 1\,.$$

Finally, the last expression can be inverted to get ν:

$$\nu = \frac{2}{B_{m,L}(\varrho_0) + \varrho_0}\frac{x_0^{m/2}e^{-x_0}}{\Gamma\left(\frac{m}{2}\right)}\,. \tag{3.15}$$

Equation (3.15) is obtained in the asymptotic limit $N \to \infty$. By replacing the asymptotic parameters x_0, ϱ_0, $B_{m,L}(\varrho_0)$ with the corresponding sample values, a sample estimate $\widehat{\nu}$ of the correlation dimension is provided. This estimator $\widehat{\nu}$ is exploited to build embedding lattice maps as discussed in Sec. 3.2.3.

4. Testing the new approach

This chapter is complementary to the previous one. In Chapter 3, a method for the identification of suitable embedding choices (m, L) to embed an input sequence was thoroughly described in its theoretical aspects. This method, which, for the sake of brevity, is henceforth referred to as the "embedding lattice method", is here applied to different synthetic and experimental benchmark sequences. The goal of these tests is to provide further insight on the kind of information delivered by the method, as well as on its limitations. Throughout this chapter, the results of the application of the embedding lattice method are displayed by means of diagrams of the kind shown in Fig. 3.7(a), which are henceforth referred to as *embedding lattice maps*. These maps encapsulate the information provided by two statistical tests, namely the compatibility of the input sequence—as a function of the embedding points (m, L)—with the null hypotheses of a GWN source ($\mathcal{H}_{0;\mathrm{GWN}}$) and of a chaotic, finite dimensional source ($\mathcal{H}_{0;\mathrm{CFD}}$), as well as the estimated $\widehat{\nu}$ value provided by the estimator in Eq. (3.15). Besides the color-scale representation within embedding lattice maps, values of $\widehat{\nu}$ are also collected within histograms such as the one shown in Fig. 3.7(b). These $\widehat{\nu}$ histograms allow a prompt identification of $\widehat{\nu}$ plateaus.

Section 4.1 presents the application of the embedding lattice method to a surrogate Lorenz sequence, which has the same Fourier components as the Lorenz sequence used to generate Fig. 3.7 despite being stochastic: in this way, the dependence of the method outcomes on the spectrum of the input sequence is discussed. Afterwards, the information provided by the embedding lattice method concerning the adequacy of the sampling period used to record an input sequence is discussed by considering data generated by a Ueda oscillator. Furthermore, a sequence generated by a Rössler system is analyzed: the corresponding results highlight the difficulties in dealing with systems that are almost periodic. In relation to this issue, a model to predict the estimated $\widehat{\nu}$ in the case of sequences having a high degree of (though not perfect) periodicity is provided in Sec. 4.2. As a last example, in Sec. 4.1 the embedding lattice method is also applied to experimental data corresponding to the recorded intensity of an NH_3 laser. Finally, Sec. 4.3 provides a qualitative analysis of the effects of observational noise on the embedding lattice maps.

4.1 Application to synthetic and experimental sequences

As a first test, it is interesting to investigate what is the role of the Fourier spectrum in determining the shape of embedding lattice maps. In other words, the aim of the following first example is to understand how much extra information do the map convey with respect to spectral analysis. To this purpose, we consider the same Lorenz sequence of Chapter 3 and we compute a *surrogate* [179] sequence out of it. Surrogate generation is a widespread tool to build sequences that preserve some properties of a given original sequence, but are otherwise stochastic [180]. This possibility is quite useful when null hypothesis distributions are unknown and significance has to be estimated by relying only on the data themelves (for a recent review, see Ref. [181]). In this example, a surrogate sequence is generated according to a "phase randomization" algorithm [180] that can be summarized as follows (see diagram in Fig. 4.1). First, the discrete Fourier transform $\{\tilde{y}_k\}$ is computed—by means of a Fast

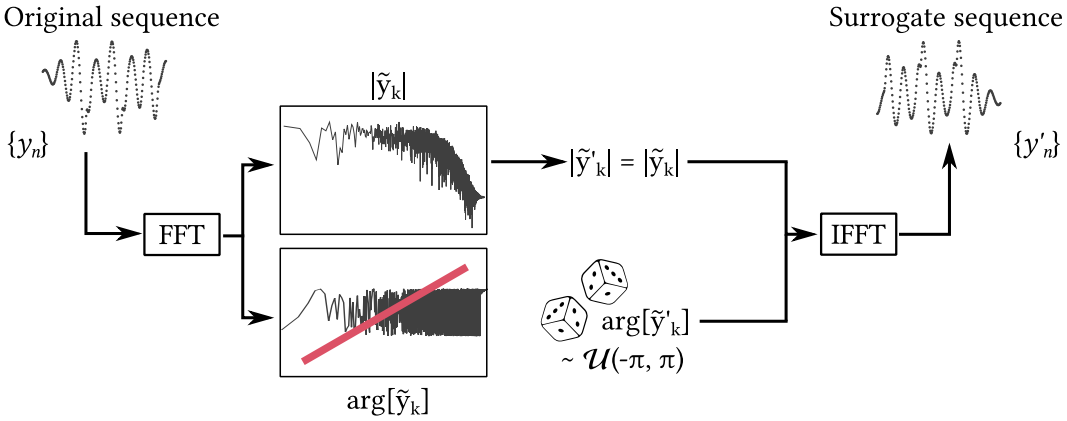


FIGURE 4.1 – *Diagram of the "phase randomization" surrogate generation algorithm.*

Fourier Transform (FFT) algorithm—out of the "original" input sequence $\{y_n\}$ of length ℓ. Thereupon, the corresponding phases $\arg[\tilde{y}_k]$ are replaced by i.i.d. random variates distributed uniformly within $[-\pi, \pi)$, while the amplitudes $|\tilde{y}_k|$ are unchanged. This procedure yields new Fourier components $\{\tilde{y}'_k\}$ having the same amplitude but randomized phases with respect to the original ones. Finally, the inverse Fourier transform is evaluated out of $\{\tilde{y}'_k\}$. The resulting "surrogate" sequence $\{y'_n\}$ thus shares the same power spectral density and—by virtue of the Wiener-Khinchin theorem—the same autocorrelation as the original sequence $\{y_n\}$. Nevertheless, because its Fourier phases are randomly assigned, the surrogate sequence is stochastic.

A comparison between the original Lorenz sequence and a phase-randomized surrogate allows testing whether the method described in Chapter 3 can distinguish a chaotic source from a stochastic one having the *same* spectrum. The embedding lattice map—obtained by means of the embedding lattice method—for a surrogate Lorenz sequence is shown in Fig. 4.2. The map is definitely different from

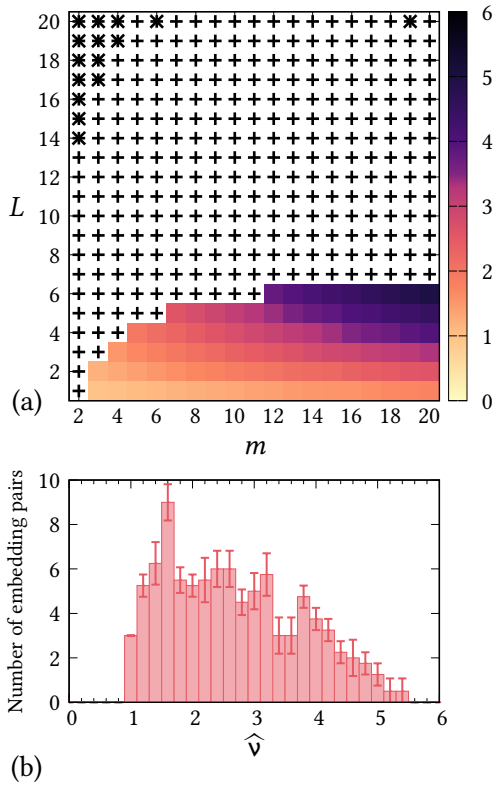


FIGURE 4.2 – *Results of the analysis by means of the embedding lattice method applied to a surrogate Lorenz sequence. (a) Map over the embedding lattice of the estimated $\widehat{\nu}$. Black crosses (×) correspond to embedding points compatible with the null hypothesis $\mathcal{H}_{0;GWN}$. Black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. (b) Histogram of the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}$ falls within 0.2-wide bins. Bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the map in (a).*

the one corresponding to the original Lorenz sequence and shown in Fig. 3.7. Most of the embedding lattice points ($L \gtrsim 5$) are not compatible with the $\mathcal{H}_{0;\mathrm{CFD}}$ null hypothesis, i.e. the sequence embedded on those (m, L) is not statistically compatible with a finite-dimensional system. In other words, the embedding lattice method correctly identifies (for $L \gtrsim 5$) the surrogate Lorenz sequence as stochastic, despite its power spectral density and its autocorrelation are *identical* to the same functions estimated out of the original Lorenz sequence. This result highlights the capability of the embedding lattice method to discriminate deterministic and stochastic sources beyond their similarity in autocorrelation (or power spectrum), a crucial issue introduced in Chapter 1 through the comparison illustrated in Fig. 1.5.

The second example considered here relies on sequences generated by a Ueda oscillator. The aim of this example is to show how the embedding lattice method can provide information on the adequacy of the sampling period used to record a sequence. To this purpose, a "Ueda sequence" is synthesized by sampling the y-coordinate of a Ueda oscillator[1]. A first sequence is tentatively sampled by setting the integration time step—and thus the sampling period—to $\delta t = 0.5$. The embedding lattice map provided by the embedding lattice method for this first sequence is shown in Fig. 4.3(a), while Fig. 4.3(b) shows the corresponding histogram of $\widehat{\nu}$ values. The $\widehat{\nu}$ histogram exhibits a growing profile that leads to a sharp peak at large dimension ($\widehat{\nu} \simeq 5$). This peak, however, corresponds to a plateau that is not well described

[1]See Chapter 1, Eq. (1.3) for the system's equations. Here the system parameters are set to $A = 7.5$, $k = 0.05$.

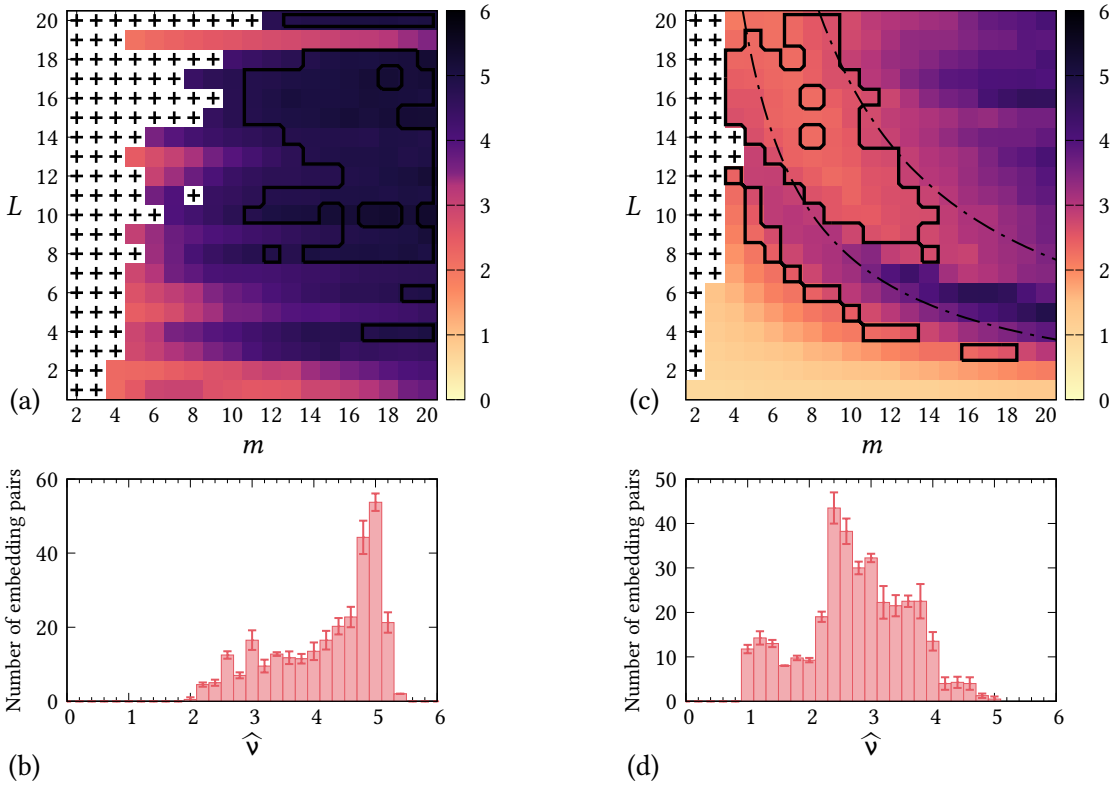


FIGURE 4.3 – *Results of the analysis by means of the embedding lattice method applied to a Ueda sequence sampled with* $\delta t = 0.5$ *(a,b) and with* $\delta t = 0.05$ *(c,d). (a) Embedding lattice map of estimated* $\widehat{\nu}$*. The black solid line identifies the embedding points corresponding to the peak at* $\widehat{\nu} = 5.0 \pm 0.1$ *within the histogram in (b). (b) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.2*-wide bins. (c) Embedding lattice map of estimated* $\widehat{\nu}$*. The black solid line identifies the embedding points corresponding to the peak at* $\widehat{\nu} = 2.5 \pm 0.2$ *within the histogram in (d). The black, dash-dotted lines correspond to the hyperbolae approximately bounding the plateau and defined by* $L = 70/(m-1)$ *and* $L = 150/(m-1)$*. (d) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.2*-wide bins. In both maps (a,c), black "+" marks denote embedding points that are incompatible with the null hypothesis* $\mathcal{H}_{0;CFD}$*. In both histograms (b,d), bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the maps in (a,c).*

by two bounding hyperbolae. The growing histogram shape is generated by two possible mechanisms: on the one hand, the system might indeed have a large ($\gtrsim 5$) dimensionality, but the currently used embedding lattice is not wide enough to correctly resolve the plateau (i.e. the sequence is oversampled). On the other hand, the opposite situation of an undersampled sequence might hold. In this case, the irrelevance time is already overcome for small embedding windows $(m-1)L \lesssim 10$: over most of the embedding lattice, the sequence then appears as noise-like and thus with a high dimensionality.

The first scenario is tackled by either widening the embedding lattice (a computationally demanding approach) or by downsampling the sequence. Following the latter strategy leads to an embedding lattice

map that is essentially the one of filtered noise. The first scenario is thus rejected.
The possibility of an undersampled sequence is instead dealt with by analyzing a Ueda sequence whose sampling period is set to $\delta t = 0.05$, i.e. one tenth of the previously used sampling period. This second sequence, sampled 10 times faster, yields the embedding lattice map of Fig. 4.3(c) and the corresponding $\widehat{\nu}$ histogram of Fig. 4.3(d). In this case, the peak at $\widehat{\nu} = 2.5 \pm 0.2$ corresponds to a $\widehat{\nu}$ plateau within the map of Fig. 4.3(c) that is bounded by two hyperbolae and thus complies with the irrelevance-redundance argument discussed in Sec. 3.2.3. The two hyperbolae shown in Fig. 4.3(c) correspond to $\tau_R \approx 3.5$ and $\tau_I \approx 7.5$. While this τ_I implies, for a sampling period equal to 0.05, an embedding window $(m-1)L \approx 150$, the undersampled sequence generated with $\delta t = 0.5$ would have an irrelevance embedding window of ≈ 15, thus compressing the plateau towards the bottom-left corner of the map, as in Fig. 4.3(a). Finally, these results are consistent with the known properties of the Ueda oscillator: the correlation dimension $\widehat{\nu} = 2.5 \pm 0.2$ corresponding to the peak in Fig. 4.3(d) is in agreement with the value reported in the literature, i.e. $\nu = 2.68 \pm 0.13$ [14].

At this point, one might wonder whether increasing the sampling frequency could yield a (spurious) plateau also in the case of filtered noise. Two filtered noise sequences are considered, each synthesized by generating $\ell = 10^5$ standard normal random variates and by piping the resulting white noise sequences into a digital first-order low-pass filter with cut-off frequency $f_{3\text{dB}}$. The first sequence is assumed[2] to be sampled with sampling frequency $T^{-1} = 20 f_{3\text{dB}}$ and the second one with sampling frequency $10T^{-1} = 200 f_{3\text{dB}}$. The embedding lattice maps obtained by means of the embedding lattice method and the related histograms are shown in Fig. 4.4(a,b) and Fig. 4.4(c,d) for the T^{-1} and the $10T^{-1}$ sequences, respectively. An increased sampling frequency yields an expansion of the region of the embedding lattice containing embedding pairs that are not incompatible with $\mathcal{H}_{0;\text{CFD}}$. Accordingly, the histogram of $\widehat{\nu}$ gets more populated, but no significant peak appears. In conclusion, despite the increased sampling frequency, maps of filtered noise do not provide any plateau, contrary to what happens in Fig. 4.3(c) for the Ueda oscillator.

In order to further test the reliability of the embedding lattice method, we now consider a "standard" Rössler sequence, namely a sequence synthesized by sampling the x-coordinate of a Rössler system[3] integrated by setting the system's parameters to $a = 0.2$, $b = 0.2$, $c = 5.7$ and with integration time step $\delta t = 0.125$. Moreover, a corresponding (phase-randomized) surrogate sequence is also generated. The resulting embedding lattice maps, as well as the related $\widehat{\nu}$ histograms, are displayed in Fig. 4.5. Despite the existence of sharp peaks at $\widehat{\nu} = 1.6 \pm 0.1$ and $\widehat{\nu} = 1.9 \pm 0.1$ within the histograms of the original and surrogate sequence, respectively, the corresponding plateaus within the maps are not bounded by hy-

[2] What really matters, here, is only the ratio between the sampling frequency and the filter cut-off frequency, and not their absolute values.

[3] See Chapter 1, Eq. (1.4) for the system's equations.

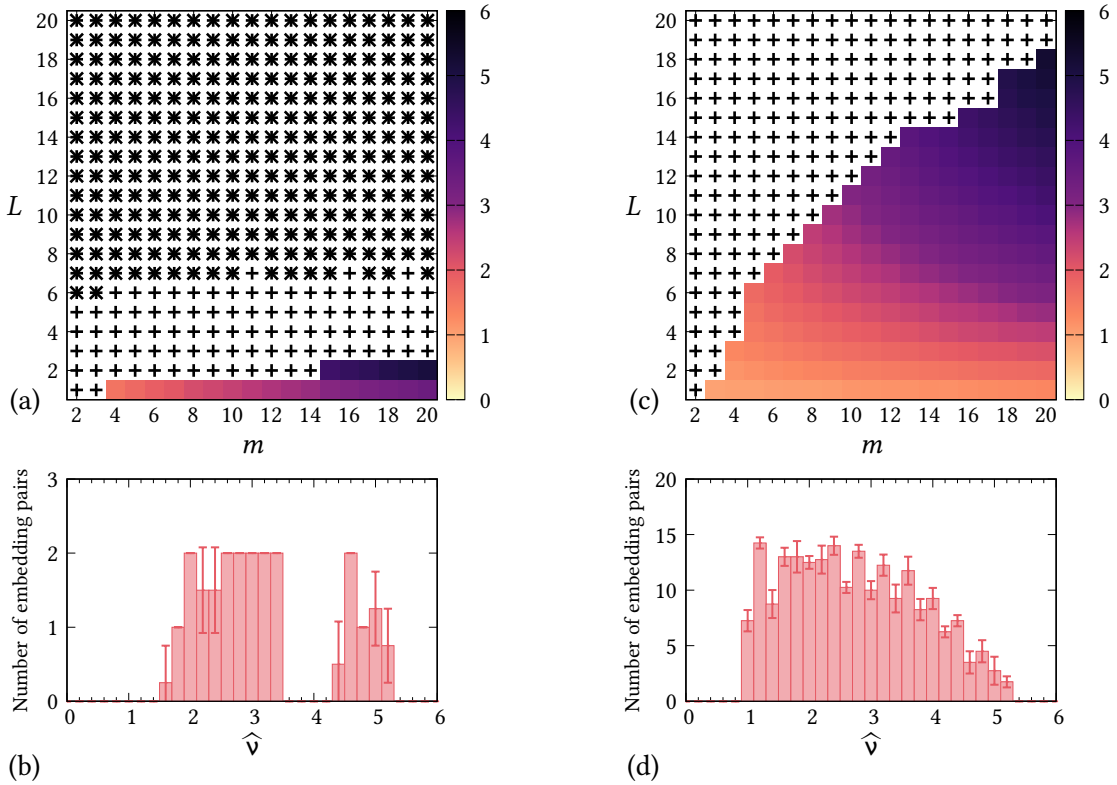


FIGURE 4.4 – *Results of the analysis by means of the embedding lattice method applied to a filtered noise sequence sampled with sampling frequency* T^{-1} *(a,b) and* $10\,T^{-1}$ *(c,d), respectively. (a,c) Embedding lattice maps of estimated* $\widehat{\nu}$*. In (a), black crosses (×) identify embedding points for which compatibility with the GWN null hypothesis* $\mathcal{H}_{0;GWN}$ *holds. In both maps, black "+" marks denote embedding points that are incompatible with the null hypothesis* $\mathcal{H}_{0;CFD}$*. In both cases, no plateau region can be identified. (b,d) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.2*-wide bins. In both histograms, bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the maps in (a,c).*

perbolae. On the contrary, the plateau shapes are "rectangular", exhibiting an almost constant $\widehat{\nu}$ values independently of L. Consequently, the maps do not provide evidence of an underlying chaotic, finite-dimensional system. While this result is expected for the surrogate sequence, in the case of the original one the underlying system is known to be chaotic with correlation dimension $\widehat{\nu} = 1.99 \pm 0.08$ [14]: the embedding lattice method fails in correctly characterizing the source. This outcome reflects the known difficulties in searching optimal embeddings for the "standard" Rössler attractor [132, 133], as well as concerning the determination of its correlation dimension with conventional approaches [14]. These difficulties arise from the high degree of periodicity and long autocorrelation time of the trajectories generated by the Rössler system. Because of these features, the "standard" Rössler sequence can be modeled as a perturbed periodic signal, for which an approximated value of the estimated $\widehat{\nu}$ is

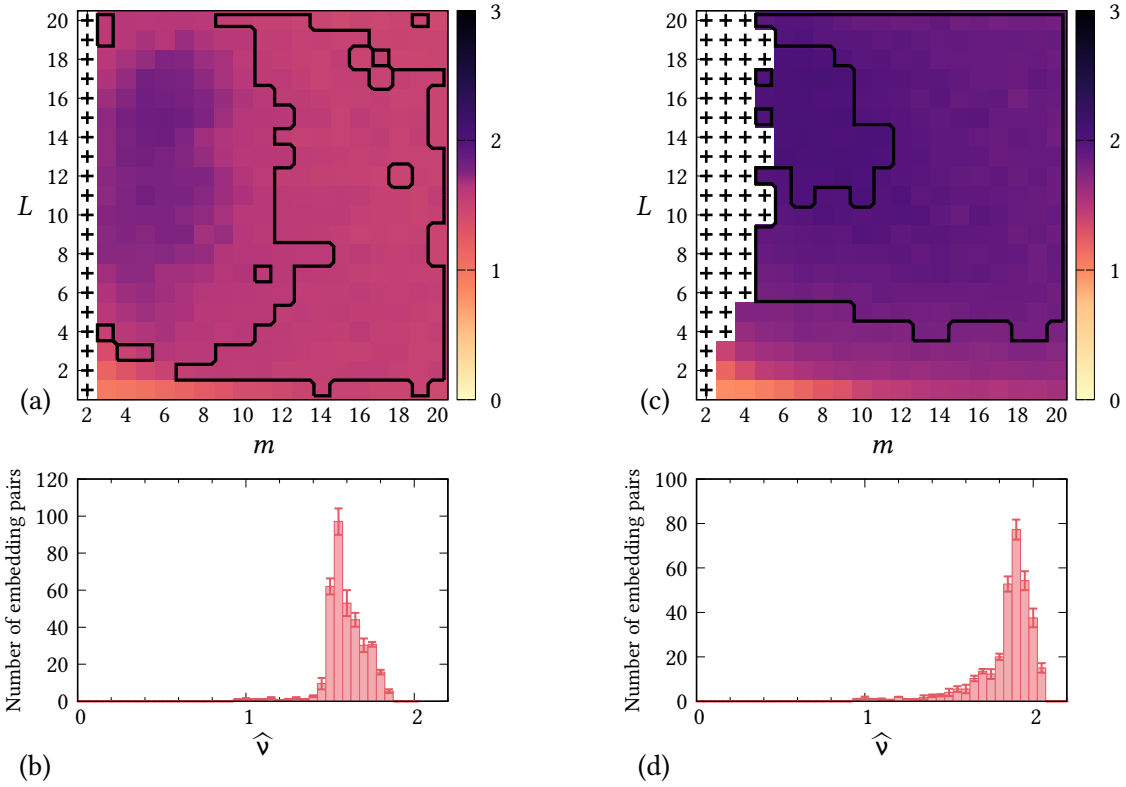


FIGURE 4.5 – *Results of the analysis by means of the embedding lattice method applied to a Rössler "standard" sequence (a,b) and to a corresponding surrogate sequence (c,d). (a) Embedding lattice map of estimated* $\widehat{\nu}$*. The black solid line identifies the embedding points corresponding to the peak at* $\widehat{\nu} = 1.55 \pm 0.05$ *within the histogram in (b). (b) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.05*-wide bins. (c) Embedding lattice map of estimated* $\widehat{\nu}$*. The black solid line identifies the embedding points corresponding to the peak at* $\widehat{\nu} = 1.9 \pm 0.1$ *within the histogram in (d). (d) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.05*-wide bins. In both maps (a,c), black "+" marks denote embedding points that are incompatible with the null hypothesis* $\mathcal{H}_{0;CFD}$*. In both histograms (b,d), bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the maps in (a,c).*

computed in Sec. 4.2 below. The discussion reported in Sec. 4.2 also provides an explanation for the rectangular plateau of Fig. 4.5.

The long autocorrelation time of the Rössler system can be reduced by tuning the parameters so as to yield the so-called "funnel" attractor: $a = 0.343$, $b = 1.82$, $c = 9.75$, and the integration time step is set to $\delta t = 0.05$. A comparison between the "standard" and the "funnel" attractors is shown in Chapter 1, Fig. 1.8. The funnel configuration provides a smoother spectrum and thus a shorter autocorrelation time. The results of the embedding lattice method on the funnel Rössler sequence, again evaluated both for the original sequence and for a surrogate one, are shown in Fig. 4.6. As for the standard

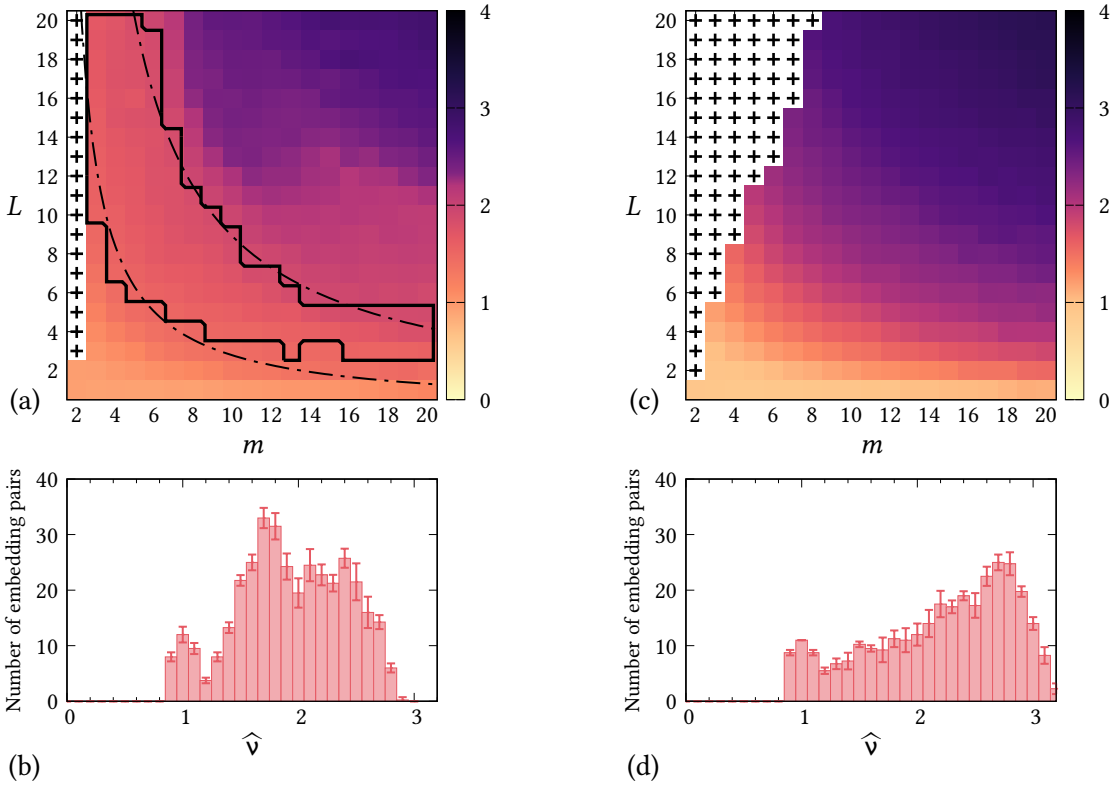


FIGURE 4.6 – *Results of the analysis by means of the embedding lattice method applied to a Rössler "funnel" sequence (a,b) and to a corresponding surrogate sequence (c,d). (a) Embedding lattice map of estimated* $\widehat{\nu}$*. The black solid line identifies the embedding points corresponding to the peak at* $\widehat{\nu} = 1.7 \pm 0.2$ *within the histogram in (b). The black, dash-dotted lines correspond to the hyperbolae approximately bounding the plateau and defined by* $L = 25/(m-1)$ *and* $L = 80/(m-1)$*. (b) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.1*-wide bins. (c) Embedding lattice map of estimated* $\widehat{\nu}$*. No plateau region can be identified. (d) Histogram of the number of embedding pairs whose estimated correlation dimension* $\widehat{\nu}$ *falls within* 0.1*-wide bins. In both maps (a,c), black "+" marks denote embedding points that are incompatible with the null hypothesis* $\mathcal{H}_{0;CFD}$*. In both histograms (b,d), bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the maps in (a,c).*

Rössler system, the results corresponding to the surrogate sequence (Fig. 4.6(c,d)) yield no evidence of an underlying chaotic system. On the contrary, the embedding lattice map is qualitatively similar to the one for filtered noise shown in Fig. 4.4(c): indeed, the surrogate sequence is stochastic, albeit with a finite autocorrelation time. The original sequence, on the other hand, provides an embedding lattice map that—as expected for a chaotic source—exhibits a plateau corresponding to a peak in the $\widehat{\nu}$ histogram centered at $\widehat{\nu} = 1.7 \pm 0.2$. The plateau turns out to be approximately bounded by two hyperbolae corresponding to a redundance time $\tau_R \approx 1$ and an irrelevance time $\tau_I \approx 4$. The embedding lattice method thus correctly identifies the sequences as generated by a chaotic system.

As a last example, the embedding lattice method is applied to an experimental benchmark sequence, namely the intensity sequence recorded out of a 81.5 µm NH_3 laser. The sequence is briefly introduced in Chapter 1, and an excerpt of the sequence is shown in Fig. 1.9. The embedding lattice map resulting from the application of the embedding lattice method is shown in Fig. 4.7(a); the histogram of $\widehat{\nu}$ is displayed in Fig. 4.7(b). A peak at $\widehat{\nu} = 2.0 \pm 0.2$ within the histogram of Fig. 4.7(b)—corresponding to

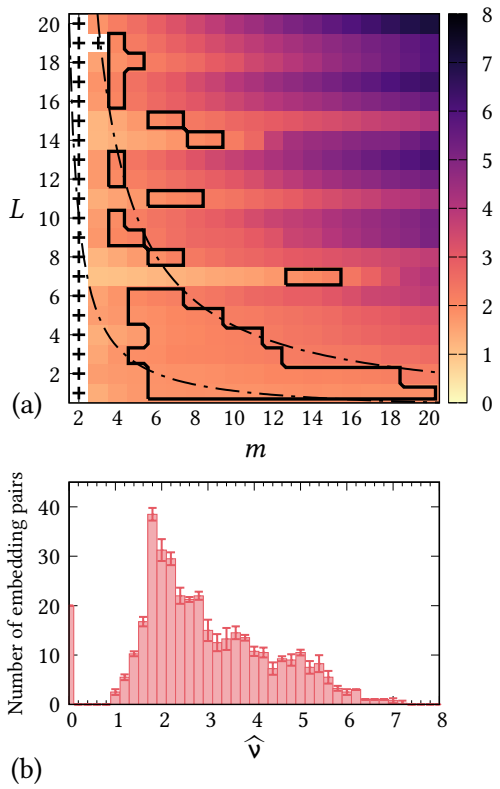


FIGURE 4.7 – *Results of the analysis by means of the embedding lattice method applied to an experimentally recorded laser intensity sequence. (a) Embedding lattice map of estimated $\widehat{\nu}$. The black solid line identifies the embedding points corresponding to the peak at $\widehat{\nu} = 2.0 \pm 0.2$ within the histogram in (b). The black, dash-dotted lines correspond to the hyperbolae approximately bounding the plateau and defined by $L = 10/(m-1)$ and $L = 40/(m-1)$. Black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. (b) Histogram of the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}$ falls within 0.2-wide bins. Bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the map in (a).*

a plateau approximately bounded by two hyperbolae—hints at an underlying chaotic source. Indeed, the $\widehat{\nu}$ estimate of the correlation dimension is in agreement with the value provided in the literature of $\nu \simeq 2.0 \div 2.3$ [40]. The two hyperbolae reported in the map of Fig. 4.7(a) correspond to estimated redundance and irrelevance times of $\tau_R/T \approx 10 = 0.4\,\mu s$ and $\tau_I/T \approx 40 = 1.6\,\mu s$, respectively.

To further check these results, the MLE of the laser intensity sequence was estimated on the 72 embedding pairs (m, L) that belong to the plateau. The divergence rate method (see Sec. 2.5) provides a nonzero value of the MLE for 48 of these embedding pairs. The average over these 48 evaluations provides an estimated MLE $\widehat{\lambda} = (0.025 \pm 0.007)\ T^{-1} = (0.63 \pm 0.18)$ MHz; this estimate is in agreement with the value obtained by Kantz & Schreiber [5].

4.2 Estimated correlation dimension in the case of noisy, periodic signals

The aim of this Section is to provide insight on the results shown in Fig. 4.5 for the "standard" Rössler sequence, which is characterized, to a first approximation, by an almost-periodic behaviour. To this purpose, the value taken on by the correlation dimension estimator of Eq. (3.15) in the case of periodic signals contaminated by observational noise is investigated. The "periodic signal + noise" model considered here is a proxy for a quasiperiodic signal: periodicity fluctuations are generated, in this model, by amplitude fluctuations. The sequence $\{y_n\}$ considered here is thus

$$y_n = \sin(\phi_0\, n) + \frac{\varepsilon_n}{\sqrt{2}}\,, \tag{4.1}$$

where the noise terms ε_n contaminating the signal are i.i.d. normal random variates with vanishing mean and standard deviation equal to ε. The factor $\sqrt{2}$ is introduced so as to provide a signal-to-noise ratio equal to ε^{-1}.

Distances between embedding vectors built out of the sequence in Eq. (4.1) embedded with parameters m, L are given, according to Eq. (3.1), by

$$d^2_{i,j} = \frac{1}{m}\sum_{k=0}^{m-1}\left\{2\cos\left[\phi_0\left(\frac{i+j}{2}+kL\right)\right]\,\sin\left[\phi_0\frac{i-j}{2}\right]+\frac{\varepsilon_{i+kL}-\varepsilon_{j+kL}}{\sqrt{2}}\right\}^2 .$$

By expanding the squared parenthesis $\{\ldots\}^2$, the expression for $d^2_{i,j}$ is recast into a sum of three terms:

$$\begin{aligned} d^2_{i,j} \;=\; & \frac{4}{m}\sin^2\left[\phi_0\frac{i-j}{2}\right]\sum_{k=0}^{m-1}\cos^2\left[\phi_0\left(\frac{i+j}{2}+kL\right)\right]+ \\ & +\frac{2\sqrt{2}}{m}\sin\left[\phi_0\frac{i-j}{2}\right]\sum_{k=0}^{m-1}\cos\left[\phi_0\left(\frac{i+j}{2}+kL\right)\right]\left(\varepsilon_{i+kL}-\varepsilon_{j+kL}\right)+ \\ & +\frac{1}{2m}\sum_{k=0}^{m-1}\left(\varepsilon_{i+kL}-\varepsilon_{j+kL}\right)^2 . \end{aligned}$$

Let us now consider the three terms separately. The first term is henceforth referred to as D_1. Because the indexes i, j are uniformly distributed, the sum of squared cosine elements is, if m is large, approximately equal to $m/2$. *Small scale* contributions to the correlation integral, then, are given by the $\sin^2(\cdot)$ factor when its argument is close to an integer multiple of π. In this approximation, $D_1 \approx 2\sin^2(\phi)$, and thus $\phi \approx \arcsin\left(\sqrt{D_1/2}\right)$: by assuming again a uniform distribution of ϕ, the cumulative distribution of

D_1 is evaluated through a change of variable and is given by

$$C_{D_1}(D_1) = \frac{2}{\pi} \arcsin\left(\sqrt{\frac{D_1}{2}}\right).$$

The factor $2/\pi$ ensures normalization (D_1 is bounded within $[0, 2]$).
The noise elements ε_n are supposed to be normally distributed according to $\varepsilon_n \sim \mathcal{N}(0, \varepsilon^2)$. Consequently, because the variances of $\sin(\cdot)$ and $\cos(\cdot)$ factors are equal to $1/2$, the second term making up $d_{i,j}^2$, which is referred to as D_2, is approximately normally distributed according to

$$D_2 \sim \mathcal{N}\left(0, \frac{4\varepsilon^2}{m}\right).$$

Finally, the third term D_3 is a χ^2 random variable with m degrees of freedom:

$$D_3 \sim \frac{\varepsilon^2}{m} \chi_m^2 .$$

By taking the limit $m \gg 1$, the term D_2 becomes negligible due to the $1/m$ factor in its variance. In the same limit of large m, the expected value of χ_m^2/m is approximately unitary, the corresponding variance vanishes, and thus D_3 tends to a constant equal to ε^2. Therefore, if $m \gg 1$, the squared distances $d_{i,j}^2$ can be written as

$$d_{i,j}^2 \simeq D_1 + \varepsilon^2 .$$

By applying again a change of variable, the corresponding cumulative distribution is finally given by

$$C(r) = \frac{2}{\pi} \arcsin\left(\sqrt{\frac{r^2 - \varepsilon^2}{2}}\right).$$

According to the argument presented in Sec. 3.3, the correlation dimension ν of the sequence y_n is estimated by

$$\begin{aligned} \nu &= \left.\frac{\mathrm{d}\log[C(r)]}{\mathrm{d}\log(r)}\right|_{r=r_0} = \left.\frac{\mathrm{d}\log\left(\arcsin\sqrt{\frac{r^2-\varepsilon^2}{2}}\right)}{\mathrm{d}\log(r)}\right|_{r=r_0} = \\ &= \frac{r_0^2}{\sqrt{r_0^2 - \varepsilon^2}\sqrt{2 + \varepsilon^2 - r_0^2}\arcsin\sqrt{\frac{r_0^2-\varepsilon^2}{2}}} . \end{aligned} \tag{4.2}$$

Here, as in Sec. 3.3, r_0 is the value of r such that $\varrho_0 = \varrho(r_0)$ is the position of the first maximum of the

corresponding correlation bridge (see Eq. (3.14)). The value of r_0 can be estimated by considering the gauge transformation $\varrho(r)$ given by Eq. (3.4) and corresponding to the cumulative distribution of a χ^2_m random variable. In the $m \gg 1$ limit, by virtue of the central limit theorem, this distribution can be approximated as a normal distribution:

$$\varrho(r) = \frac{1}{\Gamma\left(\frac{m}{2}\right)} \gamma\left(\frac{m}{2}, \frac{mr^2}{4\sigma^2}\right) \approx \Phi\left(\frac{\frac{mr^2}{2\sigma^2} - m}{\sqrt{2m}}\right) = \Phi\left(\frac{r^2 - 2\sigma^2}{2\sigma^2\sqrt{2/m}}\right),$$

where the function $\Phi(z)$ is the (cumulative) standard normal distribution. Because the distribution is significantly different from zero only for $r \simeq \sqrt{2}\sigma$, the replacement $r^2 - 2\sigma^2 \approx 2\sqrt{2}\sigma(r - \sqrt{2}\sigma)$ is carried out, leading to

$$\varrho(r) \approx \Phi\left(\frac{r - \sqrt{2}\sigma}{\sigma/\sqrt{m}}\right).$$

In the $m \gg 1$ limit, the last expression converges to a Heaviside theta centered in $\sqrt{2}\sigma$. The corresponding derivative is a Dirac delta centered in $\sqrt{2}\sigma$. Consequently, r_0, namely the root of Eq. (3.14), is equal to $\sqrt{2}\sigma$ regardless of the underlying correlation integral $C_{m,L}(r)$.

The value of σ, in the case of the "periodic signal + noise" model of Eq. (4.1), is straightforwardly given by $\sigma^2 = (1 + \varepsilon^2)/2$. Therefore,

$$r_0 = \sqrt{1 + \varepsilon^2},$$

and Eq. (4.2) provides the corresponding correlation dimension:

$$\nu = \frac{4}{\pi}(1 + \varepsilon^2). \tag{4.3}$$

This last expression holds for $m \gg 1$. In this limit, it provides the height of a plateau within the embedding lattice that is expected to be "rectangular", i.e. corresponding to a constant $\widehat{\nu}$ independently of L.

Finally, it is worth remarking that, by virtue of the central limit theorem, the result of Eq. (4.3) can be extended to noise characterized by any distribution—not necessarily Gaussian—having vanishing mean and standard deviation equal to ε.

The validity of Eq. (4.3) is numerically tested by generating a synthetic sequence according to Eq. (4.1). In Fig. 4.8, evaluations of ν by means of the embedding lattice method applied on the synthetic sequence are shown as a function of the noise standard deviation ε. The analytical expression of Eq. (4.3) is also displayed: the numerically estimated values are in good agreement with the analytical expression, thus supporting the validity of the model. In addition, a fit with a quadratic function $\nu(\varepsilon) = (4/\pi)(a + b\varepsilon^2)$ is carried out and shown for comparison. The fit parameters turn out to be

$a = 0.985 \pm 0.006$ and $b = 1.02 \pm 0.03$, again in agreement with Eq. (4.3).

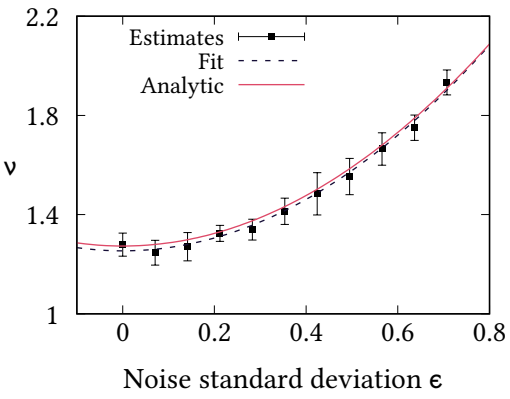


FIGURE 4.8 – *Estimation of ν out of a sinuoidal signal contaminated by Gaussian white noise as a function of the noise standard deviation ε. Black dots and the related errorbars are numerical estimates obtained by analyzing a synthetic sequence described by Eq. (4.1). Each data point is the result of the average over 11 evaluations of $\widehat{\nu}$ obtained, according to Eq. (3.15), by setting $m = 5000$ and $L = 5 \ldots 15$. Errorbars correspond to the standard deviation of the 11 evaluations. The red, solid line shows to the analytical expression of Eq. (4.3). The blue, dashed line corresponds to a fit as described in the main text.*

The observation of a rectangular (i.e. L-independent) plateau for $m \gg 1$ for the Rössler system, as shown in Fig. 4.5, can now be interpreted by means of the model discussed in this Section. According to this model, the rectangular plateau is typical of quasiperiodic systems, here simulated as periodic signals with added noise. Indeed, the Rössler sequence corresponding to the results of Fig. 4.5 exhibits a spectrum that contains a sharp line at frequency $f_0 = 0.02 T^{-1}$, where T is the sampling period, superimposed to a broader background. The periodic oscillation corresponding to f_0 is taken as the "signal", while the background is considered as a "noise" perturbation[4]. The FFT amplitude is shown in Fig. 4.9.

The "noise" standard deviation ε is estimated as follows. First, the squared FFT amplitudes falling

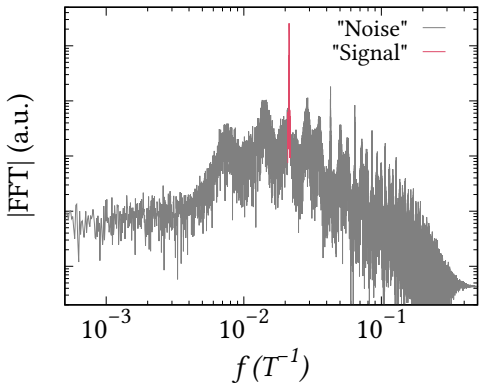


FIGURE 4.9 – *FFT amplitude of the "standard" Rössler sequence analyzed in Sec. 4.1. Frequencies are measured as multiples of the inverse sampling period T^{-1}. Following the almost-periodic model presented in this Section, the sharp spectral line at $f_0 \approx 0.02\ T^{-1}$ is assumed to be the periodic "signal", while the remaining broadband spectrum is considered as the "noise" perturbation.*

in the "noise" and "signal" categories, respectively, are summed up. Thereupon, ε is evaluated as the square root of the ratio between the two. The resulting ε turns out to be $\varepsilon \simeq 0.47$. According to the model discussed in this Section, i.e. following Eq. (4.3), this value of ε corresponds to an "effective"

[4]It is worth remarking that the sequence is purely deterministic; however, according to the almost-periodic model discussed in this Section, the spectrum of the Rössler sequence is interpreted as the sum of a periodic oscillation and a noise contribution.

correlation dimension $\nu \simeq 1.55$, which is in good agreement with the plateau value obtained for the standard Rössler sequence and shown in Fig. 4.5(a,b).

4.3 Effect of observational noise: a qualitative analysis

In this Section, the effect of observational noise on the embedding lattice maps obtained by means of the embedding lattice method is qualitatively investigated. As a reference, the Lorenz sequence analyzed in Chapter 3 is considered. The corresponding "noise-free" embedding lattice map is shown in Fig. 3.7(a). The sequence is contaminated with observational GWN having progressively higher power, and thus yielding progressively smaller signal-to-noise ratios (SNR). More precisely, the input sequences $\{s_n\}$ are given by

$$s_n = x_n + \varepsilon_n \,,$$

where x_n are the elements of a Lorenz sequence[5], while the noise elements ε_n are i.i.d. random variates distributed according to

$$\varepsilon_n \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \,.$$

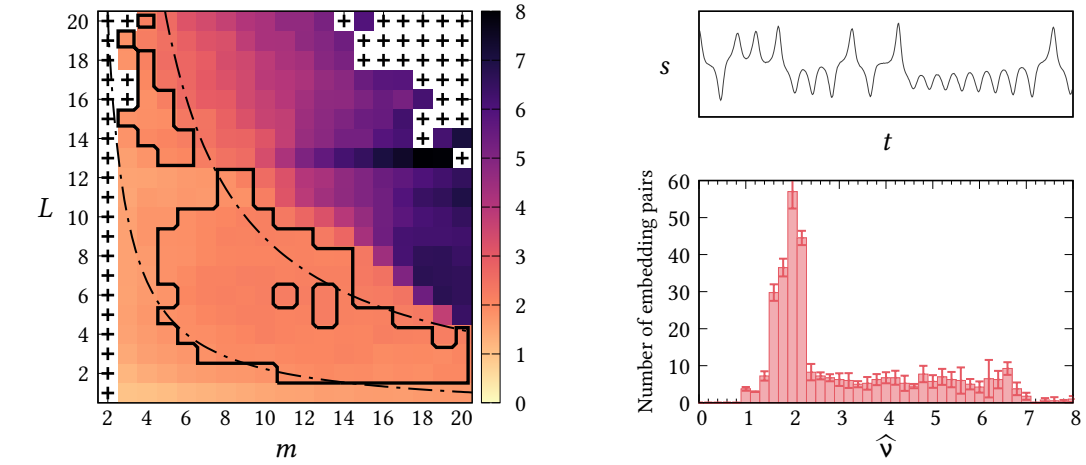
Given the standard deviation σ_x of the Lorenz sequence $\{x_n\}$, the SNR is then defined (in decibels) as

$$\mathrm{SNR} = 20 \, \log_{10}\left(\frac{\sigma_x}{\sigma_\varepsilon}\right) \, [\mathrm{dB}] \,.$$

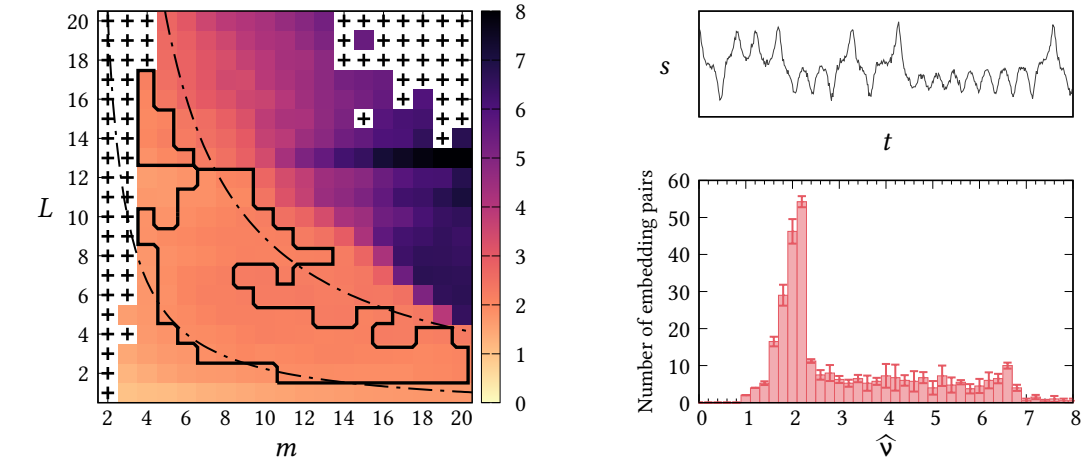
The sequences considered here have SNR equal to 40 dB, 20 dB, 15 dB, 6 dB, 0 dB, −6 dB. The corresponding embedding lattice maps, the corresponding $\widehat{\nu}$ histograms, as well as an excerpt of the $\{s_n\}$ sequences, are shown in Fig. 4.10.

FIGURE 4.10 – *(Figure is displayed in two dedicated pages below). Embedding lattice maps, $\widehat{\nu}$ histograms, and excerpts of the input $\{s_n\}$ sequences corresponding to a Lorenz sequence contaminated by observational GWN. Within maps, black crosses (×) correspond to embedding points compatible with the null hypothesis $\mathcal{H}_{0;GWN}$; black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. Histograms are built with bin width 0.2; bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the respective maps. (a) SNR = 40 dB. (b) SNR = 20 dB. (c) SNR = 15 dB. In the maps of (a,b,c), black solid lines enclose the embedding regions corresponding to the peak within the corresponding histograms. The histogram peak provides $\widehat{\nu} = 2.0 \pm 0.2$ in (a) and (b), $\widehat{\nu} = 2.2 \pm 0.2$ in (c). Within the maps in (a,b,c), the black dash-dotted lines correspond to the hyperbolae approximately bounding the plateau and defined by $L = 20/(m-1)$ and $L = 80/(m-1)$. (d) SNR = 6 dB. (e) SNR = 0 dB. (f) SNR = −6 dB. For the sequences in (d,e,f), no plateau can be identified.*

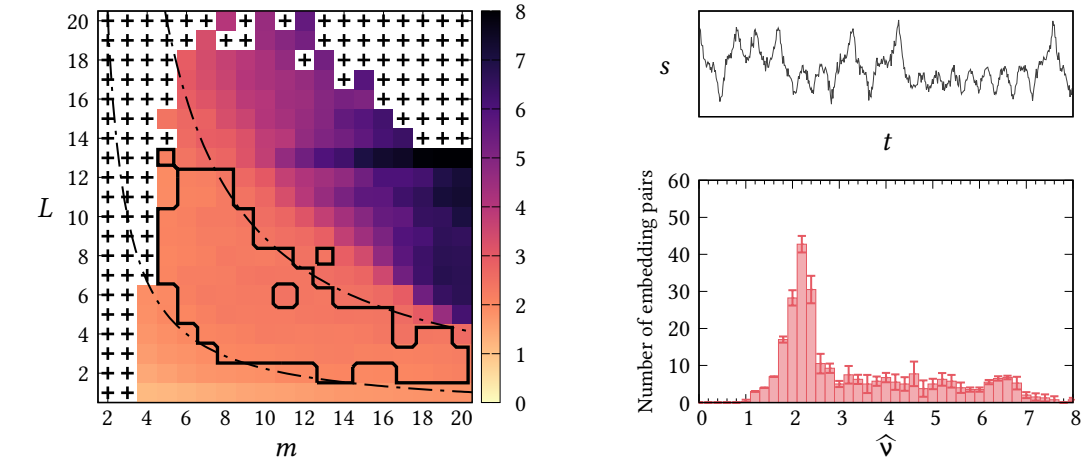
[5] The sequence corresponds to the x-coordinate of a Lorenz system. See Chapter 1, Eq. (1.2) for the system's equations; here parameters are set to $\sigma = 10$, $r = 28$, $b = 8/3$, and the integration time step is $\delta t = 0.03$, as in Chapter 3.



(a)

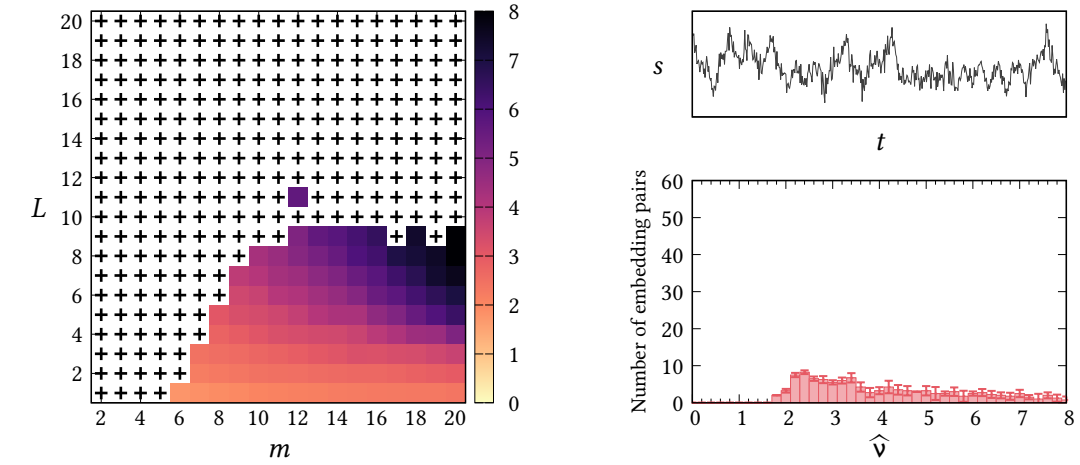


(b)

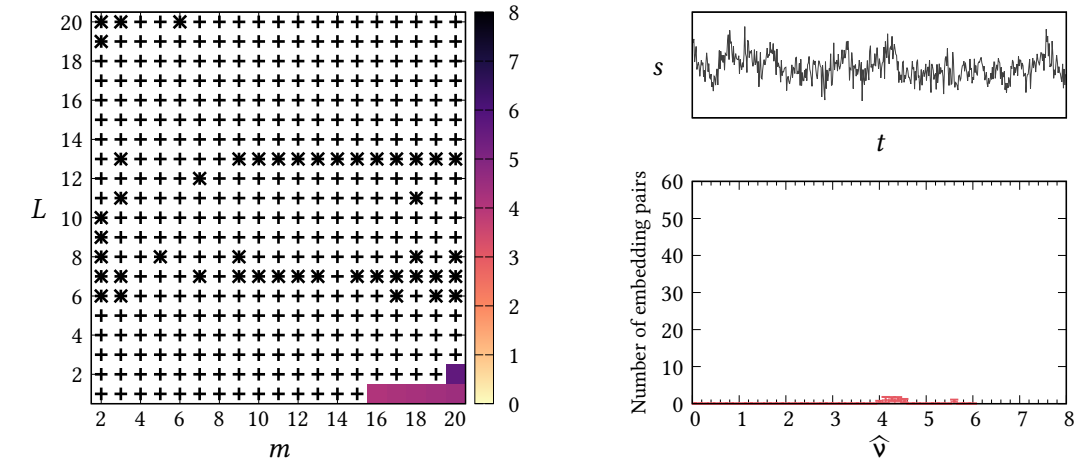


(c)

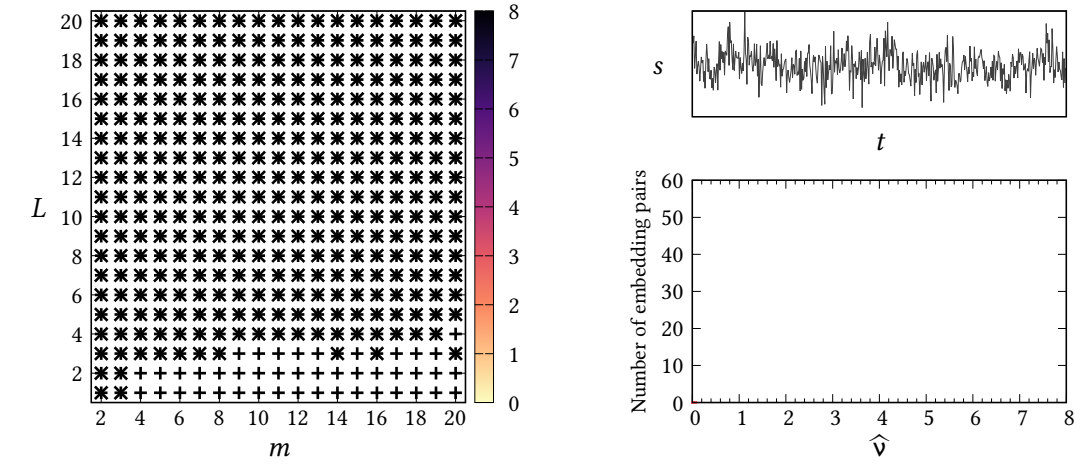
FIGURE 4.10 – *Full caption reported in main text. (a) SNR = 40 dB. (b) SNR = 20 dB. (c) SNR = 15 dB.*



(d)



(e)



(f)

FIGURE 4.10 – *Full caption reported in main text. (d) SNR = 6 dB. (e) SNR = 0 dB. (f) SNR = −6 dB.*

The noisy Lorenz sequences having SNR $\gtrsim$ 15 dB provide maps that are essentially unchanged with respect to the noiseless case (Fig. 3.7). These maps exhibit a $\widehat{\nu}$ plateau that is approximately bounded by two hyperbolae and that corresponds to $\widehat{\nu} \sim 2$. The embedding lattice method is thus resilient down to a SNR given by SNR $\approx$ 15 dB. The presence of observational noise has the expected effect of slightly increasing the estimated dimension: the peak in the $\widehat{\nu}$ histograms "drifts", for decreasing SNR, towards higher $\widehat{\nu}$ values. However, the predominant effect of observational noise on the maps is the reduction in the number of embedding points that are compatible with the null hypothesis $\mathcal{H}_{0;\mathrm{CFD}}$ of a chaotic, finite dimensional system. When the SNR is reduced to $\lesssim$ 6 dB, most embedding points are deemed to be incompatible with this null hypothesis (black "+" marks within the maps): the embedding lattice maps resembles the one obtained for filtered noise (see Fig. 4.4). In other words, for sufficiently high noise power, the embedding lattice method excludes most embedding points because the corresponding state space reconstruction is deemed to be infinite dimensional (i.e. noisy, though not white). Finally, when the power associated to the noisy contribution becomes too large (SNR $<$ 0 dB, Fig. 4.10(f)), the majority of embedding points yields statistical compatibility with a GWN sequence, as one might expect[6].

[6]This compatibility exactly holds for any (m, L) in the limit SNR $\longrightarrow -\infty$.

5. Improved estimator for correlation dimension

A relevant task in nonlinear time series analysis is the discrimination of chaos from noise. As shown in an example within Chapter 4, the method discussed in Chapter 3 is a viable tool to carry out this task. Further insights into the crucial issue of distinguishing stochastic sources from deterministic ones are provided by the analysis of the asymptotic behaviour of the so-called *time-dependent divergence exponent*. As discussed in this chapter, the results of this analysis are two-fold: on the one hand, the asymptotic behaviour of the divergence exponent indeed discriminates between stochastic sources and chaotic ones; on the other hand, in the case of chaotic sequences, the analysis also yields a precise estimation of the correlation dimension. Therefore, the investigation presented in this chapter provides an alternative way to estimate $\widehat{\nu}$, thus possibly improving the third stage of the method described in Chapter 3.

The time-dependent divergence exponent is an embedding-dependent metric. This metric is commonly used in the so-called *divergence rate method* to estimate the maximum Lyapunov exponent out of a scalar sequence, as introduced in Sec. 2.5. To this purpose, one processes the divergence exponent in the "short time" region where, in the case of a chaotic source, it exhibits a linear growth. However, relevant information on the source of the sequence under investigation can be extracted from the—usually overlooked—asymptotic behaviour of the divergence exponent. More in detail, the dependence of the asymptotic value of this metric on one of the parameters used in its estimation, namely the "starting percentile" p, is characteristic of the class of systems that generate the input sequence [175]. Two classes of systems are considered here, namely GWN and chaotic sources: in both cases, approximate expressions for the dependence of the asymptotic value of the divergence exponent on the starting percentile can be derived. Therefore, the analysis of this dependence provides a marker to identify GWN or chaotic sequences. Furthermore, in the case of a chaotic source, this dependence is linear in $\log p$, and the corresponding slope is determined by the correlation dimension of the underlying system. By virtue of this property, the assessment of the asymptotic value of the divergence exponent out

of a scalar sequence makes up an alternative tool to precisely estimate the correlation dimension of a system.

Section 5.1 introduces the definition of the divergence exponent and its asymptotic behaviour, as well as the procedure used in the present work to estimate it out of a scalar sequence. The case of a GWN sequence is presented in Sec. 5.2, along with examples concerning other kinds of noise. Section 5.3 deals with the case of a chaotic source, introducing the novel estimator of correlation dimension. Thereupon, the application to benchmark synthetic and experimental sequences is illustrated in Sec. 5.4. The improved estimator of correlation dimension is mapped, in Sec. 5.5, on the embedding lattice as an alternative implementation of the third stage of the method for the identification of suitable embedding parameters. The improved estimator yields results that are compatible–though superior in precision—with those obtained with the estimator stemming from the gauge transformation of correlation integrals.

5.1 Time-dependent divergence exponent and its *plateau*

The input to the following discussion is, as usual, a scalar sequence $\{y_n\}$. The divergence rate exponent is a state space metric. Therefore, given a choice of embedding parameters m, L, a sequence of embedding vectors $\mathbf{Y}_n = (y_n, y_{n+L}, \ldots, y_{n+(m-1)L})$ is built out of the input scalar sequence. We start by considering a pair of embedding vectors, namely $\mathbf{Y}_i, \mathbf{Y}_j$, and the corresponding Euclidean distance $\zeta_{i,j}$ in the m-dimensional space, $\zeta_{i,j} \equiv \|\mathbf{Y}_i, \mathbf{Y}_j\|$. The two vectors are chosen such that their time separation $|i - j|$ is much larger than the autocorrelation time of the sequence (see below for a formal criterion). After k time steps, the two vectors evolve into $\mathbf{Y}_{i+k}, \mathbf{Y}_{j+k}$ and the corresponding distance becomes $\zeta_{i+k,j+k}$. As introduced in Sec. 2.5, the divergence rate method relies on the estimation of the time-dependent divergence exponent $\Lambda(k; r)$ [172], namely the expected value defined by

$$\Lambda(k; r) \equiv \mathrm{E}\left[\log\left(\frac{\zeta_{i+k,j+k}}{\zeta_{i,j}}\right)\Bigg|_{\zeta_{i,j} \leqslant r}\right].$$

The parameter r, or *shell radius*, defines a maximum distance such that only initially close vectors ($\zeta_{i,j} \leqslant r$) are included in the evaluation. Typically, r is defined as follows [133]: given the distribution $f_\zeta(\zeta)$ of Euclidean distances between embedding vectors, a *starting percentile* p ($p \in [0, 1]$) defines a maximum distance $r(p)$ according to the implicit relation

$$\int_0^{r(p)} f_\zeta(\zeta)\mathrm{d}\zeta = p\,. \tag{5.1}$$

It is straightforward to see that $p \to 1 \iff r \to \infty$. By explicitly taking into account the dependence of r on p, the time-dependent divergence exponent $\Lambda(k; r)$ is rewritten as $\Lambda(k; p)$.
For chaotic systems, $\Lambda(k; p)$ grows linearly with k as a consequence of the exponential divergence of orbits. As described in Sec. 2.5, the slope of this linear growth provides an estimate of the maximum Lyapunov exponent, which is the goal of the divergence rate method. On the other hand, for large enough k, the growth stops and $\Lambda(k; p)$ saturates to an asymptotic "plateau" Λ_{pl} defined as

$$\Lambda_{\text{pl}}(p) \equiv \lim_{k \to \infty} \Lambda_{k;p}\,, \tag{5.2}$$

where the dependence on p is retained. Indeed, the method discussed in this chapter relies on the dependence of Λ_{pl} on p to provide information on the dynamical system generating the input sequence. Below, Sec. 5.2 and Sec. 5.3 provide a derivation of this dependence in the case of GWN and in the case of a chaotic source, respectively.

So far, the discussion was carried out by assuming an infinitely long sequence. The assessment of the plateau Λ_{pl} out of a finite sequence containing ℓ elements is carried out according to the following procedure [133, 175]. An embedding pair (m, L) and a starting percentile p are given. First, the sequence of embedding vectors $\mathbf{Y}_n$ is built out of the scalar sequence y_n according to (m, L). Then, five steps are implemented.

1) The shell radius $\hat{r}(p)$ for the given starting percentile p has to be estimated. To this purpose, a sample distribution of Euclidean distances is built by randomly picking η pairs of embedding vectors and thereupon collecting the corresponding η Euclidean distances. The distance value that ranks—in ascending order—at the $\rho = p\eta$ position is selected as $\hat{r}(p)$. Vice versa, one can first fix the desired rank ρ, randomly extract ρ/p distances, and finally determine $\hat{r}(p)$ as the distance value corresponding to the rank ρ. In the following, this second procedure is used.

2) Given the delay $k \in \mathbb{N}^+$, a set $\mathcal{S}$ of N pairs of embedding vectors $\mathbf{Y}_i, \mathbf{Y}_j$ is randomly extracted without replacement. Pairs have to comply with three requirements:

   - because the number of available pairs is $\sim \ell^2/2$, the number of selected pairs has to be $N \ll \ell^2$;
   - temporally correlated pairs have to be excluded from the selection (see also Sec. 2.5): the indexes i, j have to comply with $|i - j| \geqslant c_0$, where c_0 is equal to the second zero of the autocorrelation function multiplied times $m(L + 1)$ [133, 175];
   - vectors within a pair have to be closer than the shell radius $\hat{r}$, i.e. $\|\mathbf{Y}_i - \mathbf{Y}_j\| \leqslant \hat{r}(p)$.

3) An estimate of the time-dependent divergence exponent $\widehat{\Lambda}(k)$ is assessed upon the set $\mathcal{S}$ by taking

the following sample mean:

$$\widehat{\Lambda}(k) \equiv \left\langle \log\left(\frac{\|\mathbf{Y}_{i+k} - \mathbf{Y}_{j+k}\|}{\|\mathbf{Y}_i - \mathbf{Y}_j\|} \right) \right\rangle . \tag{5.3}$$

An example of the assessment of $\widehat{\Lambda}(k)$ is shown in Chapter 2, Fig. 2.6. The evaluation of $\widehat{\Lambda}(k)$ is carried out over values of k that are multiples of a delay δk, which is usually set to unity.

4) The region where $\widehat{\Lambda}(k)$ reaches its saturation (plateau) value is defined as the set $\mathcal{K}$ of k values satisfying $k \geqslant k_{\min}$. A sample plateau value $\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}}$ is computed by fitting a straight line on the points $(k, \widehat{\Lambda}(k))$ for which $k \in \mathcal{K}$:

$$\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}} \equiv s^2_{\Lambda_{\mathrm{pl};\mathcal{S}}} \sum_{k\in\mathcal{K}} \frac{\widehat{\Lambda}(k)}{s^2_\Lambda}, \tag{5.4a}$$

$$s^2_{\Lambda_{\mathrm{pl};\mathcal{S}}} \equiv \left(\sum_{k\in\mathcal{K}} \frac{1}{s^2_\Lambda} \right)^{-1} . \tag{5.4b}$$

In the last expressions, s_Λ is the sample standard deviation compute over the N elements contributing to the sample mean of Eq. (5.3).

So far, the set $\mathcal{K}$ was assumed to belong to the plateau. The appropriateness of this assumption—i.e. the quality of the straight line fit defined by Eq. (5.4)—is checked by relying on a χ^2 test. The number of degrees of freedom is $Q-1$, where Q is the number of k values within the range $\mathcal{K}$, and the sample statistic for the test is

$$\hat{\chi}^2 = \left(\sum_{k\in\mathcal{K}} \frac{\widehat{\Lambda}^2(k)}{s^2_\Lambda} \right) - \frac{\widehat{\Lambda}^2_{\mathrm{pl};\mathcal{S}}}{s^2_{\mathrm{pl};\mathcal{S}}} .$$

Given a significance level α, if $\hat{\chi}^2$ overcomes the χ^2_{Q-1} critical value corresponding to α, then the null hypothesis of $\mathcal{K}$ belonging to the plateau is rejected.

5) The major source of error on $\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}}$ is not given by the statistical fluctuations of $\widehat{\Lambda}(k)$—accounted for by $s^2_{\Lambda_{\mathrm{pl};\mathcal{S}}}$—but rather by the bias in $\widehat{\Lambda}(k)$ due to the uncertainty in the assessment of $\hat{r}(p)$. This issue is shown in Fig. 5.1: independent evaluations of $\widehat{\Lambda}(k)$, each relying on an independent assessment of $\hat{r}(p)$, provide systematically lower (or higher) values of $\widehat{\Lambda}(k)$, and thus mutually inconsistent estimated plateaus $\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}}$. To overcome this isseu, rather than generating a single set $\mathcal{S}$, the steps 1)–4) have to be repeated M times, each time independently generating a set $\mathcal{S}_h$ and evaluating the corresponding plateau according to Eq. (5.4). The sample mean and standard deviation over these

M evaluations yield the "final" value of the plateau and its uncertainty:

$$\widehat{\Lambda}_{\mathrm{pl}} \equiv \frac{1}{M}\sum_{h=1}^{M}\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}_h}\,, \tag{5.5a}$$

$$s^2_{\Lambda_{\mathrm{pl}}} \equiv \frac{1}{M-1}\sum_{h=1}^{M}\left(\widehat{\Lambda}_{\mathrm{pl};\mathcal{S}_h} - \widehat{\Lambda}_{\mathrm{pl}}\right)^2 . \tag{5.5b}$$

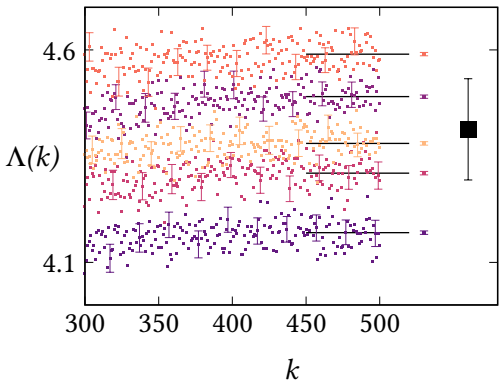


FIGURE 5.1 – *The variability of $\widehat{\Lambda}(k)$ due to statistical fluctuations and due to the uncertainty in the assessment of $\hat{r}(p)$ are compared. Data points correspond to five independent evaluations (each one plotted with a differen color) of $\widehat{\Lambda}(k)$, carried out on $k \in [300, 500]$. Each evaluation is obtained by setting $p = 10^{-4}$ and by independently estimating $\hat{r}(p)$. For the sake of clarity, errorbars are reported only once every 20 data points. Horizontal black lines correspond to the sample plateau value $\widehat{\Lambda}_{pl;\mathcal{S}}$ according to Eq. (5.4)a evaluated on each independent set of data points. The related uncertainty $s_{\Lambda_{pl;\mathcal{S}}}$ is evaluated according to Eq. (5.4)b and is shown by means of an errorbar just beside the corresponding horizontal line. Because $k_{min} \approx 420$, these evaluations are carried out over the range $k \in \mathcal{K} = [450, 500]$. The large black square dot and the related errorbar correspond to the sample mean and standard deviation (namely Eq. (5.5)) over ten evaluations of $\widehat{\Lambda}_{pl;\mathcal{S}}$, including the five shown here. For all these ten evaluations, the $\hat{\chi}^2$ sample statistic is compatible with the null hypothesis of $\mathcal{K}$ belonging to a plateau.*

Before analyzing the cases of a noisy or a chaotic source, and for reasons that will become clear later, the following function is defined:

$$I(P) \equiv \mathrm{E}\left[\log\left(\zeta_{i,j}\right)\big|_{\zeta_{i,j}\leqslant r(p)}\right] = \frac{1}{p}\int_0^{r(p)} \log(\zeta) f_\zeta(\zeta)\mathrm{d}\zeta\,, \tag{5.6}$$

where the shell radius $r(p)$ is defined by the implicit expression of Eq. (5.1). Let us now assume that the dynamical system that generates the sequence under investigation is ergodic or, alternatively, that sequence elements taken at a sufficiently large time separation are uncorrelated. If either of these conditions holds, in the $k \to \infty$ limit, the distribution of the "delayed" distances $\zeta_{i+k,j+k}$ coincides with $f_\zeta(\zeta)$ regardless of p, and thus regardless of the initial distance $\zeta_{i,j}$. Consequently, if $k \to \infty$, the expected

value of $\log\left(\zeta_{i+k,j+k}\right)$ turns out to be

$$\lim_{k\to\infty} \mathrm{E}\left[\log\left(\zeta_{i+k,j+k}\right)|_{\zeta_{i,j}\leqslant r(p)}\right] = \int_0^\infty \log(\zeta) f_\zeta(\zeta)\mathrm{d}\zeta = I(1)\,.$$

This last expression, together with Eq. (5.6), allows us to write the plateau $\Lambda_{\mathrm{pl}}(p)$, defined in Eq. (5.2), as the following difference:

$$\Lambda_{\mathrm{pl}}(p) = I(1) - I(p)\,. \tag{5.7}$$

As one might expect, $\Lambda_{\mathrm{pl}}(1) = 0$. The smaller p, the smaller is the contribution of $I(p)$, and thus the larger is the value attained by the plateau.

5.2 Plateau behaviour in the case of noise sources

The first kind of sequence that we consider is GWN: sequence elements are i.i.d. random variates distributed as $y_n \sim \mathcal{N}(\mu, \sigma^2)$. Because sequence elements are independent, the plateau starts *exactly* at $k_{\min} = (m-1)L+1$: for any value of $k \geqslant k_{\min}$, the divergence exponent $\Lambda(k)$ is equal to its plateau value Λ_{pl}. The independence of y_i, y_j for any $i \neq j$ also implies that the generic distance ζ is given by

$$\zeta = \left[\sum_{h=0}^{m-1}\left(y_{i+hL} - y_{j+hL}\right)^2\right]^{\frac{1}{2}} = \left(2\sigma^2 x\right)^{\frac{1}{2}}\,, \tag{5.8}$$

where σ^2 is the variance of the sequence elements $\{y_n\}$, the factor 2 is due to the difference between two normal random variables, and x is itself a random variable distributed according to a χ^2 distribution with m degrees of freedom. The corresponding probability density function is

$$f_{\chi^2_m}(x) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} e^{-\frac{x}{2}} x^{\frac{m}{2}-1}\,.$$

Because the divergence exponent depends on the ratio between two distances, the variance σ^2 is immaterial. Therefore, without loss of generality, we can set $\sigma^2 = 1/4$ so that $\zeta = \sqrt{x/2}$. By defining $t \equiv x/2 = \zeta^2$, $\theta \equiv r^2$, and $s \equiv m/2$, the implicit expression relating the shell radius to the starting percentile, i.e. Eq. (5.1), becomes

$$p = \frac{1}{\Gamma(s)}\int_0^\theta e^{-t} t^{s-1}\mathrm{d}t = \frac{\gamma(s,\theta)}{\Gamma(s)}\,, \tag{5.9}$$

where $\gamma(s,\theta)$ is the lower incomplete Gamma function. Inverting Eq. (5.9) would provide an expression for θ as a function of p and m.

The change of variables defined above transforms the integral $I(p)$ of Eq. (5.6) into the following one:

$$I(p) = \frac{1}{2p\Gamma(s)} \int_0^{\theta} e^{-t} t^{s-1} \log(t) \mathrm{d}t = \frac{1}{2\gamma(s,\theta)} \frac{\partial \gamma(s,\theta)}{\partial s} .$$

The limit $p \to 1$, which corresponds to a diverging shell radius $r \to \infty$, also gives $\theta \to \infty$. In this limit,

$$\lim_{\theta \to \infty} \gamma(s,\theta) = \Gamma(s) ,$$

and therefore

$$I(1) = \frac{1}{2}\psi(s) ,$$

where $\psi(s) = \frac{\mathrm{d}}{\mathrm{d}s} \log[\Gamma(s)]$ is the digamma function. By joining together the last expressions for $I(p)$ and $I(1)$, the plateau value for a GWN sequence $\Lambda_{\mathrm{pl;GWN}}(p;s)$ is written as

$$\Lambda_{\mathrm{pl;GWN}}(p;s) = \frac{1}{2} \left\{ \psi(s) - \frac{\partial}{\partial s} \log\left[\gamma(s,\theta)\right] \right\} .$$

It is worth recalling that r, and thus θ, is a function of p and s; by virtue of Eq. (5.9), θ, p, and s satisfy the equation

$$\gamma(s,\theta) = p\Gamma(s) .$$

In order to provide an explicit expression of $\Lambda_{\mathrm{pl;GWN}}(p;s)$ as a function of p, the relationship $\gamma(s,\theta) = p\Gamma(s)$ has to be inverted. However, this inversion has no analytical solution. An approximated solution, as well as its throughout derivation, is provided in Ref. [175]. The plateau in the case of GWN sequences turns out to be given by

$$\Lambda_{\mathrm{pl;GWN}}(p;s) = \Lambda' - \frac{1}{2s} \log(p) + \delta\Lambda(p;s) , \tag{5.10}$$

where Λ' is a constant term and $\delta\Lambda(p;s)$ is a correction to the linear behaviour in $\log(p)$, which vanishes when $p \to 0$ (i.e. when $\theta \to 0$). The constant term Λ' is defined as

$$\Lambda' = \frac{1}{2}\psi(s) + \frac{1}{2s} \left\{ 1 - \log\left[\Gamma(s+1)\right] \right\} .$$

On the other hand, by defining $P \equiv [p\Gamma(s+1)]^{1/s}$, the correction term $\delta\Lambda$ is approximately given by

$$\delta\Lambda(p;s) = -\frac{P}{2} \frac{s}{(s+1)^2} - \frac{P^2}{4} \frac{s(2s+3)}{(s+1)^2(s+2)^2} - \frac{P^3}{12} \frac{s(9s^2+34s+29)}{(s+1)^3(s+2)(s+3)^2} + O(P^4) . \tag{5.11}$$

This third-order approximation of the correction term $\delta\Lambda$ holds for $P \lesssim 1$, namely for starting percentiles $p \lesssim \Gamma^{-1}(\frac{m}{2}+1)$.

To test the results presented so far, the plateau $\widehat{\Lambda}_{\mathrm{pl}}$ is estimated out of a GWN sequence (length $\ell = 10^5$ points) for different starting percentile p. The estimation is carried out according to the procedure described in Sec. 5.1 and by setting the related parameters to $N = 50$, $M = 10$, $k \in [20, 40]$, $\delta k = 1$. Two sets of estimated $\widehat{\Lambda}_{\mathrm{pl}}(p)$ are shown in Fig. 5.2, corresponding to the embedding choices $(5, 1)$ and $(10, 1)$. Moreover, the exact (numerically-evaluated) expression of Eq. (5.10), as well as its approximation

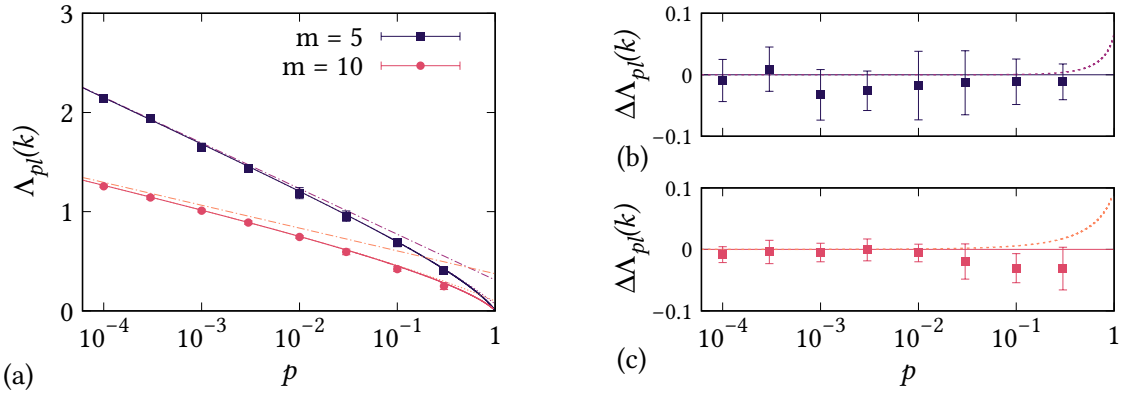


FIGURE 5.2 – *(a) Plateau value Λ_{pl} as a function of the starting percentile p for a GWN sequence. The embedding choices $(5, 1)$ and $(10, 1)$ are shown. Data points correspond to numerical evaluations of $\widehat{\Lambda}_{pl}$ as described in Sec. 5.1. Solid lines correspond to the exact expression of $\Lambda_{pl;GWN}$ of Eq. (5.10) and numerically calculated for $m = 5$ and $m = 10$, respectively. Dash-dotted and dashed lines correspond to the approximated expressions of $\Lambda_{pl;GWN}$ provided by the linear approximation ($\Lambda_{pl;GWN} \simeq \Lambda' - \frac{1}{m}\log(p)$) and the third-order approximation obtained by inserting Eq. (5.11) in Eq. (5.10), respectively. The cases $m = 5$ ($m = 10$) is plotted in purple (orange). The third-order approximation (dashed lines) is, in both cases, almost undistinguishable from the exact expression (solid lines). (b) Deviations $\Delta\Lambda_{pl}$ of data points and of the third-order approximation (dashed line) from the exact expression of $\Lambda_{pl;GWN}$ given by Eq. (5.10) in the case $m = 5$. (c) Same as (b), but in the case $m = 10$.*

relying on Eq. (5.11), are displayed as solid and dashed lines. The discrepancy between the exact and the approximated curve is shown in Fig. 5.2(b): the approximation is good up to $p \gtrsim 0.1$. Finally, Fig. 5.2(a) also shows the zeroth-order approximation $\Lambda_{\mathrm{pl;GWN}}(p; s) \simeq \Lambda' - \frac{1}{m}\log(p)$: this approximation, corresponding to $\delta\Lambda \simeq 0$, fails for $p \gtrsim 10^{-2}$.

We now briefly discuss the behaviour of the plateau Λ_{pl} in the case of noise sources different from GWN. First, we consider noise sources that are *white* but not Gaussian. By virtue of the central limit theorem, if m is large enough, the distribution of the sequence elements $\{y_n\}$ does not matter in determining the plateau value: the distances defined by Eq. (5.8) are approximately normally distributed, regardless of the distribution of $\{y_n\}$. Therefore, the results presented so far hold for white noise with *any* distribution, not only Gaussian. Figure 5.3 shows the same analysis of Fig. 5.2(a) for *uniform* white noise, i.e. a sequence of i.i.d. realizations of a random variate distributed uniformly between $[0, 1]$ ($\ell = 10^5$). The sequence is embedded with $(m, L) = (5, 1)$. Again, data points are estimated according to Sec. 5.1;

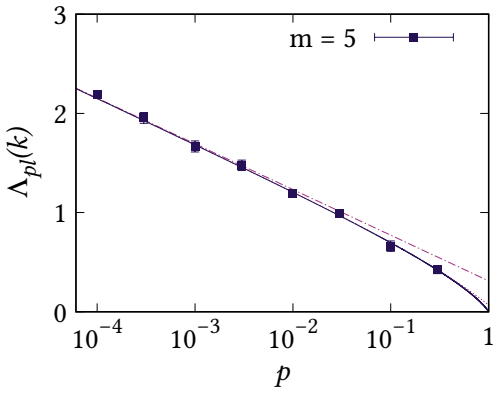


FIGURE 5.3 – *Plateau value Λ_{pl} as a function of the starting percentile p for a uniform white noise sequence embedded with $(m, L) = (5, 1)$. Data points correspond to numerical evaluations of $\widehat{\Lambda}_{pl}$ as described in Sec. 5.1. The solid line corresponds to the exact expression of $\Lambda_{pl;GWN}$ of Eq. (5.10) and numerically calculated for $m = 5$. The dash-dotted and dashed lines correspond to the approximated expressions of $\Lambda_{pl;GWN}$ provided by the linear approximation ($\Lambda_{pl;GWN} \simeq \Lambda' - \frac{1}{m}\log(p)$) and the third-order approximation obtained by inserting Eq. (5.11) in Eq. (5.10), respectively, with m set to 5. The third-order approximation (dashed line) is almost undistinguishable from the exact expression (solid line).*

parameter settings are the same as in the case of GWN. The plot turns out to be indistinguishable from the one in Fig. 5.2(a) for a GWN sequence embedded with the same embedding parameters: data points are well described by the numerically-evaluated curve given by Eq. (5.10), as well as by its third-order approximation provided by Eq. (5.11).

The second non-GWN source considered is filtered noise. Two sequences ($\ell = 10^5$) are generated by piping i.i.d. normal random variates through a first-order digital low-pass filter with cut-off frequency $f_{3\text{dB}} = 0.1\ T^{-1}$ and $f_{3\text{dB}} = 0.02\ T^{-1}$, respectively ($T$ is the sampling period). The same plateau analysis of Fig. 5.2 is carried out on both these sequences, thus producing the plots of Fig. 5.4 ($f_{3\text{dB}} = 0.1\ T^{-1}$) and Fig. 5.5 ($f_{3\text{dB}} = 0.02\ T^{-1}$). For both filtering setups, data points corresponding to the assessment of

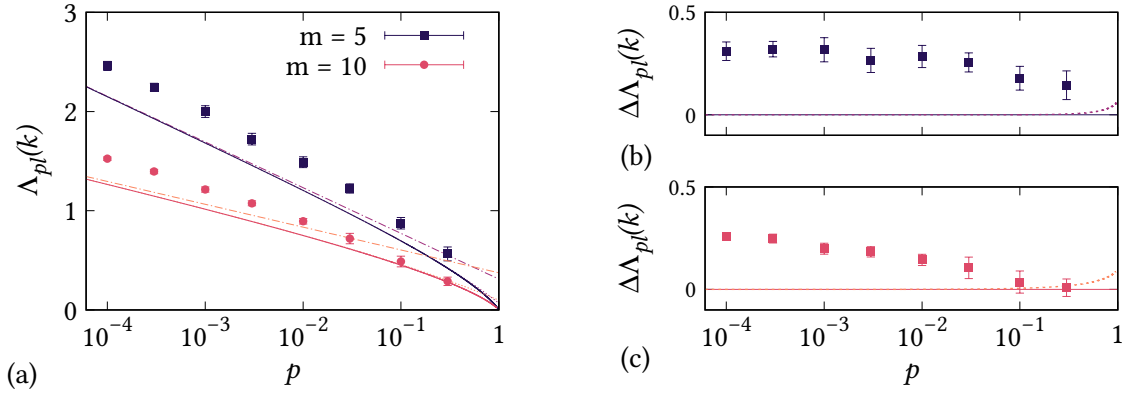


FIGURE 5.4 – *Plateau value Λ_{pl} (a) and corresponding deviations $\Delta\Lambda_{pl}$ (b,c) as a function of the starting percentile p for a filtered noise sequence. The sequence corresponds to GWN filtered by means of a first-order digital low-pass filter with $f_{3dB} = 0.1\ T^{-1}$, where T is the sampling period. The embedding choices $(5, 1)$ and $(10, 1)$ are shown. The figure description is the same as in the case of GWN shown in Fig. 5.2.*

$\widehat{\Lambda}_{\text{pl}}$ out of the scalar sequences deviate from the theoretical predictions for GWN. In other words, the

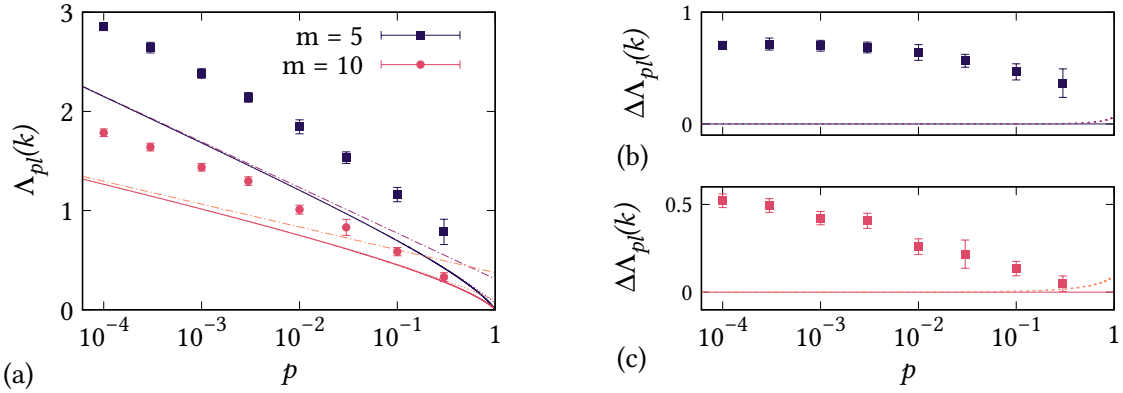


FIGURE 5.5 – *Plateau value Λ_{pl} (a) and corresponding deviations $\Delta\Lambda_{pl}$ (b,c) as a function of the starting percentile p for a filtered noise sequence. The sequence corresponds to GWN filtered by means of a first-order digital low-pass filter with $f_{3dB} = 0.02\ T^{-1}$, where T is the sampling period. The embedding choices* $(5, 1)$ *and* $(10, 1)$ *are shown. The figure description is the same as in the case of GWN shown in Fig. 5.2.*

dependence of $\widehat{\Lambda}_{\mathrm{pl}}$ on p is sensitive to the finite autocorrelation of noise sources. As one might expect, the discrepancy between data points corresponding to filtered noise and the theoretical predictions for GWN is larger if the typical autocorrelation time of the input sequence is longer. Indeed, the difference between data points and theoretical prediction is almost double in Fig. 5.5(b) with respect to Fig. 5.4(b), as a consequence of the longer autocorrelation time of the sequence filtered with $f_{3\mathrm{dB}} = 0.02\ T^{-1}$ with respect to the sequence filtered with $f_{3\mathrm{dB}} = 0.1\ T^{-1}$.

5.3 Plateau behaviour in the case of chaotic, finite-dimensional sources

The case of a chaotic dynamical system being the source that generates the sequence y_n is now examined. To evaluate the plateau value $\Lambda_{\text{pl;chaos}}$, we can assume the distribution of distances $f_\zeta(\zeta)$ to be determined (at small distances, $\zeta \to 0$) by the Grassberger-Procaccia power-law scaling (see Sec. 2.4):

$$C_\zeta(\zeta) \simeq \beta \zeta^\nu \,,$$

The exponent ν is the correlation dimension and β is an embedding-dependent constant. Differentiating the last expression provides the probability density function of distances:

$$f_\zeta(\zeta) = \frac{\mathrm{d}C_\zeta(\zeta)}{\mathrm{d}\zeta} \simeq \nu\beta\zeta^{\nu-1} \,.$$

It is worth remarking that this expression for $f_\zeta(\zeta)$ holds only in the small-scale regime.
According to its definition in Sec. 5.1, given a value of the starting percentile p, the shell radius $r(p)$ is determined by

$$C\left[r(p)\right] = p \quad \Longrightarrow \quad r(p) \simeq \left(\frac{p}{\beta}\right)^{\frac{1}{\nu}} \,.$$

This expression for $r(p)$ allows evaluating directly the integral $I(p)$ defined in Eq. (5.6). This (approximated) evaluation reads:

$$I(p) \simeq \frac{1}{\nu}\left[\log(p) - \log(e\beta)\right] \,.$$

Again, the validity of the last expression is limited to the small scale regime, where the Grassberger-Procaccia scaling holds. Therefore, because $p \to 1$ corresponds to $r \to \infty$, $I(1)$ *cannot* be written as $-\log(e\beta)/\nu$, and the constant $I(1)$ remains, in general, unknown. Nevertheless, provided that $m > \nu$, and in the small scale regime, the plateau value of a chaotic system can be written as

$$\Lambda_{\text{pl;chaos}}(p) \simeq \Lambda' - \frac{1}{\nu}\log(p) \,, \tag{5.12}$$

where the constant Λ' is given by $I(1) + \log(e\beta)/\nu$. Therefore, in the case of a chaotic source, the plateau value is given by an embedding-dependent constant minus a term which is linear in $\log(p)$ and whose slope is determined solely by the correlation dimension of the system, and not by the embedding choice, as long as $m > \nu$. This behaviour clearly differs from the GWN case thus allowing to distinguish the two classes of systems. Moreover, the $\log(p)$ dependence given by Eq. (5.12) makes up a tool to estimate the correlation dimension of a system out of an input sequence, for example by means of a linear fit.

As a first example, a sequence corresponding to the x-coordinate of a Lorenz system[1] is considered.

[1] See Chapter 1, Eq. (1.2) for the system's equations. Here, the system parameters are set to $\sigma = 10$, $r = 28$, $b = 8/3$ and the

The sequence is embedded with $(m, L) = (4, 6)$ and $(8, 6)$, respectively. These embedding choices were selected according to the results of the method discussed in Chapter 3 (see Fig. 3.7). While $(8, 6)$ is roughly in the middle of the region corresponding to the set of "suitable" embedding parameters, $(4, 6)$ is just outside that region. The results of the assessment of the plateau values $\widehat{\Lambda}_{\mathrm{pl}}$ according to the procedure described in Sec. 5.1 are reported in Fig. 5.6. The figure also shows, by means of dashed

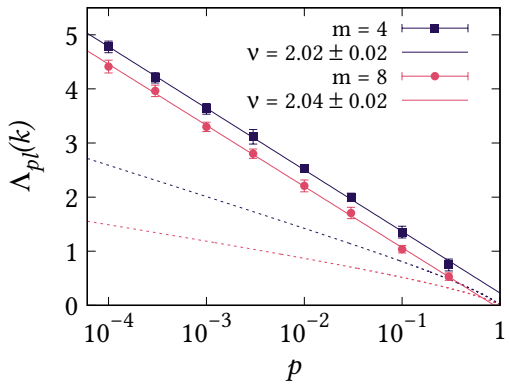


FIGURE 5.6 – *Plateau value Λ_{pl} as a function of the starting percentile p for a Lorenz sequence. The embedding choices $(4, 6)$ and $(8, 6)$ are shown. Data points correspond to numerical evaluations of $\widehat{\Lambda}_{pl}$ as described in Sec. 5.1 ($N = 500$, $M = 10$, $k \in [450, 500]$, $\delta k = 1$). The solid lines each correspond to a fit with a straight line ($\log(p)$ is the independent variable) on the data points plotted with the same color. The slopes of the straight lines provide an estimate of the correlation dimension, which is reported in the figure. Dashed lines correspond to the exact expression of $\Lambda_{pl;GWN}$ given by Eq. (5.10) and numerically calculated for $m = 4$ and $m = 8$, respectively (line colors match the color of data points corresponding to the same embedding dimension).*

lines, the expected plateau values in the case of GWN sequences: data are definitely not compatible with a GWN source. In addition, linear fits are carried out the sets of data corresponding to both embedding choices. The resulting values of ν are reported in the figure and provide a precise estimate of the correlation dimension, which is in agreement with the accepted value for the Lorenz attractor given by $\nu = 2.049 \pm 0.096$ [14].

Going back to the results for filtered noise shown in Fig. 5.4 and Fig. 5.5, one might argue whether those values of $\widehat{\Lambda}_{\mathrm{pl}}$ can be interpreted as corresponding to a chaotic source. Assuming that no other analysis (e.g. by relying on the method discussed in Chapter 3) is performed, the only way to answer this question is to check whether the slope of the $\widehat{\Lambda}_{\mathrm{pl}}$ vs $\log(p)$ curve is embedding-independent, as predicted by Eq. (5.12). In the case of the data of Fig. 5.4, a straight line fit on the data corresponding to $m = 10$ provides a slope which differs by a factor ~ 1.6 from the slope corresponding to $m = 5$. Similarly, in Fig. 5.5 the ratio between the slopes provided by fitting is ~ 1.4. Consequently, one can conclude that those data do not correspond to an underlying chaotic source.

integration time step is $\delta t = 0.03$.

5.4 Application to synthetic and experimental sequences

In this Section, the plateau analysis relying on the $\widehat{\Lambda}_{\text{pl}}$ assessment protocol of Sec. 5.1 and on the results of Sec. 5.3 is further tested on a synthetic and an experimental sequence. The first sequence corresponds to the y-coordinate of a Ueda oscillator[2]. According to the results concerning this sequence that are presented in Chapter 4, and in particular in Fig. 4.3(c,d), three embedding choices are identified: $(m, L) = (2, 1)$, which was deemed to be an unsuitable embedding choice, $(6, 8)$ and $(10, 12)$, both supposed to be suitable to reconstruct the system's dynamics. The plateau analysis is carried out by setting $N = 500$, $M = 10$, $k \in [2000, 3000]$, $\delta k = 5$. The results of the analysis for the Ueda sequence are shown in Fig. 5.7. In the case of the embedding parameters set to $(2, 1)$, despite data are fit by a straight line, the corresponding $\nu = 2.02 \pm 0.02$ turns out to be equal to the embedding dimension $m = 2$. This results hints at the fact that the condition $m > \nu$ is not satisfied, i.e. $m = 2$ is insufficient to reconstruct correctly the dynamics, and Eq. (5.12) does not hold. On the other hand, selecting the embedding parameters equal to $(6, 8)$ or to $(10, 12)$ provides consistent, embedding-independent estimates of the correlation dimension. By virtue of these estimates, one can correctly conclude that the source of the input sequence is indeed chaotic, with $\nu \simeq 2.7$. Finally, the values of ν estimated through fitting are in agreement with the value reported in the literature equal to 2.675 ± 0.132 [14].

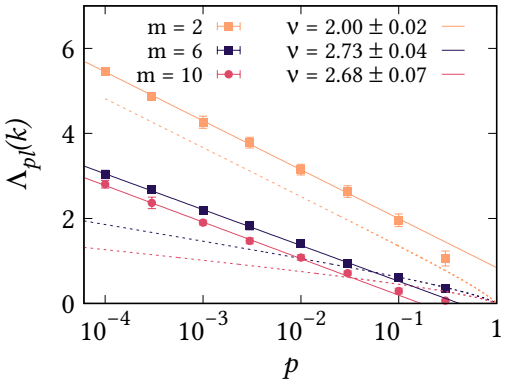


FIGURE 5.7 – *Plateau value Λ_{pl} as a function of the starting percentile p for a Ueda sequence. The embedding choices $(2, 1)$, $(6, 8)$ and $(10, 12)$ are shown. Data points correspond to numerical evaluations of $\widehat{\Lambda}_{pl}$ as described in Sec. 5.1. The solid lines each correspond to a fit with a straight line ($\log(p)$ is the independent variable) on the data points plotted with the same color. The slopes of the straight lines provide an estimate of the correlation dimension, which is reported in the figure. Dashed lines correspond to the exact expression of $\Lambda_{pl;GWN}$ given by Eq. (5.10) and numerically calculated for $m = 2$, $m = 6$ and $m = 10$, respectively (line colors match the color of data points corresponding to the same embedding dimension).*

The second sequence considered as example is the intensity sequence experimentally recorded out of a 81.5 µm NH_3 laser. The sequence was introduced in Chapter 1 and already analyzed in Chapter 4 (see Fig. 4.7). The sequence is embedded according to two different choices of parameters (m, L), namely $(5, 5)$ and $(8, 3)$. Both choices belong to the region of suitable embedding parameters identified in Chap-

[2]See Chapter 1, Eq. (1.3) for the system's equations. Here the system parameters are set to $A = 7.5$, $k = 0.05$, and the integration time step is set to $\delta t = 0.05$.

ter 4 and highlighted in Fig. 4.7. The results of the plateau analysis for the laser sequence are shown in

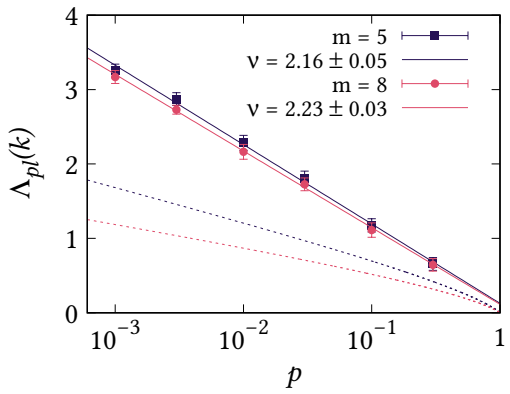


FIGURE 5.8 – *Plateau value Λ_{pl} as a function of the starting percentile p for an experimentally recorded laser intensity sequence. The embedding choices* $(5, 5)$ *and* $(8, 3)$ *are shown. Data points correspond to numerical evaluations of $\widehat{\Lambda}_{pl}$ as described in Sec. 5.1 ($N = 500$, $M = 10$, $k \in [450, 500]$, $\delta k = 1$). The solid lines each correspond to a fit with a straight line ($\log(p)$ is the independent variable) on the data points plotted with the same color. The slopes of the straight lines provide an estimate of the correlation dimension, which is reported in the figure. Dashed lines correspond to the exact expression of $\Lambda_{pl;GWN}$ given by Eq. (5.10) and numerically calculated for $m = 5$ and $m = 8$, respectively (line colors match the color of data points corresponding to the same embedding dimension).*

Fig. 5.8. Both choices of embedding parameters yield a linear dependence of $\widehat{\Lambda}_{\mathrm{pl}}$ on $\log p$, thus supporting the hypothesis on an underlying chaotic source. The corresponding values of ν, which are reported in the figure, are compatible with each other and in agreement with the estimates of Chapter 4.

5.5 Closing the circle: application of the improved estimator in the identification of suitable embedding regions

The main result of Sec. 5.3, namely Eq. (5.12), makes up a tool to estimate the correlation dimension of a system out of a scalar sequence. As shown in Sections. 5.3, 5.4, this estimator—henceforth referred to as $\widehat{\nu}_{\text{pl}}$—provides a precise value of ν, with an uncertainty in the range $1 \div 2$ %. Therefore, the purpose of the present Section is to investigate the possibility of using this estimator to improve the assessment of suitable embedding pairs discussed in Chapter 3. In particular, $\widehat{\nu}_{\text{pl}}$ can be used as a replacement for the estimator—henceforth referred to as $\widehat{\nu}_{\text{gauge}}$—provided by Eq. (3.15) and used to build embedding lattice maps. It is worth noting that the two estimators $\widehat{\nu}_{\text{pl}}$ and $\widehat{\nu}_{\text{gauge}}$ are derived from the very same property, namely the Grassberger-Procaccia power-law scaling of correlation integrals at small distances, while differ in *how* the exponent ν corresponding to this scaling is retrieved: the two estimators are thus expected to provide compatible results.

As a test bench, we consider the same Lorenz sequence analyzed in Chapter 3 and in Sec. 5.3 above[3]. An embedding lattice map of estimated correlation dimension is built by relying on the results of Sec. 5.3 as follows. For each embedding choice (m, L), the plateau level $\widehat{\Lambda}_{\text{pl}}$, defined as in Eq. (5.5), is computed according to the procedure of Sec. 5.1 for different values of the starting percentile p: $p = 10^{-4}, 3 \cdot 10^{-4}, 10^{-3}, 3 \cdot 10^{-3}, 10^{-2}, 3 \cdot 10^{-2}, 10^{-1}, 3 \cdot 10^{-1}$. In all these evaluations, the other parameters of the algorithm are set to $N = 500$, $M = 10$, $k \in [450, 500]$, $\delta k = 1$. As a result, for each embedding point (m, L), a set of $(p, \widehat{\Lambda}_{\text{pl}})$ pairs are available, where each value of $\widehat{\Lambda}_{\text{pl}}$ is also characterized by a corresponding uncertainty $s_{\Lambda_{\text{pl}}}$ (see Eq. (5.5)). Thereupon, a linear fit is carried out on the points $(\log(p), \widehat{\Lambda}_{\text{pl}})$ according to Eq. (5.12): the slope of the straight line provides an estimate of $\widehat{\nu}_{\text{pl}}$ for that embedding point.
In analogy with the method of Chapters 3 and 4, all estimates of $\widehat{\nu}_{\text{pl}}$ are displayed by means of an embedding lattice map, which is shown in Fig. 5.9(a). For the sake of comparison, the embedding lattice map of Fig. 3.7(a) is also reported in Fig. 5.9(c). Related histograms of the estimated $\widehat{\nu}$ are shown in Fig. 5.9(b) and Fig. 5.9(d). For some embedding points, marked by black triangles within the map and corresponding to large embedding windows, the $\widehat{\nu}_{\text{pl}}$ estimate is unavailable. This unavailability is not due to a skimming procedure, such as the one described in Sec. 3.2.2 that yields the black "+" marks in Fig. 5.9(c): black triangles mark embedding points for which the algorithm did not deliver, for some values of p, the evaluation of $\widehat{\Lambda}_{\text{pl}}$ within a maximum computational time limit of 6 hours (see the discussion on computational cost below).

To some extent, the results of the two estimators—namely $\widehat{\nu}_{\text{pl}}$ and $\widehat{\nu}_{\text{gauge}}$—reported in Fig. 5.9(a) and

[3]The sequence is synthesized by integrating a Lorenz system (see Eq. (1.2)) with parameters $\sigma = 10$, $r = 28$, $b = 8/3$, with integration time step $\delta t = 0.03$, and by sampling its x-coordinate with sampling period $T = \delta t$. The sequence length is $\ell = 10^5$ points

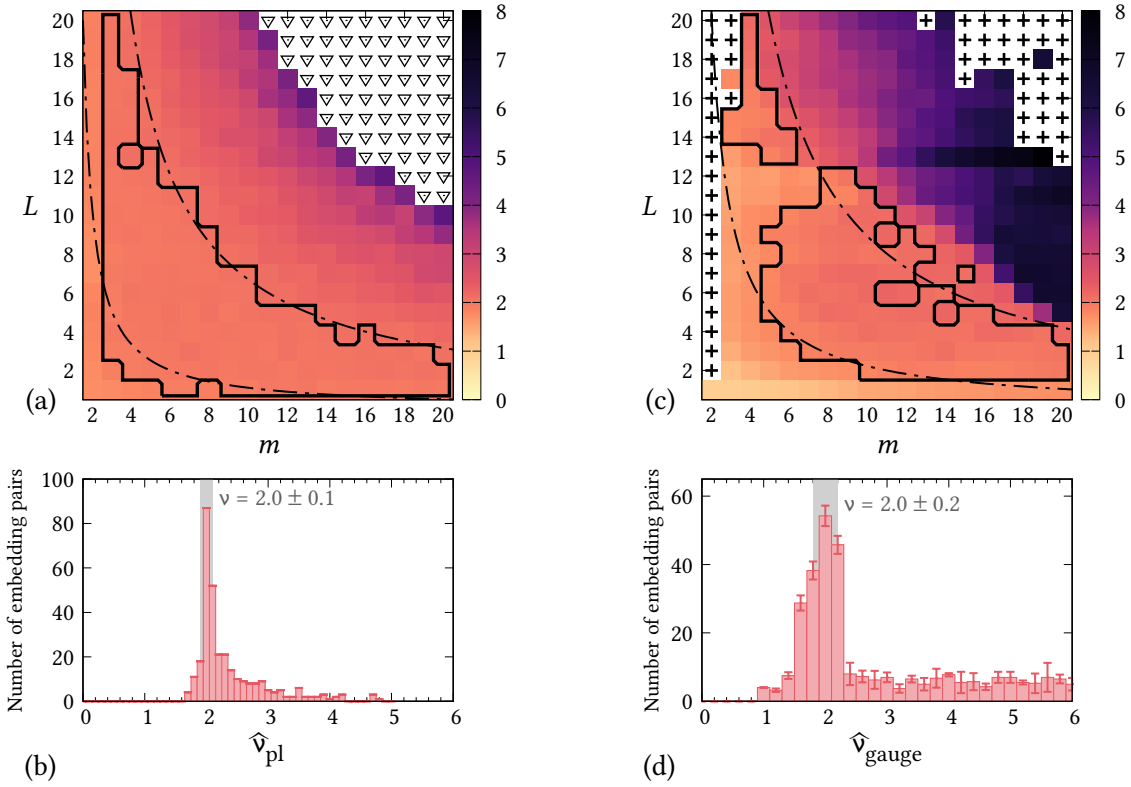


FIGURE 5.9 – *(a) Embedding lattice map of estimated $\widehat{\nu}_{pl}$ according to the procedure discussed in the present Section. The black solid line identifies the set of suitable embedding points corresponding to the peak at $\widehat{\nu}_{pl} = 2.0 \pm 0.1$ within the histogram in (b). The black, dash-dotted lines correspond to the hyperbolae approximately bounding the set of suitable embedding points and defined by $L = 10/(m-1)$ and $L = 60/(m-1)$. Black triangles mark embedding points for which no evaluation is available because the algorithm exceeded the maximum computational time limit. (b) Histogram of the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}_{pl}$ falls within 0.1-wide bins. No errorbars are provided: because of the high computational cost, only one evaluation of the map in (a) was carried out. The range $\widehat{\nu}_{pl} = 2.0 \pm 0.1$, which includes embedding points within the set bounded by a black contour in (a), is highlighted by a gray shaded area. (c) Embedding lattice map of estimated $\widehat{\nu}_{gauge}$ built according to the method discussed in Chapter 3. Black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. The black solid line identifies set of suitable embedding points corresponding to the peak at $\widehat{\nu} = 2.0 \pm 0.2$ within the histogram in (d). The black, dash-dotted lines correspond to the hyperbolae approximately bounding the the set of suitable embedding points and defined by $L = 20/(m-1)$ and $L = 80/(m-1)$. (d) Histogram of the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}_{gauge}$ falls within 0.2-wide bins. Bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the map in (c). The range $\widehat{\nu}_{pl} = 2.0 \pm 0.2$, which includes embedding points within the set bounded by a black contour in (c), is highlighted by a gray shaded area.*

Fig. 5.9(c), respectively, are compatible. In both cases, a region of suitable embedding points bounded by two hyperbolae is found, corresponding to a peak in the related $\widehat{\nu}$ histogram. The related value of ν,

which is assessed as $\nu = 2.0 \pm 0.1$ for the $\widehat{\nu}_{\text{pl}}$ estimator, is compatible with the former figure of 2.0 ± 0.2 provided by the $\widehat{\nu}_{\text{gauge}}$ estimator. However, the $\widehat{\nu}_{\text{pl}}$ estimator is superior in precision and accuracy: the corresponding peak in the histogram in narrower compared to the one for the $\widehat{\nu}_{\text{gauge}}$ estimator, and the region of suitable embedding points in the map is flatter.

A relevant difference between the maps is in the *position* of the region of suitable embedding points, or, equivalently, on the estimated irrelevance and redundance times corresponding to the two bounding hyperbolae. While the hyperbolae in the case of the $\widehat{\nu}_{\text{gauge}}$ estimator correspond to $(m-1)L = 80$ and $(m-1)L = 20$, for the $\widehat{\nu}_{\text{pl}}$ estimator the embedding windows defining the bounding curves are $(m-1)L = 60$ and $(m-1)L = 10$. In other words, the method relying on $\widehat{\nu}_{\text{pl}}$ requires a smaller embedding window to achieve a good estimate of the correlation dimension. A comparison between the regions highlighted by a black contour in Fig. 5.9(a) and Fig. 5.9(c) is shown in Fig. 5.10. The region of suitable embedding points

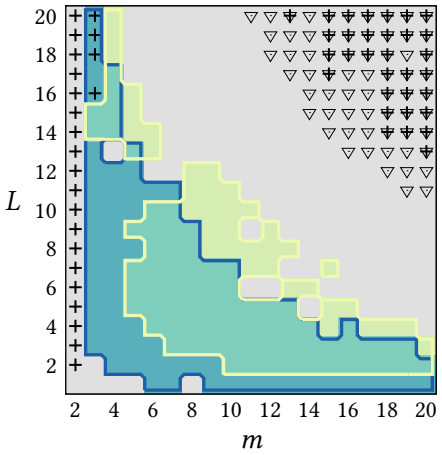


FIGURE 5.10 – *Comparison of the regions of suitable embedding points provided by the two estimators $\widehat{\nu}_{pl}$ and $\widehat{\nu}_{gauge}$. The blue shaded area corresponds to the embedding lattice region highlighted by a black solid line in Fig. 5.9(a), i.e. the region of suitable embedding choices according to the $\widehat{\nu}_{pl}$ estimator. The yellow shaded area corresponds to the analogous region in Fig. 5.9(c) yielded by the $\widehat{\nu}_{gauge}$ estimator. The green shaded area corresponds to the overlap between the blue and yellow regions. Black triangles and "+" marks have the same meaning as in Fig. 5.9.*

of the $\widehat{\nu}_{\text{pl}}$ estimator (blue) and the same region for the $\widehat{\nu}_{\text{gauge}}$ estimator (yellow) partially overlap (green). More in detail, the overlapping set corresponds to 55% and 65% of the region of suitable embedding points for $\widehat{\nu}_{\text{pl}}$ and $\widehat{\nu}_{\text{gauge}}$, respectively. Interestingly, the region of embedding points for which the $\widehat{\nu}_{\text{pl}}$ estimator cannot be evaluated because of a exceedingly long computational time (black triangles) partially overlaps with the region corresponding to black "+" marks, i.e. the set of embedding points that are deemed to be incompatible with the null hypothesis of a finite-dimensional system by the method discussed in Chapter 3. The reason behind this failure, which concerns both methods, might lay in the excessive spreading of embedding vectors within the embedding space as a consequence of a too large embedding window. If there are too few embedding vectors close to each other, the small-distance features of the system are poorly sampled. Therefore, the method discussed in Chapter 3 will fail in recognizing the system as finite dimensional, as it occurs for $(m-1)L \gtrsim 250$ (top-right of the map). Under the same circumstances, the method discussed in this chapter will be hampered as well: for small

enough values of the shell radius r, the search for embedding vectors such that $\|\mathbf{Y}_i - \mathbf{Y}_j\| < r$ will take a very long time, as it happens for $(m-1)L \gtrsim 200$. To summarize, when large embedding windows are examined, and thus the small-distance regime is poorly sampled, the method discussed in Chapter 3 fails in identifying the system as finite dimensional; on the other hand, the method discussed in this chapter fails—at least for small p—to deliver an estimate of the correlation dimension in a reasonable time.

As far as the points at $m = 2$ and marked with a "+" are concerned, the method of Chapter 3 provides a correct classification by marking these embedding points as incompatible with a finite-dimensional system. Indeed, Lorenz system's trajectories reconstructed with $m = 2$ will have crossings, which are forbidden by determinism, and thus the system *cannot* be embedded with $m = 2$. On the contrary, the $\widehat{\nu}_{\mathrm{pl}}$ estimator provides a finite dimension $\lesssim 2$ for the points with $m = 2$. This outcome might be mistakenly mixed up with the histogram's peak at $\widehat{\nu}_{\mathrm{pl}} \simeq 2$, and it does not rule out, in principle, those unsuitable embedding choices. Consequently, a correct use of the $\widehat{\nu}_{\mathrm{pl}}$ estimator still requires the preliminary skimming of embedding points discussed in Sec. 3.2.1 and Sec. 3.2.2.

A relevant issue that hinders the use of $\widehat{\nu}_{\mathrm{pl}}$ concerns its computational cost. The computational time required by the two methods to deliver a map of correlation dimension on the embedding lattice was evaluated by running the related algorithms on a workstation hosting an Intel®Xeon®E5-2650 CPU, 48 GB RAM, running Ubuntu Linux 16.04 LTS 64-bit. The code was written in C++ [182] and parallelized as much as possible by means of the OpenMP library [183]. Random number generation, as well as the evaluation of Gamma functions, were carried out by means of the GNU Scientific Library [184]. Executable files were assembled via the GNU Compiler Collection (gcc), version 5.4.0 [185]; the optimization flag was set to `-O3`. Execution time was assessed via the `time` utility provided by the Linux *bash* shell.

The method of Chapter 3 takes 10.8 ± 0.1 minutes (mean and standard deviation over 4 repetitions) to provide a single map of the kind shown in Fig. 5.9(c). On the other hand, the method discussed in this chapter takes 171 hours (only one calculation was performed) to provide a single map of $\widehat{\nu}_{\mathrm{pl}}$ like the one shown in Fig. 5.9(a). This estimate does not include the points marked by the black triangles, for which the program fails to provide a $\widehat{\Lambda}_{\mathrm{pl}}$ value (for some values of p) within 6 hours. Therefore, there is a significant difference in computational cost between the two methods, which can be quantified in about three orders of magnitude in execution time.

Finally, it is worth highlighting another issue related to the $\widehat{\nu}_{\mathrm{pl}}$ estimator. To properly assess the plateau value, the value of $k_{\min}$ has to be properly chosen: this choice requires at least a preliminary inspection of the $\widehat{\Lambda}(k)$ curve, and thus an additional stage prior to the actual execution of the algorithm.

In conclusion, the correlation dimension estimator $\widehat{\nu}_{\mathrm{pl}}$, which relies on the asymptotic behaviour of the time-dependent divergence exponent, is more precise than the estimator $\widehat{\nu}_{\mathrm{gauge}}$ introduced in

Chapter 3. However, its implementation is far more demanding from a computational point of view. Therefore, $\widehat{\nu}_{\mathrm{pl}}$ might be a good alternative to $\widehat{\nu}_{\mathrm{gauge}}$ for the third stage of the method for the identification of suitable embedding parameters, to be used when a preliminary (and inexpensive) analysis based on $\widehat{\nu}_{\mathrm{gauge}}$ already gave positive results.

6. Chasing chaos in experimental recordings

The goal of the present chapter is to discuss the issue of searching chaos in experimental recordings. As introduced in Chapter 1, the detection of chaotic behaviour in nature is a pivotal problem in nonlinear science. In many research fields, most notably in neuroscience, the identification of chaos is still an open point, and conclusive evidence of existence or absence of chaos is lacking. The difficulties arising in this quest concern, for example, low signal-to-noise ratio, insufficient length of analyzed sequences, nonstationarity, as well as the inadequacy of acquisition devices. Here, the possible use of the approach described in the previous chapters as a chaos detection method is discussed.

The discussion relies on two examples: in the first example, I will analyze signals generated by an implementation of Chua's circuit, while in the second one I will discuss the case of brain activity recorded via magnetoencephalography (MEG). Electronic circuits and the human brain can be regarded as opposite within the spectrum of nonlinear—and potentially chaotic—systems. Systems belonging to the former class are easy to tune and control, and the effective degrees of freedom involved in the dynamics are straightforwardly identified with electric potentials and currents. The situation is reversed in the case of the human brain: no direct control or tuning of the system is possible, and the observability of its "degrees of freedom" is technologically limited.

Despite the two dynamical systems seem apparently unrelated, networks of coupled Chua's circuits were studied as models to simulate phenomenological aspects of the human brain. More in detail, an array of Chua's circuits was claimed to exhibit transitions between dynamical regimes similar to those observed through EEG in epileptic seizures [186]. More recently, experimental investigations were carried out on a small network of Chua oscillators to demonstrate the capabilities of such configuration in implementing information processing functions [187]. Both these examples stem from the idea that "mesoscopic" complex behavior (e.g. in the human brain) can be qualitatively modelled by means of networks of interacting dynamical units, also called "cellular neural networks" [188, 189].
It is worth remarking that, while Chua's circuit was chosen as example in the present chapter, a possible future investigation—beyond the present work—could concern the study of models of neuron dynamics such as the FitzHugh-Nagumo oscillator [190], for which a more direct relationship with the human

brain can be drawn. The FitzHugh-Nagumo system, which can be driven into a chaotic regime by applying an external periodic forcing, was extensively studied in the framework of synchronization: the study of coupled FitzHugh-Nagumo oscillators can provide information on the dynamics of ensembles of biological neurons [191, 192, 193].

Chua's circuit is here implemented by means of standard, off-the-shelf electronic components, as described in Sec. 6.1. Assembly of the circuit and data acquisition were carried out within the NSE laboratory I am working in. On the other hand, MEG recordings were extracted from the public database of the Human Connectome Project [194, 195], as in two works from our research group in which brain networks were investigated [196, 197]. Data collected from the circuit and from MEG recordings were analyzed by means of the method discussed in Chapter 3. Results concerning the electronic system (Sec. 6.1) are substantially different from those obtained for brain signals (Sec. 6.2). In particular, while the outcomes of the analysis performed on the circuit's recorded potentials provide evidence for an underlying chaotic dynamics, such evidence is lacking in the case of brain activity. Some comments concerning this difference, as well as some observations on the detection of chaos in experimental recordings, are reported at the end of Sec. 6.2.

6.1 An implementation of Chua's circuit

Chua's circuit is an autonomous oscillating circuit [48] that, in its standard form shown in Fig. 6.1, relies on an inductor and two capacitors to generate oscillations, and on an active component, "Chua's diode", to provide nonlinearity by virtue of the piecewise-linear I-V characteristic of the diode. The

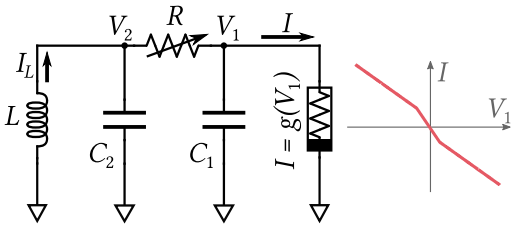


FIGURE 6.1 – *Standard form of Chua's circuit. The rightmost element, namely Chua's diode, is characterized by the piecewise-linear I-V curve* $I = g(V_1)$ *sketched in the graph: Chua's diode is the circuit's sole nonlinear element.*

system's equations are readily obtained through standard circuit analysis [47]:

$$C_1 \frac{dV_1}{dt} = \frac{1}{R}(V_2 - V_1) - g(V_1), \tag{6.1a}$$

$$C_2 \frac{dV_2}{dt} = \frac{1}{R}(V_1 - V_2) + I_L, \tag{6.1b}$$

$$L \frac{dI_L}{dt} = -V_2, \tag{6.1c}$$

where $g(V_1)$ is the current flowing through Chua's diode when a potential difference V_1 is applied across it. The term $g(V_1)$ is the only nonlinearity of the system. In the present investigation, Chua's diode is practically implemented by means of two negative impedance converters [198]. Moreover, the discrete inductor L is replaced by a *gyrator* circuit. The gyrator relies on two operational amplifiers to provide an inductor-like impedance $-i\omega L$, with $L = 20\,\text{mH}$, out of a $C = 20\,\text{nF}$ capacitor, having impedance $i/\omega C$. As a result, in the present example, Chua's circuit is built out of four operational amplifiers, three capacitors and a handful of resistors: the complete circuit diagram, which also reports the nominal values of components, is shown in Fig. 6.2. Upon delivering suitable power supplies to the

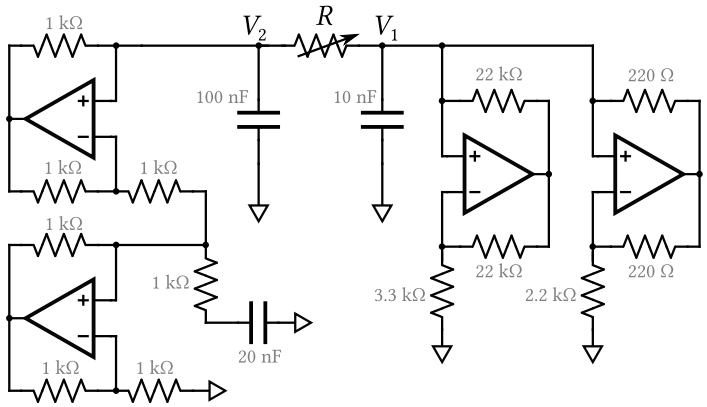


FIGURE 6.2 – *Implementation of Chua's circuit used in the present example. The leftmost network containing two op-amps corresponds to a gyrator circuit that provides an equivalent impedance $-i\omega L$ with $L = 20$ mH. The rightmost network containing two op-amps corresponds to Chua's diode. R is given by a variable resistor.*

operational amplifiers, the circuit oscillates autonomously. Adjusting the resistor R allows to achieve different dynamical regimes. Figure 6.3 shows sequences of the potentials V_1 and V_2 recorded in two noticeably different regimes, namely a periodic limit cycle—Fig. 6.3(a)—and an "irregular" aperiodic evolution—Fig. 6.3(b). While it is well known that the latter regime corresponds to chaos, for the time being such knowledge is neglected. It is worth noting that the values of R giving rise to the two regimes in Fig. 6.3 are approximately 1.8 kΩ and 1.7 kΩ for the periodic and the aperiodic evolution, respectively. The transition between the two different regimes occurs due to a variation of $\sim 5\%$ of the resistance R. Moreover, reducing R below approximately 1.6 kΩ yields again a limit cycle: in other words, the value of R is critical in determining the dynamical regime, and an irregular evolution it is achieved only for a range of R values between 1.6 kΩ and 1.8 kΩ.

For each of the two dynamical regimes shown in Fig. 6.3, sequences of both potentials V_1 and V_2 are recorded for 0.5 s (10^5 samples, $T = 5\,\mu\text{s}$). In the following, for the sake of this example, no knowledge is assumed about the two dynamical regimes, and the sequences corresponding to V_1 and V_2 are independently analyzed, thus neglecting the information provided by Eq. (6.1).

The method discussed in Chapter 3 is applied to the four sequences, thus generating the embedding lattice maps shown in Fig. 6.4 and in Fig. 6.5 for the sequences in Fig. 6.3(a) and Fig. 6.3(b), respectively. The outcomes of the analysis concerning the first dynamical regime, which are shown in Fig. 6.4, are

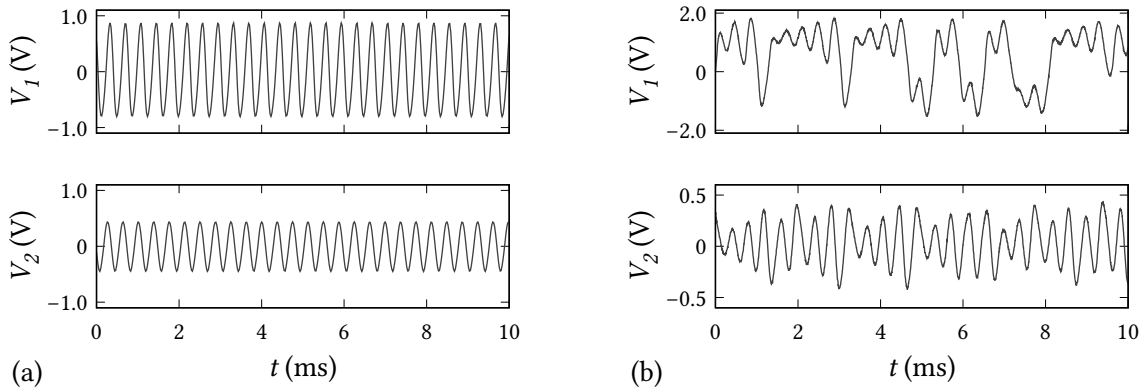
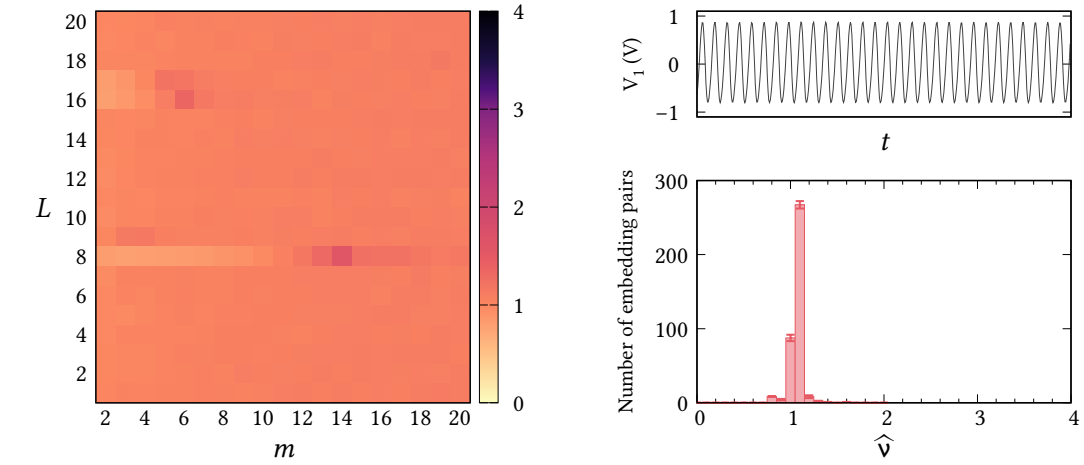


FIGURE 6.3 – *Recorded voltages V_1 and V_2 for two different values of R. (a) $R \simeq 1.8$ kΩ, the system oscillates periodically. (b) $R \simeq 1.7$ kΩ, the system's oscillations are aperiodic.*

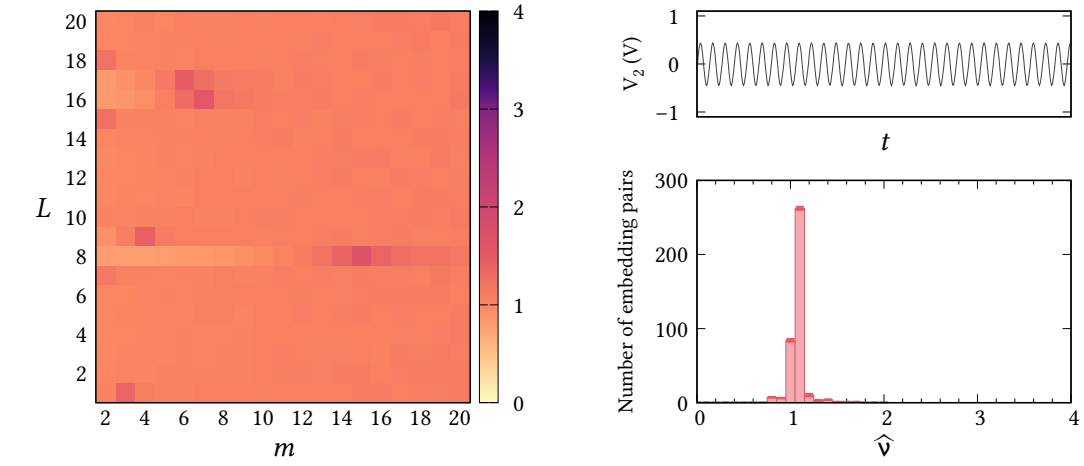
straightforwardly interpreted. The estimated $\widehat{\nu}$ attains, independently of the embedding parameters, a value close to unity, compatibly with a periodic limit cycle. This outcome simply confirms, on the one hand, what we could have expected for a periodic signal, and, on the other hand, it validates the ability of the method to correctly identify periodic signals.

Figure 6.5 shows the results concerning the "irregular" dynamical regime. To build the map in Fig. 6.5(a) for the potential V_1, the corresponding sequence was first downsampled by selecting one sample out of three (i.e. the new sampling period is $T = 15\,\mu s$). The map obtained without downsampling exhibits the same hyperbolic-like transition from $\widehat{\nu} \approx 1$ to $\widehat{\nu} \approx 4$, although shifted towards the top-right corner, at $m, L \approx 17$. The aim of downsampling is to "center" this possibly interesting transition region within the embedding lattice. Equivalently, one could have extended the embedding lattice to $m, L \approx 35$.

Compared to maps in Fig. 6.4, embedding lattice maps in Fig. 6.5 show more complex features. First of all, both maps exhibit a hyperbolic-like profile of $\widehat{\nu}$, as expected in the case of chaotic sources. Interestingly, in both maps, the embedding points corresponding to $m = 2$ are ruled out by the analysis because of their incompatibility with the null hypothesis $\mathcal{H}_{0;\mathrm{CFD}}$ of an underlying finite-dimensional source: the same result was found for the Lorenz sequence (see Fig. 3.7(a)), the Ueda sequence (Fig. 4.3(c)), the Rössler sequence (Fig. 4.6(a)) and the sequence of laser intensity (Fig. 4.7(a)). In the case of V_2, the $\widehat{\nu}$ histogram exhibits a peak at $\widehat{\nu} = 2.0 \pm 0.2$, and the corresponding region within the embedding lattice is approximately bounded by two hyperbolae. These results support the hypothesis of a chaotic system generating the sequence V_2. However, the same is not true for the sequence V_1. In this case, the peak within the $\widehat{\nu}$ histogram corresponds to $\widehat{\nu} \simeq 1$, which would rather suggest a periodic regime. This hypothesis, however, does not agree with the fact that embedding points with $m = 2$ are ruled out by the algorithm, as well as with the fact that $\widehat{\nu} > 1$ are estimated for $m \gtrsim 10$, $L \gtrsim 10$.



(a)



(b)

FIGURE 6.4 – *Embedding lattice maps, $\widehat{\nu}$ histograms, and excerpts of the input sequences corresponding to the signals V_1 (a) and V_2 (b) generated by Chua's circuits upon setting $R \simeq 1.8\,k\Omega$. Histograms are built with bin width* 0.1*; bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the respective maps. The histogram peaks at $\widehat{\nu} = 1.1 \pm 0.1$ and the related flatness of the map reflect the fact that the system's dynamical regime is a (one-dimensional) limit cycle.*

More realistically, the $\widehat{\nu}$ estimator of Eq. (3.15) used to build the maps in Fig. 6.5 might fail to provide reliable values of the system's correlation dimension out of the sequence V_1. To check if this is the case, the improved estimator introduced in Chapter 5 is used. As discussed in Sec. 5.5, although computationally demanding, it is worth evaluating ν by means of the improved estimator to obtain a more reliable map, given that we already have some evidence hinting at an underlying chaotic dynamics.

Similarly to the example of Sec. 5.5, an embedding lattice map of estimated correlation dimension

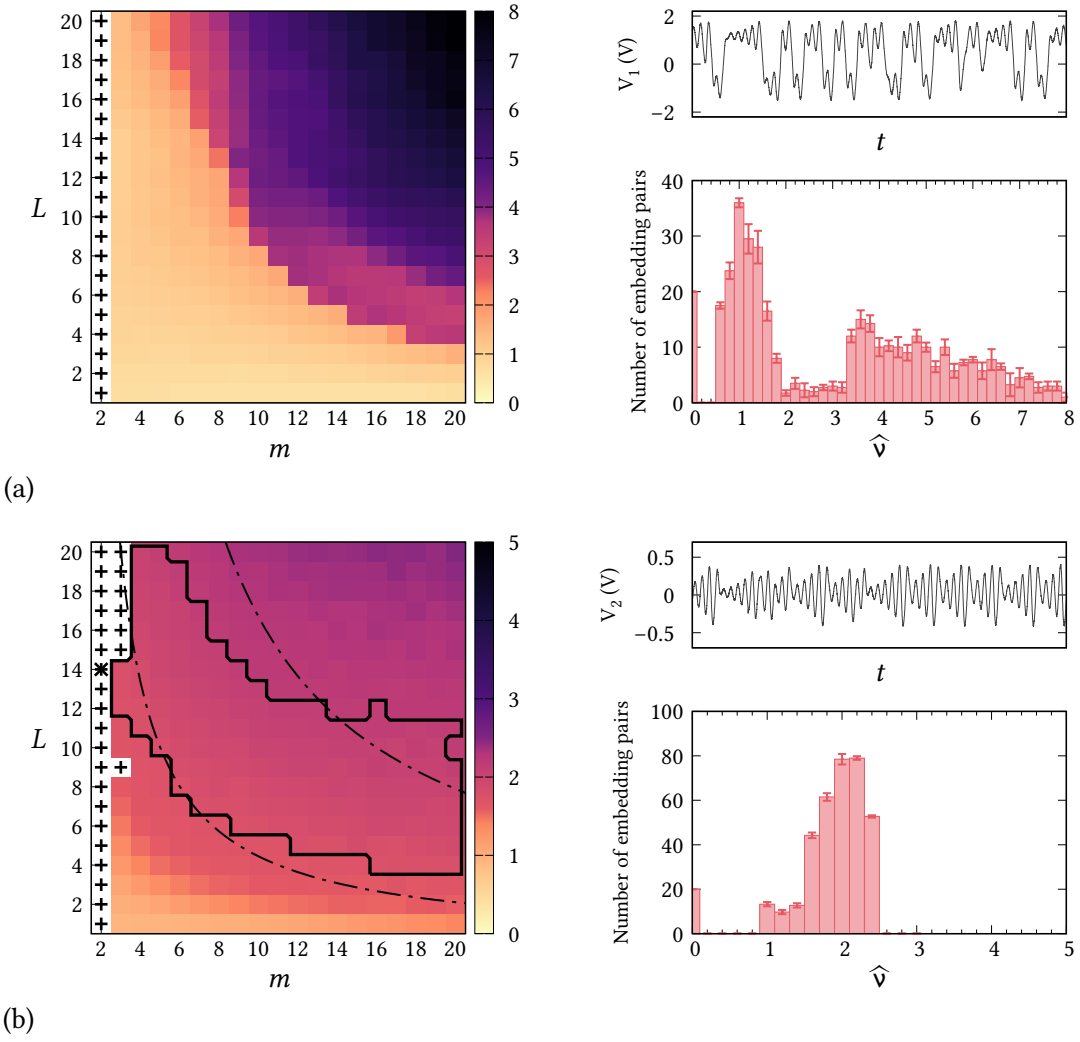


FIGURE 6.5 – *Embedding lattice maps, $\widehat{\nu}$ histograms, and excerpts of the input sequences corresponding to the signals V_1 (a) and V_2 (b) generated by Chua's circuits upon setting $R \simeq 1.7\ k\Omega$. Within maps, black crosses (×) correspond to embedding points compatible with the null hypothesis $\mathcal{H}_{0;GWN}$; black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. Histograms are built with bin width 0.2; bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the respective maps. In (b), the black solid line encloses the embedding region corresponding to the peak at $\widehat{\nu} = 2.0 \pm 0.2$ highlighted by a gray shaded area within the related histogram. The black dash-dotted lines correspond to the hyperbolae approximately bounding the region of suitable embedding points and defined by $L = 40/(m-1)$ and $L = 150/(m-1)$.*

$\widehat{\nu}_{\mathrm{pl}}$ is built by evaluating the dependence of the asymptotic divergence exponent on the starting percentile p, according to the results of Sec. 5.3. The analysis discussed in Chapter 5, and in particular the evaluation of sample plateau levels $\widehat{\Lambda}_{\mathrm{pl}}$ described in Sec. 5.1, is carried out on the two input sequences (V_1 and V_2) analyzed in Fig. 6.5. The plateau level $\widehat{\Lambda}_{\mathrm{pl}}$, defined in Eq. (5.5), is computed for different

values of the starting percentile p: $3 \cdot 10^{-3}$, 10^{-2}, $3 \cdot 10^{-2}$, 10^{-1}, $3 \cdot 10^{-1}$. In all these evaluations, the other parameters of the algorithm are set to $N = 500$, $M = 10$. For the V_1 sequence, $k \in [450, 500]$, $\delta k = 1$; for the V_2 sequence, $k \in [1800, 2000]$, $\delta k = 5$. As it was done in Sec. 5.5, for each embedding choice (m, L), a linear fit carried out on the points $(\log(p), \widehat{\Lambda}_{\mathrm{pl}})$ according to Eq. (5.12) provides an estimate of $\widehat{\nu}_{\mathrm{pl}}$ for that (m, L).

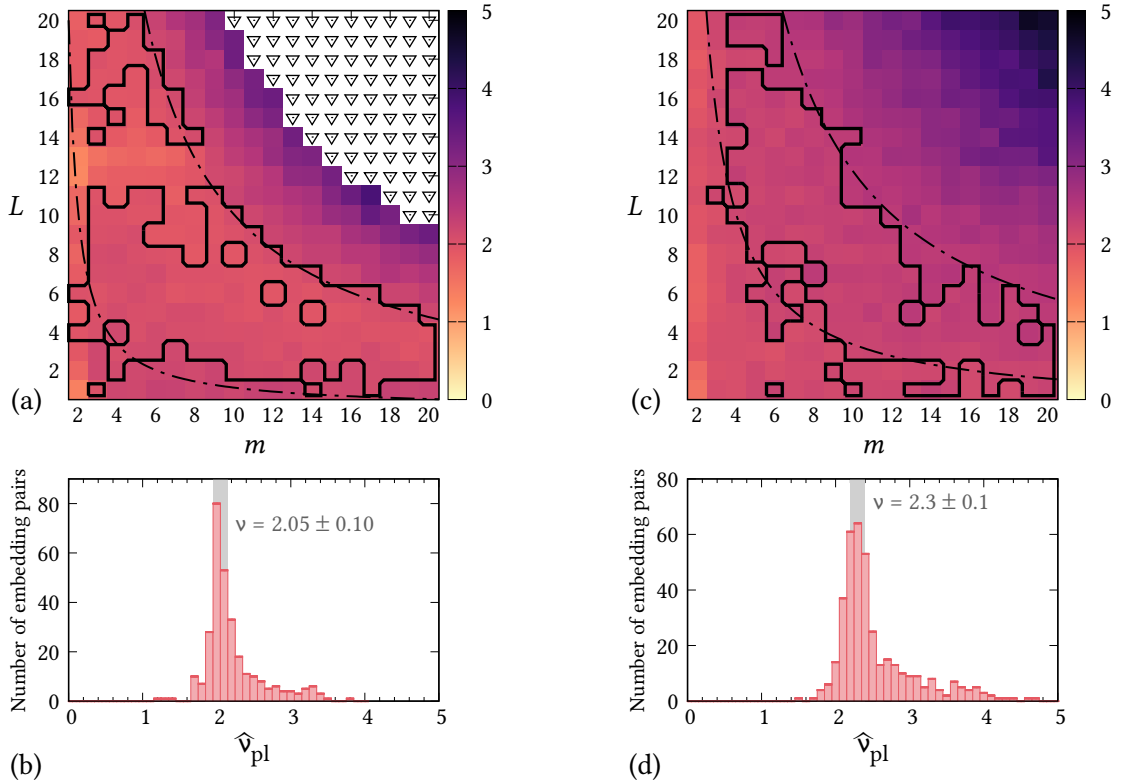


FIGURE 6.6 – *Embedding lattice maps—and related histograms—of correlation dimension $\widehat{\nu}_{pl}$ estimated by means of the improved estimator discussed in Chapter 5. (a,b) V_1 sequence. (c,d) V_2 sequence. Within maps, black solid lines identify the sets of suitable embedding points corresponding to the peaks at $\widehat{\nu}_{pl} = 2.3 \pm 0.1$ within the related histograms. Black, dash-dotted lines correspond to the hyperbolae approximately bounding the sets of suitable embedding points: hyperbolae are defined by $L = 10/(m-1)$ and $L = 90/(m-1)$ in (a), while $L = 30/(m-1)$ and $L = 110/(m-1)$ in (c). Black triangles mark embedding points for which no evaluation is available because the algorithm exceeded the maximum computational time limit. Histograms (b,d) report the number of embedding pairs whose estimated correlation dimension $\widehat{\nu}_{pl}$ falls within 0.1-wide bins. The ranges $\widehat{\nu}_{pl} = 2.05 \pm 0.10$ and $\widehat{\nu}_{pl} = 2.3 \pm 0.1$, which include embedding points within the sets bounded by black contours in (a,c), are highlighted by a gray shaded area in both histograms.*

The embedding lattice maps of the estimated correlation dimension $\widehat{\nu}_{\mathrm{pl}}$ obtained through the improved estimator are shown in Fig. 6.6(a) and Fig. 6.6(c) for the V_1 and V_2 sequences, respectively, while the related $\widehat{\nu}_{\mathrm{pl}}$ histograms are shown in Fig. 6.6(b) and Fig. 6.6(d). Both embedding lattice maps exhibit

a region of suitable embedding points bounded by a pair of hyperbolae. The related estimates of correlation dimension are in reasonable agreement and equal to $\widehat{\nu}_{\mathrm{pl}} = 2.05 \pm 0.10$ and $\widehat{\nu}_{\mathrm{pl}} = 2.3 \pm 0.1$. In conclusion, both sequences V_1 and V_2 provide reliable evidence of an underlying chaotic source. The related correlation dimension is estimated, by averaging the two values for V_1 and V_2, as $\nu = 2.18 \pm 0.14$, in agreement with values reported in literature [199].

6.2 Magnetoencephalographic recordings

In the present Section, a very different system, namely a human brain, is investigated. As in the previous Section, the method presented in Chapter 3 is applied to sequences generated by the system with the purpose of characterizing it and, possibly, find evidence of an underlying chaotic dynamics. As discussed in Chapter 1, the search for chaotic regimes within human brain activity yielded conflicting results. While the present Section is definitely not intended to settle this issue, its aim is to highlight differences with respect to the case of a small circuit. In particular, while chaos can be robustly identified out of a tabletop setup such as the one in Fig. 6.2, attempting a similar identification out of brain activity is shown to be far less rewarding.

Data under investigation are time series of brain activity reconstructed out of magnetoencephalographic (MEG) recordings. MEG data considered in the following are a portion of the dataset employed by our research group to investigate brain connectivity [196] and its dependence on geometric distance [197]. The dataset was downloaded from the public database of the Human Connectome Project (HCP)[1] [194, 195]. A single MEG resting state acquisition—corresponding to one randomly chosen subject out of the 20 used in Ref. [197]—is selected.

MEG recordings pick up brain activity by measuring magnetic fields generated by neural currents. Detectable fields, which are of the order of $10^{-12} \div 10^{-14}$ T, are the result of thousands of neurons firing [99]. MEG devices record these magnetic fields by means of $200 \div 300$ superconducting sensors: field sources (currents) are reconstructed at various brain locations out of the recorded fields by estimating an approximate solution to the (unsolvable) inverse problem [100]. Details of this preprocessing stage in the case of the present dataset can be found in Ref. [197] (section "Dataset and preprocessing" therein). The outcome of source reconstruction are sequences of current dipole amplitudes comprising 75000 elements each and sampled with sampling period $T = 4$ ms (sequence duration is 5 minutes). Each sequence $s_{(i)}$ is associated to one location (or *source*) within the human brain. Four among the 72 brain locations used in the study of Ref. [197] are considered here: the corresponding anatomical regions are shown in Fig. 6.7. These four locations are selected as representative of four different areas of the brain concerning different tasks. According to a standard atlas [200], these four regions are

[1] Ethical approval and consent needed for study and dissemination are provided by the HCP (see also Ref. [197]).

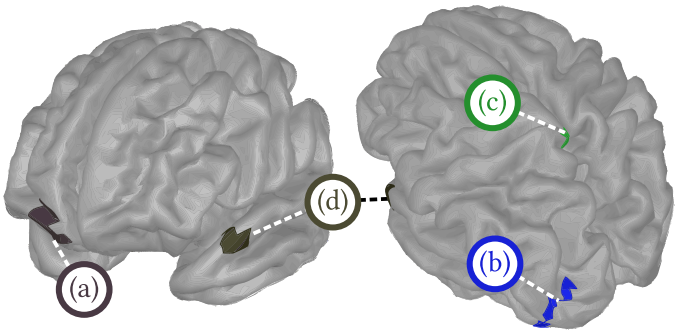


FIGURE 6.7 – *Anatomical position of the four brain regions considered in the present example. Colors correspond to those used in a standard atlas [200] and are representative of the region's functional group: (a) and (d) are "task-negative" regions, (b) is part of the visual cortex, (c) is part of the motor/somatosensory cortex.*

conventionally classified in terms of the related functions: (a) is "task-negative", i.e. activates during resting-state; (b) is part of the visual cortex; (c) is part of the motor/somatosensory cortex; (d) is again task-negative. The set is thus supposed to cover several possible dynamical regimes within the brain.

The method of Chapter 3 is first applied to the whole sequences $s_{(a)}, \dots s_{(d)}$ corresponding to 5 minutes of resting-state brain activity reconstructed at the four locations (a)-(d). The results of this analysis are shown in Fig. 6.8, where embedding lattice maps are reported alongside the corresponding histograms of $\widehat{\nu}$ and the plots of the sequences.

FIGURE 6.8 – *(Figure is displayed in a dedicated page below). Embedding lattice maps, $\widehat{\nu}$ histograms, and input sequences corresponding to brain activity reconstructed out of MEG resting-state recordings of 5 minutes length. Within maps, black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. Histograms are built with bin width 0.05; bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the respective maps. Each panel (a)-(d) corresponds to the sequence associated to the brain region marked with the same letter in Fig. 6.7. Sequence plots report the whole, 5-minutes-long sequences.*

The embedding lattice maps of Fig. 6.8 do not provide evidence of an underlying chaotic system. Despite the presence of peaks within the $\widehat{\nu}$ histograms, the corresponding embedding points do not make up a region of the embedding lattice delimited by two hyperbolae. Indeed, for those embedding points that are compatible with $\mathcal{H}_{0;\mathrm{CFD}}$, $\widehat{\nu}$ appears essentially embedding-independent. Moreover, regions (a) and (b) provide an estimated $\widehat{\nu}$ lower than 1, which is meaningless. In the case of region (c), the histogram exhibits a peak at $\widehat{\nu} \approx 2.6$; however, for embedding choices having $m \lesssim 10$, the sequence appears to be incompatible with a finite-dimensional source, independently of L. This result is in contradiction with the fact that $m > D$ is sufficient to provide estimates of a system's correlation dimension [126]. A similar argument applies to the map corresponding to region (d), which would apparently yield $\widehat{\nu} \approx 1.5$.

A possible interpretation for the maps in Fig. 6.8(c) and Fig. 6.8(d) relies on the model of an almost-periodic signal discussed in Sec. 4.2: the model would explain both the values of $\widehat{\nu}$, in terms of an effective signal-to-noise ratio, as well as the rectangular shape of the corresponding regions of the em-

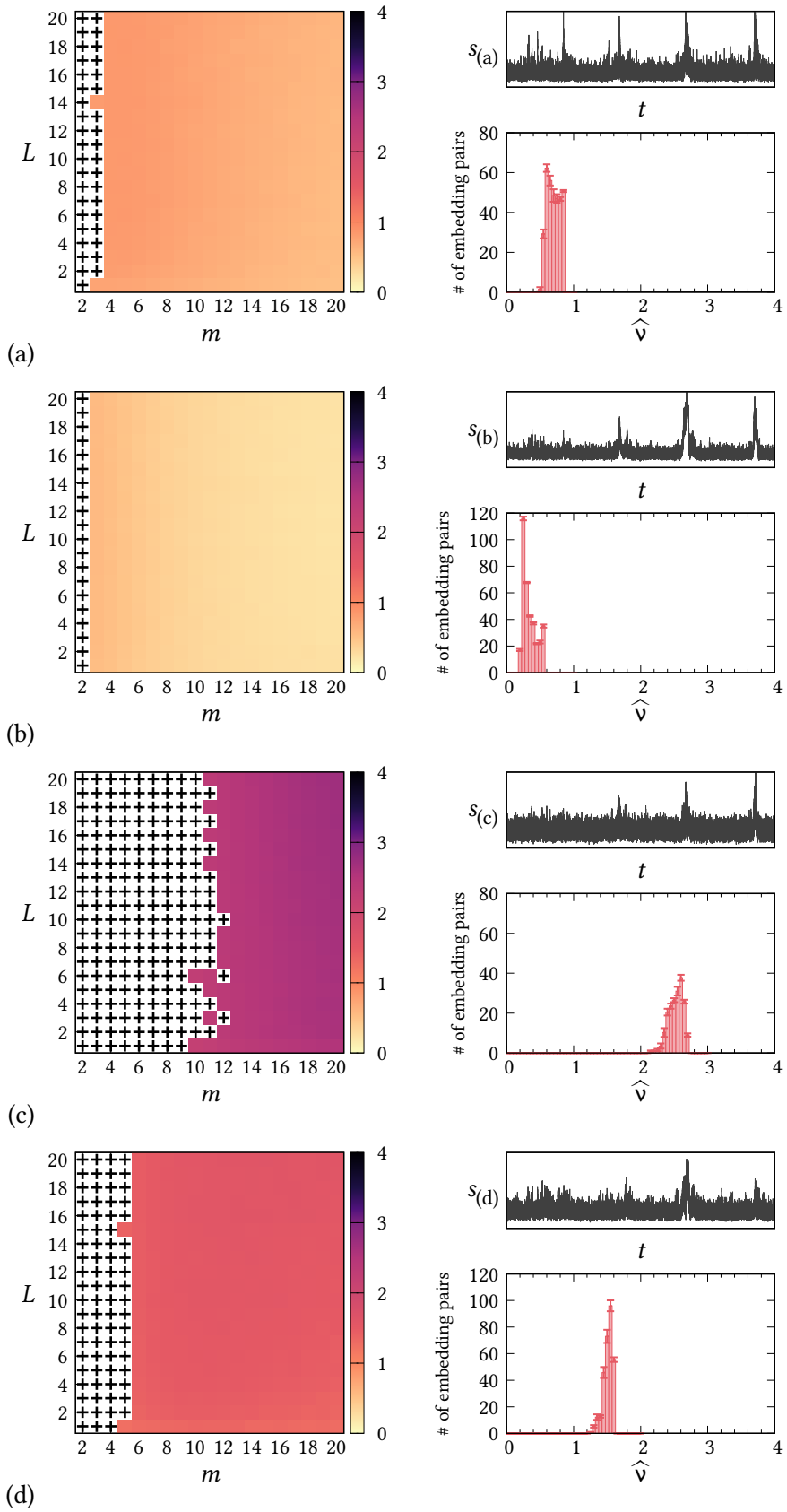


FIGURE 6.8 – *(Full caption reported in the main text). Each panel (a)-(d) corresponds to the sequence associated to the brain region marked with the same letter in Fig. 6.7.*

bedding lattice. However, this interpretation has to be discarded: the sequences associated to the brain regions (c) and (d) do not exhibit any significant peak in their spectra, and thus cannot be modelled as almost-periodic signals.

Another possible explanation for the results shown in Fig. 6.8 is that those maps exhibit spurious estimates of ν because the corresponding sequences are nonstationary. Indeed, by looking at the sequences plots within the figure, one can identify several "bumps" of activity that are well distinguished from the lower amplitude of most remaining parts of the recordings. These events, which are more clearly visible in Fig. 6.9, might be due to artifacts (e.g. head movements), or related to a strong sporadic increase in the activity of the related brain region. As an attempt to circumvent this problem, 1 minute out of 5 is extracted from the sequences so that the "bumps" described above are avoided. Figure 6.9 highlights the time interval corresponding to the extracted segments, which is chosen between $t = 210$ s and $t = 270$ s. The same time interval is selected from all sequences, yielding four segments $s'_{(a)}, \dots s'_{(d)}$.

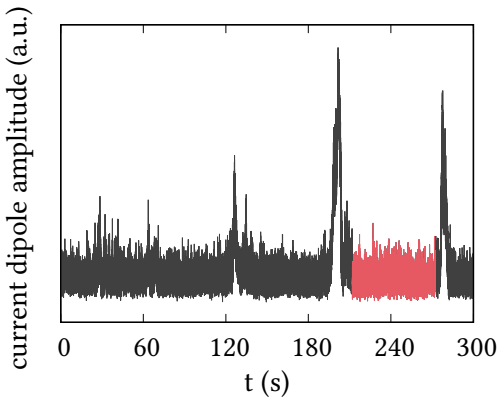


FIGURE 6.9 – *Sequence $s_{(b)}$ corresponding to brain region (b)—see Fig. 6.7. The segment highlighted in red corresponds to 1 minute, from 210 s to 270 s, without any "bump" in activity such as those occurring at $t \approx 200$ s and $t \approx 280$ s.*

The analysis carried out on the whole sequences—and whose results are shown in Fig. 6.8—is then repeated on the four segments. The results of this second analysis are shown in Fig. 6.10, again displaying embedding lattice maps alongside the corresponding histograms of $\widehat{\nu}$ and the plots of the sequences.

In this case, the interpretation of the outcomes is quite straightforward: according to the method described in Chapter 3, the sequence segments considered here are essentially noise. For almost every

FIGURE 6.10 – *(Figure is displayed in a dedicated page below). Embedding lattice maps, $\widehat{\nu}$ histograms, and input sequences corresponding to brain activity reconstructed out of MEG resting-state recordings. The sequences considered here correspond to 1 minute (between 210 s and 270 s) selected out of the 5 minutes available—see Fig. 6.9. Within maps, black crosses (×) correspond to embedding points compatible with the null hypothesis $\mathcal{H}_{0;GWN}$; black "+" marks denote embedding points that are incompatible with the null hypothesis $\mathcal{H}_{0;CFD}$. Histograms are built with bin width 0.25; bin heights and errorbars are sample mean and standard deviation computed by repeating four times the evaluation of the respective maps. Each figure panel (a-d) corresponds to the brain region marked with the same letter in Fig. 6.7. Sequence plots report the whole, 1-minute-long sequence segments.*

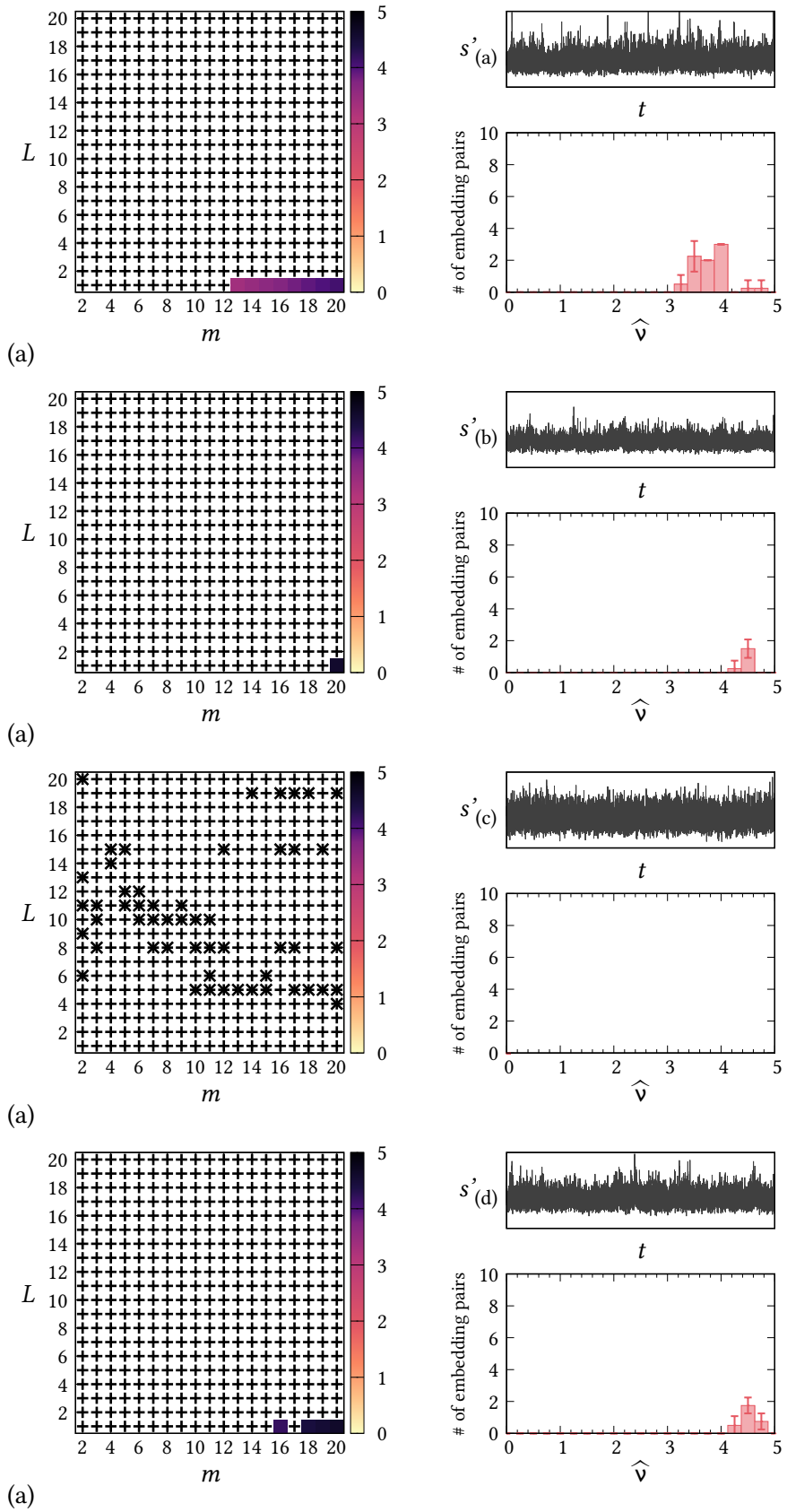


FIGURE 6.10 – *(Full caption reported in the main text). Each figure panel (a-d) corresponds to the brain region marked with the same letter in Fig. 6.7.*

embedding choice, sequences turn out to be incompatible with the null hypothesis $\mathcal{H}_{0;\mathrm{CFD}}$ of an underlying finite-dimensional source. This result supports the hypothesis that the maps of Fig. 6.8 provided a spurious finite correlation dimension due to nonstationary fluctuations within the sequences.

While conclusive evidences of a chaotic source are provided by analyzing sequences recorded out of the implementation of Chua's circuit, in the case of MEG recordings such evidence is absent. As a matter of fact, the 1-minute-long segments analyzed in Fig. 6.10 are compatible with a stochastic source. This outcome might indeed be due to a stochastic evolution of a human brain in resting state. On the other hand, the detection of an underlying deterministic dynamics might be hindered by other factors, most notably the fact that electromagnetic fields generated by a human brain are an *indirect* measure of the generating dynamics. Fields recorded by MEG devices correspond to the superposition of the activity of thousands of neurons: the possible deterministic dynamics of single units or small sets of neurons is washed out by the overwhelming background. This problem also concerns electroencephalography (EEG): large portions of the brain have to be synchronized, thus yielding a small number of effective degrees of freedom out of a huge number of units, in order to produce observable fields that faithfully represent the ongoing dynamics.

Although Takens' theorem tells us that *any* observable signal can be used to study a system, this is not true in experimental practice: the unavoidable presence of noise can hide signals and render them undetectable by analytical algorithms; at the same time, nonstationarity can lead analytical algorithms astray. These issues are not critical if direct access to a system's degrees of freedom is available, as shown for the electronic circuit in Sec. 6.1, where noise is certainly present but does not prevent the identification of chaos. On the contrary, indirect, noisy and nonstationary measurements can provide spurious or no results at all.

7. Conclusions

The analysis of nonlinear dynamical systems relies on the possibility of reconstructing proxy state space evolutions out of recorded scalar sequences, a possibility granted by Takens' theorem. Time-delay embedding is a viable and widespread method to carry out this reconstruction, although its correct implementation is not trivial. As discussed in Chapter 2, no definitive answer to the issue of optimal embedding exists. Throughout the present thesis work, the issue of optimal embedding was tackled by proposing a new method in the attempt to overcome the drawbacks of traditional approaches. First of all, our approach does not provide—as it is conventionally done—a single choice of embedding parameters m, L. More realistically, a set of statistically equivalent choices is identified. This identification proceeds by first ruling out those choices that yield a *bad* reconstruction, either because the reconstructed state space trajectory turns out to be compatible with a Gaussian white noise source, or because it is incompatible with a finite-dimensional system. The use of statistical testing as a criterion to exclude embedding choices that are not suitable to reconstruct the dynamics is one of the perks of the method, as it bypasses the issue of choosing and trimming thresholds. In the method's last step, the correlation dimension ν is estimated for the embedding pairs m, L that survived the previous skimming steps: the uniformity of this estimate is taken as a marker of good embedding choices.

Our approach provides, at the same time, some information on the sequence source and—if the latter is chaotic—a set of optimal embedding choices to reconstruct the dynamics. Therefore, the approach makes up not only an *optimal embedding criterion*, but also a tool to characterize nonlinear systems. For example, in Chapter 4, by analyzing synthetic sequences, it was shown that the method can distinguish chaotic sequences from stochastic ones having the same power spectrum (or autocorrelation). Identification of chaotic sources was successfully shown for various synthetic and experimental sequences. As an exception, the case of a standard Rössler sequence presented some difficulties caused by the system being almost periodic, an issue that prompted the modelling of correlation dimension of noisy periodic signals.

Concerning the last stage of the method, two estimators of ν are proposed in this thesis: a computationally inexpensive—though less precise—one, which can provide a preliminary assessment of ν, and

a more precise one, introduced in Chapter 5, which can be used to improve the initial assessment. The improvement granted by the more precise estimator can be quite crucial for a robust characterization of dynamical systems, as shown in Chapter 6 for the analysis of experimentally recorded sequences. In the analysis of recordings from an electronic circuit, the preliminary evaluation that relies on the less precise ν estimator suggested the existence of an underlying chaotic source, but only the application of the more precise estimator led to a definitive characterization.

An interesting result of our approach concerns the role of the embedding window $(m-1)L$. As mentioned in Chapter 2, the embedding window is considered as a relevant parameter to the purpose of optimal embedding, but most existing criteria neglect its role. Sets of suitable embedding choices identified by our approach, on the other hand, correspond to embedding windows that are intermediate between two extrema, which are taken as estimates of the *irrelevance* and *redundance* times of the input sequence. This condition makes up a further check of the quality of the method's outcomes and allows discarding spurious results.

The approach described in this thesis work was applied to two sets of experimentally recorded sequences, the first one concerning an electronic circuit and the second one corresponding to brain activity measured via magnetoencephalography. Besides providing further examples of the method's capabilities, these two tests allow to make some considerations. On the one hand, chaos is robustly identified when a system is characterized by few effective degrees of freedom and sufficiently high signal-to-noise ratio, as in the case of Chua's circuit implemented in Sec. 6.1. The opposite is true for complex systems having many more degrees of freedom: subsets of constituent units have to act synchronously in order to overcome the noisy background, which is partially contributed by the incoherent superposition of the motion of all the other units. A further complication is given by the nonstationarity of systems such as the human brain: as opposite to an electronic circuit, one has to be more careful in recording, preprocessing and selecting data, otherwise analytical algorithms might end up providing spurious results.

As a final remark, a feature of the method described in this thesis work is that it provides a "snapshot" of the system generating an input sequence in terms of an embedding lattice map. This approach focused on "mapping" the system onto the embedding lattice can be extended also to other metrics beyond the statistical tests and the estimation of ν employed in this work. Analyzing a system reconstructed through many different choices of embedding parameters provides a more complete view of its features, in analogy with pictures of an elaborate object taken from different positions. This idea is in line with the concept expressed above that, rather than seeking a single choice m, L that provides the best reconstruction of the dynamics, a set of statistically equivalent choices exists, and the ascertainment of this set is, at the same time, a way to study the properties of the source itself.

Bibliography

[1] J. Theiler, *Estimating fractal dimension*, J. Opt. Soc. Am. A **7** (1990) 1055–1073. doi:10.1364/JOSAA.7.001055.

[2] E. Bradley, H. Kantz, *Nonlinear time-series analysis revisited*, Chaos **25** (2015) 097610. doi:10.1063/1.4917289.

[3] D. K. C. MacDonald, Noise and Fluctuations: An Introduction, 1st Edition, John Wiley & Sons, New York, 1962.

[4] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Time Series Analysis: Forecasting and Control, 3rd Edition, Prentice-Hall, New Jersey, 1994.

[5] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, Cambridge, UK, 2004.

[6] S. H. Strogatz, Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, Addison-Wesley, Reading, Massachusetts, 1994.

[7] J. C. Sprott, *Simplest dissipative chaotic flow*, Phys. Lett. A **228** (1997) 271–274. doi:10.1016/S0375-9601(97)00088-1.

[8] R. Devaney, An Introduction to Chaotic Dynamical Systems, 2nd Edition, Addison-Wesley, Redwood City, CA, 1989.

[9] E. N. 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, 1972.

[10] C. Grebogi, E. Ott, S. Pelikan, J. A. Yorke, *Strange attractors that are not chaotic*, Physica D **13** (1984) 261–268. doi:10.1016/0167-2789(84)90282-3.

[11] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, L. S. Tsimring, *The analysis of observed chaotic data in physical systems*, Rev. Mod. Phys. **65** (1993) 1331–1392. doi:10.1103/RevModPhys.65.1331.

[12] J. C. Sprott, Chaos and Time-Series Analysis, Oxford University Press, Oxford, UK, 2003.

[13] E. N. Lorenz, *Deterministic Nonperiodic Flow*, J. Atmos. Sci. **20** (1963) 130–141. doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

[14] J. C. Sprott, G. Rowlands, *Improved correlation dimension calculation*, Int. J. Bifurcat. Chaos **11** (2001) 1865–1880. doi:10.1142/S021812740100305X.

[15] Y. Ueda, *Randomly transitional phenomena in the system governed by Duffing's equation*, J. Stat. Phys. **20** (1979) 181–196. doi:10.1007/BF01011512.

[16] O. E. Rössler, *An equation for continuous chaos*, Phys. Lett. A **57** (1976) 397–398. doi:10.1016/0375-9601(76)90101-8.

[17] D. Farmer, J. Crutchfield, H. Froehling, N. Packard, R. Shaw, *Power spectra and mixing properties of strange attractors*, Ann. NY Acad. Sci. **357** (1980) 453–471. doi:10.1111/j.1749-6632.1980.tb29710.x.

[18] P. J. Prince, J. R. Dormand, *High order embedded Runge-Kutta formulae*, J. Comput. Appl. Math. **7** (1981) 67–75. doi:10.1016/0771-050X(81)90010-3.

[19] J. L. Hudson, J. C. Mankin, *Chaos in the Belousov-Zhabotinskii reaction*, J. Chem. Phys. **74** (1981) 6171–6177. doi:10.1063/1.441007.

[20] J. Roux, *Experimental studies of bifurcations leading to chaos in the Belousof-Zhabotinsky reaction*, Physica D **7** (1983) 57–68. doi:10.1016/0167-2789(83)90115-x.

[21] I. R. Epstein, *Oscillations and chaos in chemical systems*, Physica D **7** (1983) 47–56. doi:10.1016/0167-2789(83)90114-8.

[22] I. R. Epstein, K. Showalter, *Nonlinear Chemical Dynamics: Oscillations, Patterns, and Chaos*, J. Phys. Chem. **100** (1996) 13132–13147. doi:10.1021/jp953547m.

[23] N. Yoshida, *Characterization of patterns, oscillations, and chaos in chemical systems*, J. Chem. Phys. **115** (2001) 8327–8330. doi:10.1063/1.1410378.

[24] D. Poland, *Cooperative catalysis and chemical chaos: a chemical model for the Lorenz equations*, Physica D **65** (1993) 86–99. doi:10.1016/0167-2789(93)90006-m.

[25] P. E. King, T. L. Ochs, A. D. Hartman, *Chaotic responses in electric arc furnaces*, J. App. Phys. **76** (1994) 2059–2065. doi:10.1063/1.357639.

[26] I. Grabec, *Chaos generated by the cutting process*, Phys. Lett. A **117** (1986) 384–386. doi:10.1016/0375-9601(86)90003-4.

[27] R. Rusinek, J. Warminski, *Attractor reconstruction of self-excited mechanical systems*, Chaos Soliton. Fract. **40** (2009) 172–182. doi:10.1016/j.chaos.2007.07.040.

[28] M. Sayah, M. S. Baptista, J. Ing, M. Wiercigroch, *Attractor reconstruction of an impact oscillator for parameter identification*, Int. J. Mech. Sci. **103** (2015) 212–223. doi:10.1016/j.ijmecsci.2015.08.005.

[29] C. Craig, R. D. Neilson, J. Penman, *The Use of Correlation Dimension in Condition Monitoring of Systems with Clearance*, J. Sound. Vib. **231** (2000) 1–17. doi:10.1006/jsvi.1998.2713.

[30] S. Hu, A. Raman, *Chaos in Atomic Force Microscopy*, Phys. Rev. Lett. **96** (2006). doi:10.1103/PhysRevLett.96.036107.

[31] K. Ikeda, *Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system*, Opt. Commun. **30** (1979) 257–261. doi:10.1016/0030-4018(79)90090-7.

[32] K. Ikeda, H. Daido, O. Akimoto, *Optical Turbulence: Chaotic Behavior of Transmitted Light from a Ring Cavity*, Phys. Rev. Lett. **45** (1980) 709–712. doi:10.1103/PhysRevLett.45.709.

[33] F. A. Hopf, D. L. Kaplan, H. M. Gibbs, R. L. Shoemaker, *Bifurcations to chaos in optical bistability*, Phys. Rev. A **25** (1982) 2172–2182. doi:10.1103/PhysRevA.25.2172.

[34] Y. Cho, T. Umeda, *Observation of chaos in a semiconductor laser with delayed feedback*, Opt. Commun. **59** (1986) 131–136. doi:10.1016/0030-4018(86)90464-5.

[35] F. T. Arecchi, R. Meucci, G. Puccioni, J. Tredicce, *Experimental Evidence of Subharmonic Bifurcations, Multistability, and Turbulence in a Q-Switched Gas Laser*, Phys. Rev. Lett. **49** (1982) 1217–1220. doi:10.1103/PhysRevLett.49.1217.

[36] T. Midavaine, D. Dangoisse, P. Glorieux, *Observation of Chaos in a Frequency-Modulated CO_2 Laser*, Phys. Rev. Lett. **55** (1985) 1989–1992. doi:10.1103/PhysRevLett.55.1989.

[37] D. J. Biswas, V. Dev, U. K. Chatterjee, *Experimental observation of oscillatory instabilities and chaos in a gain-modulated single-mode cw CO_2 laser*, Phys. Rev. A **35** (1987) 456–458. doi:10.1103/PhysRevA.35.456.

[38] H. Haken, *Analogy between higher instabilities in fluids and lasers*, Phys. Lett. A **53** (1975) 77–78. doi:10.1016/0375-9601(75)90353-9.

[39] E. H. M. Hogenboom, W. Klische, C. O. Weiss, A. Godone, *Instabilities of a Homogeneously Broadened Laser*, Phys. Rev. Lett. **55** (1985) 2571–2574. doi:10.1103/PhysRevLett.55.2571.

[40] U. Hübner, N. B. Abraham, C. O. Weiss, *Dimensions and entropies of chaotic intensity pulsations in a single-mode far-infrared NH_3 laser*, Phys. Rev. A **40** (1989) 6354–6365. doi:10.1103/PhysRevA.40.6354.

[41] See https://www.comp-engine.org/#!visualize/4057f818-3874-11e8-8680-0242ac120002 for public research data of intensity from a far-infrared NH_3 laser. The time series contains 9093 points. The signal is digitized with 8 bit resolution and with a sampling period of 40 ns.

[42] S. Sunada, T. Harayama, P. Davis, K. Tsuzuki, K. Arai, K. Yoshimura, A. Uchida, *Noise amplification by chaotic dynamics in a delayed feedback laser system and its application to nondeterministic random bit generation*, Chaos **22** (2012) 047513. doi:10.1063/1.4754872.

[43] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, P. Davis, *Fast physical random bit generation with chaotic semiconductor lasers*, Nat. Photonics **2** (2008) 728–732. doi:10.1038/nphoton.2008.227.

[44] X. Li, J. Zhuang, S. Li, J. Gao, S. Chan, *Randomness evaluation for an optically injected chaotic semiconductor laser by attractor reconstruction*, Phys. Rev. E **94** (2016). doi:10.1103/PhysReve.94.042214.

[45] M. Sciamanna, K. A. Shore, *Physics and applications of laser diode chaos*, Nat. Photonics **9** (2015) 151–162. doi:10.1038/nphoton.2014.326.

[46] T. Matsumoto, L. O. Chua, S. Tanaka, *Simplest chaotic nonautonomous circuit*, Phys. Rev. A **30** (1984) 1155–1157. doi:10.1103/PhysRevA.30.1155.

[47] T. Matsumoto, *A chaotic attractor from Chua's circuit*, IEEE T. Circuits Syst. **31** (1984) 1055–1058. doi:10.1109/TCS.1984.1085459.

[48] R. N. Madan, Chua's Circuit: A Paradigm for Chaos, World Scientific Publishing, 1993. doi:10.1142/1997.

[49] L. Minati, M. Frasca, P. Oświęcimka, L. Faes, S. Drożdż, *Atypical transistor-based chaotic oscillators: Design, realization, and diversity*, Chaos **27** (2017) 073113. doi:10.1063/1.4994815.

[50] L. Minati, M. Frasca, N. Yoshimura, L. Ricci, P. Oświęcimka, Y. Koike, K. Masu, H. Ito, *Current-Starved Cross-Coupled CMOS Inverter Rings as Versatile Generators of Chaotic and Neural-Like Dynamics Over Multiple Frequency Decades*, IEEE Access **7** (2019) 54638–54657. doi:10.1109/access.2019.2912903.

[51] B. Muthuswamy, *Implementing Memristor Based Chaotic Circuits*, Int. J. Bifurcat. Chaos **20** (2010) 1335–1350. doi:10.1142/s0218127410026514.

[52] L. Minati, L. V. Gambuzza, W. J. Thio, J. C. Sprott, M. Frasca, *A chaotic circuit based on a physical memristor*, Chaos Soliton. Fract. **138** (2020) 109990. doi:10.1016/j.chaos.2020.109990.

[53] A. Namajūnas, K. Pyragas, A. Tamaševičius, *An electronic analog of the mackey-glass system*, Phys. Lett. A **201** (1995) 42–46. doi:10.1016/0375-9601(95)00208-k.

[54] L. Minati, *Experimental synchronization of chaos in a large ring of mutually coupled single-transistor oscillators: amplitude, phase, and clustering effects*, Chaos **24** (2014) 043108. doi:10.1063/1.4896815.

[55] L. Minati, *Remote synchronization of amplitudes across an experimental ring of non-linear oscillators*, Chaos **25** (2015) 123107. doi:10.1063/1.4936791.

[56] L. Minati, M. Frasca, G. Giustolisi, P. Oświęcimka, S. Drożdż, L. Ricci, *High-dimensional dynamics in a single-transistor oscillator containing feynman-sierpiński resonators: Effect of fractal depth and irregularity*, Chaos **28** (2018) 093112. doi:10.1063/1.5047481.

[57] M. Lei, G. Meng, *Detecting Nonlinearity of Sunspot Number*, Int. J. Nonlin. Sci. Num. **5** (2004). doi:10.1515/ijnsns.2004.5.4.321.

[58] R. A. Greenkorn, *Analysis of Sunspot Activity Cycles*, Sol. Phys. **255** (2009) 301–323. doi:10.1007/s11207-009-9331-z.

[59] N. Jevtić, J. S. Schweitzer, C. J. Cellucci, *Research Note: Nonlinear time series analysis of northern and southern solar hemisphere daily sunspot numbers in search of short-term chaotic behavior*, Astron. Astrophys. **379** (2001) 611–615. doi:10.1051/0004-6361:20011277.

[60] P. D. Mininni, D. O. Gómez, G. B. Mindlin, *Stochastic Relaxation Oscillator Model for the Solar Cycle*, Phys. Rev. Lett. **85** (2000) 5476–5479. doi:10.1103/PhysRevLett.85.5476.

[61] J. F. Lindner, V. Kohar, B. Kia, M. Hippke, J. G. Learned, W. L. Ditto, *Strange Nonchaotic Stars*, Phys. Rev. Lett. **114** (2015). doi:10.1103/PhysRevLett.114.054101.

[62] B. Sivakumar, *Chaos theory in geophysics: past, present and future*, Chaos Soliton. Fract. **19** (2004) 441–462. doi:10.1016/s0960-0779(03)00055-9.

[63] M. De Domenico, M. A. Ghorbani, O. Makarynskyy, D. Makarynska, H. Asadi, *Chaos and reproduction in sea level*, Appl. Math. Model. **37** (2013) 3687–3697. doi:10.1016/j.apm.2012.08.018.

[64] R. Khatibi, M. A. Ghorbani, M. T. Aalami, K. Kocak, O. Makarynskyy, D. Makarynska, M. Aalinezhad, *Dynamics of hourly sea level at Hillarys Boat Harbour, Western Australia: a chaos theory perspective*, Ocean Dynam. **61** (2011) 1797–1807. doi:10.1007/s10236-011-0466-8.

[65] J. Huang, D. L. Turcotte, *Are earthquakes an example of deterministic chaos?*, Geophys. Res. Lett. **17** (1990) 223–226. doi:10.1029/gl017i003p00223.

[66] A. C. Iliopoulos, G. P. Pavlos, E. E. Papadimitriou, D. S. Sfiris, M. A. Athanasiou, V. G. Tsoutsouras, *Chaos Self Organized Criticality Intermittent Turbulence And Nonextensivity Revealed From Seismogenesis In North Aegean Area*, Int. J. Bifurcat. Chaos **22** (2012) 1250224. doi:10.1142/s0218127412502240.

[67] F. B. Pelap, L. Y. Kagho, C. F. Fogang, *Chaotic behavior of earthquakes induced by a nonlinear magma up flow*, Chaos Soliton. Fract. **87** (2016) 71–83. doi:10.1016/j.chaos.2016.03.018.

[68] A. Gualandi, J. Avouac, S. Michel, D. Faranda, *The predictable chaos of slow earthquakes*, Sci. Adv. **6** (2020) eaaz5548. doi:10.1126/sciadv.aaz5548.

[69] A. Sornette, J. Dubois, J. L. Cheminée, D. Sornette, *Are sequences of volcanic eruptions deterministically chaotic?*, J. Geophys. Res. **96** (1991) 11931. doi:10.1029/91jb00515.

[70] W. Marzocchi, F. Mulargia, G. Gonzato, *Detecting low-dimensional chaos in geophysical time series*, J. Geophys. Res. **102** (1997) 3195–3209. doi:10.1029/96jb03268.

[71] W. Marzocchi, M. S. Bebbington, *Probabilistic eruption forecasting at short and long time scales*, B. Volcanol. **74** (2012) 1777–1805. doi:10.1007/s00445-012-0633-x.

[72] V. Cuomo, V. Lapenna, M. Macchiato, C. Serio, L. Telesca, *Stochastic behaviour and scaling laws in geoelectrical signals measured in a seismic area of southern Italy*, Geophys. J. Int. **139** (1999) 889–894. doi:10.1046/j.1365-246x.1999.00997.x.

[73] A. A. Tsonis, J. B. Elsner, *Chaos, Strange Attractors, and Weather*, B. Am. Meteorol. Soc. **70** (1989) 14–23. doi:10.1175/1520-0477(1989)070<0014:csaaw>2.0.co;2.

[74] J. Slingo, T. Palmer, *Uncertainty in weather and climate prediction*, P. Roy. Soc. A **369** (2011) 4751–4767. doi:10.1098/rsta.2011.0161.

[75] C. L. Keppenne, C. Nicolis, *Global Properties and Local Structure of the Weather Attractor over Western Europe*, J. Atmos. Sci. **46** (1989) 2356–2370. doi:10.1175/1520-0469(1989)046<2356:gpalso>2.0.co;2.

[76] H. N. Shirer, C. J. Fosmire, R. Wells, L. Suciu, *Estimating the Correlation Dimension of Atmospheric Time Series*, J. Atmos. Sci. **54** (1997) 211–230. doi:10.1175/1520-0469(1997)054<0211:etcdoa>2.0.co;2.

[77] T. E. Karakasidis, A. Charakopoulos, *Detection of low-dimensional chaos in wind time series*, Chaos Soliton. Fract. **41** (2009) 1723–1732. doi:10.1016/j.chaos.2008.07.020.

[78] E. N. Lorenz, *Dimension of weather and climate attractors*, Nature **353** (1991) 241–244. doi:10.1038/353241a0.

[79] A. A. Tsonis, J. B. Elsner, K. P. Georgakakos, *Estimating the Dimension of Weather and Climate Attractors: Important Issues about the Procedure and Interpretation*, J. Atmos. Sci. **50** (1993) 2549–2555. doi:10.1175/1520-0469(1993)050<2549:etdowa>2.0.co;2.

[80] T. Reyes, B. Shen, *A recurrence analysis of chaotic and non-chaotic solutions within a generalized nine-dimensional Lorenz model*, Chaos Soliton. Fract. **125** (2019) 1–12. doi:10.1016/j.chaos.2019.05.003.

[81] C. Nicolis, N. G., *Is there a climatic attractor?*, Nature **311** (1984) 529–532. doi:10.1038/311529a0.

[82] P. Grassberger, *Do climatic attractors exist?*, Nature **323** (1986) 609–612. doi:10.1038/323609a0.

[83] A. Tsonis, J. Elsner, *The weather attractor over very short timescales*, Nature **333** (1988) 545–547. doi:10.1038/333545a0.

[84] I. Procaccia, *Complex or just complicated?*, Nature **333** (1988) 498–499. doi:10.1038/333498a0.

[85] A. Király, I. M. Jánosi, *Stochastic modeling of daily temperature fluctuations*, Phys. Rev. E **65** (2002) 051102. doi:10.1103/PhysRevE.65.051102.

[86] M. Mackey, L. Glass, *Oscillation and chaos in physiological control systems*, Science **197** (1977) 287–289. doi:10.1126/science.267326.

[87] R. Pool, *Is it healthy to be chaotic?*, Science **243** (1989) 604–607. doi:10.1126/science.2916117.

[88] M. R. Guevara, L. Glass, A. Shrier, *Phase locking, period-doubling bifurcations, and irregular dynamics in periodically stimulated cardiac cells*, Science **214** (1981) 1350–1353. doi:10.1126/science.7313693.

[89] L. Glass, M. R. Guevara, A. Shrier, R. Perez, *Bifurcation and chaos in a periodically stimulated cardiac oscillator*, Physica D **7** (1983) 89–101. doi:10.1016/0167-2789(83)90119-7.

[90] A. L. Goldberger, B. J. West, *Applications of Nonlinear Dynamics to Clinical Cardiology*, Ann. NY Acad. Sci. **504** (1987) 195–213. doi:10.1111/j.1749-6632.1987.tb48733.x.

[91] A. Babloyantz, A. Destexhe, *Is the normal heart a periodic oscillator?*, Biol. Cybern. **58** (1988) 203–211. doi:10.1007/bf00364139.

[92] R. B. Govindan, K. Narayanan, M. S. Gopinathan, *On the evidence of deterministic chaos in ECG: Surrogate and predictability analysis*, Chaos **8** (1998) 495–502. doi:10.1063/1.166330.

[93] J. Q. Zhang, A. V. Holden, O. Monfredi, M. R. Boyett, H. Zhang, *Stochastic vagal modulation of cardiac pacemaking may lead to erroneous identification of cardiac "chaos"*, Chaos **19** (2009) 028509. doi:10.1063/1.3141426.

[94] P. C. Ivanov, L. A. Nunes Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, H. E. Stanley, *Multifractality in human heartbeat dynamics*, Nature **399** (1999) 461–465. doi:10.1038/20924.

[95] R. T. Baillie, A. A. Cecen, C. Erkal, *Normal heartbeat series are nonchaotic, nonlinear, and multifractal: New evidence from semiparametric and parametric tests*, Chaos **19** (2009) 028503. doi:10.1063/1.3152006.

[96] D. Ghosh, S. Dutta, S. Chakraborty, S. Samanta, *Chaos based nonlinear analysis to study cardiovascular responses to changes in posture*, Physica A **512** (2018) 392–403. doi:10.1016/j.physa.2018.08.125.

[97] T. Henriques, M. Ribeiro, A. Teixeira, L. Castro, L. Antunes, C. Costa-Santos, *Nonlinear Methods Most Applied to Heart-Rate Time Series: A Review*, Entropy **22** (2020) 309. doi:10.3390/e22030309.

[98] E. Niedermeyer, F. H. L. da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 5th Edition, Lippincott Williams & Wilkins, Philadelphia, 2005.

[99] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, O. V. Lounasmaa, *Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain*, Rev. Mod. Phys. **65** (1993) 413–497. doi:10.1103/RevModPhys.65.413.

[100] S. Supek, C. J. Aine, Magnetoencephalography: From Signals to Dynamic Cortical Networks, 1st Edition, Springer-Verlag, Berlin Heidelberg, 2014. doi:10.1007/978-3-642-33045-2.

[101] M. R. Guevara, L. Glass, M. C. Mackey, A. Shrier, *Chaos in neurobiology*, IEEE T. Syst. Man. Cyb. **SMC-13** (1983) 790–798. doi:10.1109/tsmc.1983.6313073.

[102] A. Babloyantz, J. M. Salazar, C. Nicolis, *Evidence of chaotic dynamics of brain activity during the sleep cycle*, Phys. Lett. A **111** (1985) 152–156. doi:10.1016/0375-9601(85)90444-x.

[103] J. Fell, J. Röschke, P. Beckmann, *Deterministic chaos and the first positive Lyapunov exponent: a nonlinear analysis of the human electroencephalogram during sleep*, Biol. Cybern. **69** (1993) 139–146. doi:10.1007/bf00226197.

[104] J. P. Pijn, J. Van Neerven, A. Noest, L. F. H., *Chaos or noise in EEG signals; dependence on state and brain site*, Electroen. Clin. Neuro. **79** (1991) 371–381. doi:10.1016/0013-4694(91)90202-f.

[105] K. Lehnertz, C. E. Elger, *Can Epileptic Seizures be Predicted? Evidence from Nonlinear Time Series Analysis of Brain Electrical Activity*, Phys. Rev. Lett. **80** (1998) 5019–5022. doi:10.1103/physrevlett.80.5019.

[106] A. Babloyantz, A. Destexhe, *Low-dimensional chaos in an instance of epilepsy.*, P. Natl. Acad. Sci. USA **83** (1986) 3513–3517. doi:10.1073/pnas.83.10.3513.

[107] K. Lehnertz, C. E. Elger, *Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss*, Electroen. Clin. Neuro. **95** (1995) 108–117. doi:10.1016/0013-4694(95)00071-6.

[108] C. A. Skarda, W. J. Freeman, *How brains make chaos in order to make sense of the world*, Behav. Brain. Sci. **10** (1987) 161–173. doi:10.1017/s0140525x00047336.

[109] J. Jeong, M. K. Joung, S. Y. Kim, *Quantification of emotion by nonlinear analysis of the chaotic dynamics of electroencephalograms during perception of 1/f music*, Biol. Cybern. **78** (1998) 217–225. doi:10.1007/s004220050428.

[110] X. Xie, Z. Cao, X. Weng, *Spatiotemporal nonlinearity in resting-state fMRI of the human brain*, Neuroimage **40** (2008) 1672–1685. doi:10.1016/j.neuroimage.2008.01.007.

[111] C. J. Stam, *Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field*, Clin. Neurophysiol. **116** (2005) 2266–2301. doi:10.1016/j.clinph.2005.06.011.

[112] M. I. Rabinovich, H. D. I. Abarbanel, *The role of chaos in neural systems*, Neuroscience **87** (1998) 5–14. doi:10.1016/s0306-4522(98)00091-8.

[113] H. Korn, P. Faure, *Is there chaos in the brain? II. Experimental evidence and related models*, C. R. Biol. **326** (2003) 787–840. doi:10.1016/j.crvi.2003.09.011.

[114] H. Hayashi, M. Nakao, K. Hirakawa, *Chaos in the self-sustained oscillation of an excitable biological membrane under sinusoidal stimulation*, Phys. Lett. A **88** (1982) 265–266. doi:10.1016/0375-9601(82)90245-6.

[115] H. Hayashi, S. Ishizuka, *Chaotic nature of bursting discharges in the Onchidium pacemaker neuron*, J. Theor. Biol. **156** (1992) 269–291. doi:10.1016/s0022-5193(05)80676-9.

[116] F. Pittorino, M. Ibáñez-Berganza, M. di Volo, A. Vezzani, R. Burioni, *Chaos and Correlated Avalanches in Excitatory Neural Networks with Synaptic Plasticity*, Phys. Rev. Lett. **118** (2017) 098102. doi:10.1103/physrevlett.118.098102.

[117] W. S. Pritchard, D. W. Duke, K. K. Krieble, *Dimensional analysis of resting human EEG II: Surrogate-data testing indicates nonlinearity but not low-dimensional chaos*, Psychophysiology **32** (1995) 486–491. doi:10.1111/j.1469-8986.1995.tb02100.x.

[118] J. Theiler, P. E. Rapp, *Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram*, Electroen. Clin. Neuro. **98** (1996) 213–222. doi:10.1016/0013-4694(95)00240-5.

[119] R. Hornero, D. Abásolo, J. Escudero, C. Gómez, *Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with alzheimer's disease*, P. Roy. Soc. A **367** (2008) 317–336. doi:10.1098/rsta.2008.0197.

[120] M. J. Brookes, E. L. Hall, S. E. Robson, D. Price, L. Palaniyappan, E. B. Liddle, P. F. Liddle, S. E. Robinson, P. G. Morris, *Complexity Measures in Magnetoencephalography: Measuring "Disorder" in Schizophrenia*, PLoS ONE **10** (2015) e0120991. doi:10.1371/journal.pone.0120991.

[121] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, B. Vanrumste, *Review on solving the inverse problem in EEG source analysis*, J. Neuroeng. Rehabil. **5** (2008) 25. doi:10.1186/1743-0003-5-25.

[122] H. Kantz, E. Olbrich, *Scalar observations from a class of high-dimensional chaotic systems: Limitations of the time delay embedding*, Chaos **7** (1997) 423–429. doi:10.1063/1.166215.

[123] N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw, *Geometry from a Time Series*, Phys. Rev. Lett. **45** (1980) 712–716. doi:10.1103/PhysRevLett.45.712.

[124] F. Takens, *Detecting strange attractors in turbolence*, in: D. A. Rand, L. S. Young (Eds.), Dynamical Systems and Turbolence (Warwick 1980), Springer Verlag, Berlin, Heidelberg, 1980, pp. 366–381.

[125] R. Mañé, *On the dimension of the compact invariant sets of certain non-linear maps*, in: D. A. Rand, L. S. Young (Eds.), Dynamical Systems and Turbolence (Warwick 1980), Springer Verlag, Berlin, Heidelberg, 1980, pp. 230–242.

[126] M. Ding, C. Grebogi, E. Ott, T. Sauer, J. A. Yorke, *Estimating correlation dimension from a chaotic time series: when does plateau onset occur?*, Physica D **69** (1993) 404–424. doi:10.1016/0167-2789(93)90103-8.

[127] P. Grassberger, T. Schreiber, C. Schaffrath, *Nonlinear time sequence analysis*, Int. J. Bifurcat. Chaos **1** (1991) 521–547. doi:10.1142/S0218127491000403.

[128] D. S. Broomhead, G. P. King, *Extracting qualitative dynamics from experimental data*, Physica D **20** (1986) 217–236. doi:10.1016/0167-2789(86)90031-X.

[129] P. S. Landa, M. G. Rosenblum, *Time series analysis for system identification and diagnostics*, Physica D **48** (1991) 232–254. doi:10.1016/0167-2789(91)90059-I.

[130] M. Casdagli, S. Eubank, J. D. Farmer, J. Gibson, *State space reconstruction in the presence of noise*, Physica D **51** (1991) 52–98. doi:10.1016/0167-2789(91)90222-U.

[131] T. Sauer, J. A. Yorke, M. Casdagli, *Embedology*, J. Stat. Phys. **65** (1991) 579–616. doi:10.1007/BF01053745.

[132] C. J. Cellucci, A. M. Albano, P. E. Rapp, *Comparative study of embedding methods*, Phys. Rev. E **67** (2003) 066210. doi:10.1103/PhysRevE.67.066210.

[133] M. Franchi, L. Ricci, *Statistical properties of the maximum Lyapunov exponent calculated via the divergence rate method*, Phys. Rev. E **90** (2014) 062920. doi:10.1103/PhysRevE.90.062920.

[134] A. M. Fraser, H. L. Swinney, *Independent coordinates for strange attractors from mutual information*, Phys. Rev. A **33** (1986) 1134–1140. doi:10.1103/PhysRevA.33.1134.

[135] W. Liebert, H. G. Schuster, *Proper choice of the time delay for the analysis of chaotic time series*, Phys. Lett. A **142** (1989) 107–111. doi:10.1016/0375-9601(89)90169-2.

[136] T. Buzug, G. Pfister, *Comparison of algorithms calculating optimal embedding parameters for delay time coordinates*, Physica D **58** (1992) 127–137. doi:10.1016/0167-2789(92)90104-U.

[137] M. B. Kennel, R. Brown, H. D. I. Abarbanel, *Determining embedding dimension for phase-space reconstruction using a geometrical construction*, Phys. Rev. A **45** (1992) 3403–3411. doi:10.1103/PhysRevA.45.3403.

[138] R. Hegger, H. Kantz, T. Schreiber, *Practical implementation of nonlinear time series methods: The TISEAN package*, Chaos **9** (1999) 413–435. doi:10.1063/1.166424.

[139] L. M. Pecora, L. Moniz, J. Nichols, T. L. Carroll, *A unified approach to attractor reconstruction*, Chaos **17** (2007) 013110. doi:10.1063/1.2430294.

[140] S. P. Garcia, J. S. Almeida, *Nearest neighbor embedding with different time delays*, Phys. Rev. E **71** (2005) 037204. doi:10.1103/PhysRevE.71.037204.

[141] J. Gao, Z. Zheng, *Local exponential divergence plot and optimal embedding of a chaotic time series*, Phys. Lett. A **181** (1993) 153–158. doi:10.1016/0375-9601(93)90913-K.

[142] J. Gao, Z. Zheng, *Direct Dynamical Test for Deterministic Chaos*, Europhys. Lett. **25** (1994) 485–490. doi:10.1209/0295-5075/25/7/002.

[143] W. Liebert, K. Pawelzik, H. G. Schuster, *Optimal Embeddings of Chaotic Attractors from Topological Considerations*, Europhys. Lett. **14** (1991) 521–526. doi:10.1209/0295-5075/14/6/004.

[144] K. Judd, A. Mees, *Embedding as a modeling problem*, Physica D **120** (1998) 273–286. doi:10.1016/S0167-2789(98)00089-X.

[145] Y. Hirata, H. Suzuki, K. Aihara, *Reconstructing state spaces from multivariate data using variable delays*, Phys. Rev. E **74** (2006) 026202. doi:10.1103/PhysRevE.74.026202.

[146] R. Hegger, H. Kantz, *Improved false nearest neighbor method to detect determinism in time series data*, Phys. Rev. E **60** (1999) 4970–4973. doi:10.1103/PhysRevE.60.4970.

[147] M. B. Kennel, H. D. I. Abarbanel, *False neighbors and false strands: A reliable minimum embedding dimension algorithm*, Phys. Rev. E **66** (2002) 026209. doi:10.1103/PhysRevE.66.026209.

[148] D. Chelidze, *Reliable Estimation of Minimum Embedding Dimension Through Statistical Analysis of Nearest Neighbors*, J. Comput. Nonlin. Dyn. **12** (2017) 051024. doi:10.1115/1.4036814.

[149] C. Nichkawde, *Optimal state-space reconstruction using derivatives on projected manifold*, Phys. Rev. E **87** (2013) 022905. doi:10.1103/PhysRevE.87.022905.

[150] L. C. Uzal, G. L. Grinblat, P. F. Verdes, *Optimal reconstruction of dynamical systems: A noise amplification approach*, Phys. Rev. E **84** (2011) 016223. doi:10.1103/PhysRevE.84.016223.

[151] J. Garland, R. G. James, E. Bradley, *Leveraging information storage to select forecast-optimal parameters for delay-coordinate reconstructions*, Phys. Rev. E **93** (2016) 022221. doi:10.1103/PhysRevE.93.022221.

[152] V. Deshmukh, E. Bradley, J. Garland, J. D. Meiss, *Using curvature to select the time lag for delay reconstruction*, Chaos **30** (2020) 063143. doi:10.1063/5.0005890.

[153] D. Kugiumtzis, *State space reconstruction parameters in the analysis of chaotic time series – the role of the time window length*, Physica D **95** (1996) 13–28. doi:10.1016/0167-2789(96)00054-1.

[154] Y. Manabe, B. Chakraborty, *A novel approach for estimation of optimal embedding parameters of nonlinear time series by structural learning of neural network*, Neurocomputing **70** (2007) 1360–1371. doi:10.1016/j.neucom.2006.06.005.

[155] A. Maus, J. C. Sprott, *Neural network method for determining embedding dimension of a time series*, Commun. Nonlinear Sci. **16** (2011) 3294–3302. doi:10.1016/j.cnsns.2010.10.030.

[156] L. Moniz, L. Pecora, J. Nichols, M. Todd, J. R. Wait, *Dynamical Assessment of Structural Damage Using the Continuity Statistic*, Struct. Health Monit. **3** (2004) 199–212. doi:10.1177/1475921704042681.

[157] T. L. Carroll, J. M. Byers, *Dimension from covariance matrices*, Chaos **27** (2017) 023101. doi:10.1063/1.4975063.

[158] H. S. Greenside, A. Wolf, J. Swift, T. Pignataro, *Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors*, Phys. Rev. A **25** (1982) 3453–3456. doi:10.1103/PhysRevA.25.3453.

[159] P. Grassberger, I. Procaccia, *Characterization of Strange Attractors*, Phys. Rev. Lett. **50** (1983) 346–349. doi:10.1103/PhysRevLett.50.346.

[160] P. Grassberger, I. Procaccia, *Measuring the strangeness of strange attractors*, Physica D **9** (1983) 189–208. doi:10.1016/0167-2789(83)90298-1.

[161] A. Perinelli, L. Ricci, *Identification of suitable embedding dimensions and lags for time series generated by chaotic, finite-dimensional systems*, Phys. Rev. E **98** (2018) 052226. doi:10.1103/PhysRevE.98.052226.

[162] L. A. Smith, *Intrinsic limits on dimension calculations*, Phys. Lett. A **133** (1988) 283–288. doi:10.1016/0375-9601(88)90445-8.

[163] I. Dvořák, J. Klaschka, *Modification of the Grassberger-Procaccia algorithm for estimating the correlation exponent of chaotic systems with high embedding dimension*, Phys. Lett. A **145** (1990) 225–231. doi:10.1016/0375-9601(90)90355-R.

[164] R. Badii, A. Politi, *Intrinsic oscillations in measuring the fractal dimension*, Phys. Lett. A **104** (1984) 303–305. doi:10.1016/0375-9601(84)90801-6.

[165] J. Theiler, *Statistical precision of dimension estimators*, Phys. Rev. A **41** (1990) 3038–3051. doi:10.1103/PhysRevA.41.3038.

[166] J. Theiler, *Spurious dimension from correlation algorithms applied to limited time-series data*, Phys. Rev. A **34** (1986) 2427–2432. doi:10.1103/PhysRevA.34.2427.

[167] A. R. Osborne, A. Provenzale, *Finite correlation dimension for stochastic systems with power-law spectra*, Physica D **35** (1989) 357–381. doi:10.1016/0167-2789(89)90075-4.

[168] J. Theiler, *Some comments on the correlation dimension of $1/f^{\alpha}$ noise*, Phys. Lett. A **155** (1991) 480–493. doi:10.1016/0375-9601(91)90651-N.

[169] A. M. Albano, P. E. Rapp, A. Passamante, *Kolmogorov-Smirnov test distinguishes attractors with similar dimensions*, Phys. Rev. E **52** (1995) 196–206. doi:10.1103/PhysRevE.52.196.

[170] G. Benettin, L. Galgani, A. Giorgilli, J. Strelcyn, *Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory*, Meccanica **15** (1980) 9–20. doi:10.1007/BF02128236.

[171] G. Benettin, L. Galgani, A. Giorgilli, J. 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** (1980) 21–30. doi:10.1007/BF02128237.

[172] J. Gao, Z. Zheng, *Local exponential divergence plot and optimal embedding of a chaotic time-series*, Phys. Lett. A **181** (1993) 153–158. doi:10.1016/0375-9601(93)90913-K.

[173] M. T. Rosenstein, J. J. Collins, C. J. De Luca, *A practical method for calculating largest Lyapunov exponents from small data sets*, Physica D **65** (1993) 117–134. doi:10.1016/0167-2789(93)90009-P.

[174] H. Kantz, *A robust method to estimate the maximal Lyapunov exponent of a time series*, Phys. Lett. A **185** (1994) 77–87. doi:10.1016/0375-9601(94)90991-1.

[175] L. Ricci, A. Perinelli, M. Franchi, *Asymptotic behavior of the time-dependent divergence exponent*, Phys. Rev. E **101** (2020) 042211. doi:10.1103/PhysRevE.101.042211.

[176] J. E. Angus, *The Probability Integral Transform and Related Results*, SIAM Rev. **36** (1994) 652–654. doi:10.1137/1036146.

[177] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C, Cambridge University Press, Cambridge, UK, 1997.

[178] J. Durbin, Distribution Theory for Tests Based on the Sample Distribution Function, Society for Industrial and Applied Mathematics, 1973. doi:10.1137/1.9781611970586.

[179] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J. Doyne Farmer, *Testing for nonlinearity in time series: the method of surrogate data*, Physica D **58** (1992) 77–94. doi:10.1016/0167-2789(92)90102-S.

[180] T. Schreiber, A. Schmitz, *Surrogate time series*, Physica D **142** (2000) 346–382. doi:10.1016/S0167-2789(00)00043-9.

[181] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, A. Stefanovska, *Surrogate data for hypothesis testing of physical systems*, Phys. Rep. **748** (2018) 1–60. doi:10.1016/j.physrep.2018.06.001.

[182] See https://isocpp.org/ for the C++ Programming Language ISO Standard.

[183] OpenMP library website: https://www.openmp.org/.

[184] GNU Scientific Library website: https://www.gnu.org/software/gsl/.

[185] See https://gcc.gnu.org/ for the GNU Compiler Collection.

[186] A. P. Muñuzuri, J. A. K. Suykens, L. O. Chua, *A CNN approach to brain-like chaos-periodicity transitions*, Int. J. Bifurcat. Chaos **8** (1998) 2263–2278. doi:10.1142/S0218127498001856.

[187] D. Malagarriga, M. García-Vellisca, A. E. Villa, J. Buldú, J. García-Ojalvo, A. Pons, *Synchronization-based computation through networks of coupled oscillators*, Front. Comput. Neurosc. **9** (2015) 97. doi:10.3389/fncom.2015.00097.

[188] L. O. Chua, T. Roska, *The CNN paradigm*, IEEE T. Circuits-I **40** (1993) 147–156. doi:10.1109/81.222795.

[189] V. Pérez-Muñuzuri, A. P. Muñuzuri, M. Gómez-Gesteira, V. Pérez-Villar, L. Pivka, L. O. Chua, *Nonlinear waves, patterns and spatio-temporal chaos in cellular neural networks*, Philos. Tr. R. Soc. S-A **353** (1701) (1995) 101–113. doi:10.1098/rsta.1995.0093.

[190] R. FitzHugh, *Impulses and Physiological States in Theoretical Models of Nerve Membrane*, Biophys. J. **1** (1961) 445–466. doi:10.1016/S0006-3495(61)86902-6.

[191] S. Zambrano, I. P. Mariño, J. M. Seoane, M. A. F. Sanjuán, S. Euzzor, A. Geltrude, R. Meucci, F. T. Arecchi, *Synchronization of uncoupled excitable systems induced by white and coloured noise*, New J. Phys. **12** (2010) 053040. doi:10.1088/1367-2630/12/5/053040.

[192] M. Ciszak, S. Euzzor, A. Geltrude, F. T. Arecchi, R. Meucci, *Noise and coupling induced synchronization in a network of chaotic neurons*, Commun. Nonlinear Sci. Numer. Simulat. **18** (2013) 938–945. doi:10.1016/j.cnsns.2012.08.038.

[193] M. Ciszak, F. T. Arecchi, S. Euzzor, R. Meucci, *Experimental characterization of the dynamics in a network of chaotic FitzHugh-Nagumo neurons*, in: 2014 Complexity in Engineering (COMPENG), 2014, pp. 1–6. doi:10.1109/CompEng.2014.6994677.

[194] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, K. Ugurbil, The WU-Minn Human Connectome Project: An overview (2013). doi:10.1016/j.neuroimage.2013.05.041.

[195] See https://db.humanconnectome.org/ for public research data.

[196] A. Perinelli, D. E. Chiari, L. Ricci, *Correlation in brain networks at different time scale resolution*, Chaos **28** (2018) 063127. doi:10.1063/1.5025242.

[197] A. Perinelli, D. Tabarelli, C. Miniussi, L. Ricci, *Dependence of connectivity on geometric distance in brain networks*, Sci. Rep. **9** (2019) 13412. doi:10.1038/s41598-019-50106-2.

[198] M. P. Kennedy, *Robust op amp realization of Chua's circuit*, Frequenz. **46** (1992) 66–80. doi:10.1515/FREQ.1992.46.3-4.66.

[199] H. A. Albuquerque, R. M. Rubinger, P. C. Rech, *Theoretical and experimental time series analysis of an inductorless Chua's circuit*, Physica D **233** (2007) 66–72. doi:10.1016/j.physd.2007.06.018.

[200] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith, D. C. Van Essen, *A multi-modal parcellation of human cerebral cortex*, Nature **536** (2016) 171–178. doi:10.1038/nature18933.

List of publications

The list of publications that I co-authored during my Ph. D. is reported below (ordered by topic). For journal articles the abstract is also reported.

Pertaining this thesis

A. Perinelli and L. Ricci,
Identification of suitable embedding dimensions and lags for time series generated by chaotic, finite-dimensional systems,
Phys. Rev. E **98** (2018), 052226, doi:10.1103/PhysRevE.98.052226

In the field of nonlinear dynamics, many methods have been proposed to tackle the issue of optimally setting embedding dimension and lag in order to analyze sampled scalar signals. However, intrinsic statistical uncertainties due to the finiteness of input sequences severely hinder a general solution to the problem. A more achievable approach consists of assessing sets of dimension and lag pairs that are equivalently suitable to embed a time series. Here we present a method to identify these sets of embedding pairs, under the hypothesis that the time series of interest is generated by a chaotic, finite-dimensional dynamical system. We first introduce a "distance gauge transformation" based on the analytical forms of correlation integrals corresponding to a Gaussian white noise source. We show that in this new distance gauge, correlation integrals generated by chaotic, finite-dimensional systems are characterized by distinctive features, whose absence is a marker of the unsuitability of the underlying embedding choice. By means of a new estimator of the correlation dimension that relies on the new distance gauge, sets of suitable embedding pairs are finally identified by looking at the uniformity of the estimation. The method is completely automatic and was successfully tested on both synthetic and experimental time series. It also provides a tool to estimate the redundance and irrelevance timescales of the system that underlie the time series as well as a lower constraint to the sampling rate. The method is suitable for applications in research fields where a chaotic behavior has to be identified, such as neuroscience, geophysics, and economics.

L. Ricci, A. Perinelli and M. Franchi,
Asymptotic behavior of the time-dependent divergence exponent,
Phys. Rev. E **101** (2020), 042211, doi:10.1103/PhysRevE.101.042211

The divergence rate method, which is used to determine the maximum Lyapunov exponent out of time series, is based on the evaluation of the time-dependent divergence exponent. For chaotic systems and in the small time regime, this exponent grows linearly in time. The asymptotic regime is instead characterized by a time-independent behavior due to the system eventually losing its memory of the starting conditions. The amplitude of this "plateau"-like divergence exponent depends both on the choice of the embedding dimension and lag and on the maximum distance of nearby starting trajectories in a way that is characteristic of the underlying dynamical system. In this paper, upon introducing the basic mathematical tools, we address the plateau evaluation for two classes of time series, those generated by a white noise source and those generated by a finite-dimensional chaotic system. The different behavior provides a novel tool to distinguish purely stochastic sources from deterministic ones, as well as to provide a precise estimate of the correlation dimension in the latter case. The method is also sensitive to correlated noise sources.

A. Perinelli and L. Ricci,
Embedding-Dependent, Full Scale Characterization of Sample Correlation Integrals,
in Chaos 2018 - Book of Abstracts - 11 th Chaotic Modeling and Simulation International Conference: Springer proceedings in complexity. - ISBN: 978-3-030-15296-3. Proceedings of: Chaos 2018 - 11 th Chaotic Modeling and Simulation International Conference, Rome, Italy, 05th-08th June 2018. (2018), p. 89-89..

Parallel to the thesis

Spike train sequences

L. Ricci, M. Castelluzzo, L. Minati and A. Perinelli,
Generation of surrogate event sequences via joint distribution of successive inter-event intervals,
Chaos **29** (2019), 121102, doi:10.1063/1.5138250

The study of many dynamical systems relies on the analysis of experimentally-recorded sequences of events for which information is encoded in the sequence of interevent intervals. A correct interpretation of the results of the application of analytical techniques to these sequences requires the assessment of statistical significance. In most cases, the corresponding null-hypothesis distribution is unknown, thus forbidding an evaluation of the significance. An alternative solution, which is efficient in the case of continuous signals, is provided by the generation of surrogate data that share statistical and spectral properties with the original dataset. However, in the case of event sequences, the available algorithms for the generation of surrogate data can become cumbersome and computationally demanding. In this work, we present a new method for the generation of surrogate event sequences that relies on the joint distribution of successive interevent intervals. Our method, which was tested on both synthetic and experimental sequences, performs equally well or even better than conventional methods in terms of interevent interval distribution and autocorrelation while abating the computational time by at least one order of magnitude.

A. Perinelli, M. Castelluzzo, L. Minati and L. Ricci,
SpiSeMe: A multi-language package for spike train surrogate generation,
Chaos **30** (2020), 073120, doi:10.1063/5.0011328

Many studies in nonlinear science heavily rely on surrogate-based hypothesis testing to provide significance estimations of analysis results. Among the complex data produced by nonlinear systems, spike trains are a class of sequences requiring algorithms for surrogate generation that are typically more sophisticated and computationally demanding than methods developed for continuous signals. Although algorithms to specifically generate surrogate spike trains exist, the availability of open-source, portable implementations is still incomplete. In this paper, we introduce the SpiSeMe (Spike Sequence Mime) software package that implements four algorithms for the generation of surrogate data out of spike trains and more generally out of any sequence of discrete events. The purpose of the package is to provide a unified and portable toolbox to carry out surrogate generation on point-process data. Code is provided in three languages, namely, C++, Matlab, and Python, thus allowing straightforward integration of package functions into most analysis pipelines.

L. Minati, H. Ito, A. Perinelli, L. Ricci, L. Faes, N. Yoshimura, Y. Koike and M. Frasca,
Connectivity Influences on Nonlinear Dynamics in Weakly-Synchronized Networks: Insights From Rössler Systems, Electronic Chaotic Oscillators, Model and Biological Neurons,
IEEE Access **7** (2019), 174793–174821, doi:10.1109/ACCESS.2019.2957014

Natural and engineered networks, such as interconnected neurons, ecological and social networks, coupled oscillators, wireless terminals and power loads, are characterized by an appreciable heterogeneity in the local connectivity around each node. For instance, in both elementary structures such as stars and complex graphs having scale-free topology, a minority of elements are linked to the rest of the network disproportionately strongly. While the effect of the arrangement of structural connections on the emergent synchronization pattern has been studied extensively, considerably less is known about its influence on the temporal dynamics unfolding within each node. Here, we present a comprehensive investigation across diverse simulated and experimental systems, encompassing star and complex networks of Rössler systems, coupled hysteresis-based electronic oscillators, microcircuits of leaky integrate-and-fire model neurons, and finally recordings from in-vitro cultures of spontaneously-growing neuronal networks. We systematically consider a range of dynamical measures, including the correlation dimension, nonlinear prediction error, permutation entropy, and other information-theoretical indices. The empirical evidence gathered reveals that under situations of weak synchronization, wherein rather than a collective behavior one observes significantly differentiated dynamics, denser connectivity tends to locally promote the emergence of stronger signatures of nonlinear dynamics. In deterministic systems, transition to chaos and generation of higher-dimensional signals were observed; however, when the coupling is stronger, this relationship may be lost or even inverted. In systems with a strong stochastic component, the generation of more temporally-organized activity could be induced. These observations have many potential implications across diverse fields of basic and applied science, for example, in the design

of distributed sensing systems based on wireless coupled oscillators, in network identification and control, as well as in the interpretation of neuroscientific and other dynamical data.

Networks out of time series

A. Perinelli, D. E. Chiari and L. Ricci,
Correlation in brain networks at different time scale resolution,
Chaos **28** (2018), 063127, doi:10.1063/1.5025242

Assessing brain connectivity makes up a major issue in the field of network dynamics and neuroscience. Conventional experimental techniques are based on functional imaging and magnetoencephalography, allowing to reconstruct the activity of relatively small brain volume elements. A common approach to identify networks consists in singling out sets of elements that maintain a correlated activity over time. Despite the general consensus that these networks are detectable on a time window of 10 s, no study is presently available on the distribution and thus the reliability of this time scale. In this work, we describe a new method to assess time scales on which correlations between network elements occur and to consequently identify the underlying network structures. The analysis relies on the evaluation of quasi-zero-delay cross-correlation between power sequences associated with distinct volume elements. By changing the width of the running window used to analyze successive segments of time series, the behavior of cross-correlation at different time scales was investigated. The onset of connectivity was estimated to be observable at about 30 s. The method was applied to a set of volume elements that are supposed to belong to a known resting-state network, namely the Default Mode Network. Fully connected networks were identified, provided that a sufficiently long time scale is considered. Our method makes up a new tool for the investigation of the temporal dynamics of networks.

A. Perinelli, D. Tabarelli, C. Miniussi and L. Ricci,
Dependence of connectivity on geometric distance in brain networks,
Scientific Reports **9** (2019), 13412, doi:10.1038/s41598-019-50106-2

In any network, the dependence of connectivity on physical distance between nodes is a direct consequence of trade-off mechanisms between costs of establishing and sustaining links, processing rates, propagation speed of signals between nodes. Despite its universality, there are still few studies addressing this issue. Here we apply a recently–developed method to infer links between nodes, and possibly subnetwork structures, to determine connectivity strength as a function of physical distance between nodes. The model system we investigate is brain activity reconstructed on the cortex out of magnetoencephalography recordings sampled on a set of healthy subjects in resting state. We found that the dependence of the time scale of observability of a link on its geometric length follows a power-law characterized by an exponent whose extent is inversely proportional to connectivity. Our method provides a new tool to highlight and investigate networks in neuroscience.

A. Perinelli and L. Ricci,
NetOnZeroDXC: A package for the identification of networks out of multivariate time series via zero-delay cross-correlation,
SoftwareX **10** (2019), 100316, doi:10.1016/j.softx.2019.100316

The identification of networks is an issue that is relevant in many research fields. Assessing links between nodes of a network by analyzing time series associated to those nodes is a crucial topic in fields as diverse as neuroscience and climate research. In this work we present NetOnZeroDXC, a package that implements a recently published method for the assessment of networks out of multivariate time series. The method relies on the computation of zero-delay cross-correlation between pairs of time series. The software implements the various stages of the algorithm in a general way, thus making up a novel tool that can be used for network analysis in different fields.

A. Perinelli, D. Tabarelli and L. Ricci,
Are observable links between brain nodes due to geometric proximity?,
in Second International Summer Institute on Network Physiology (ISINP). Proceedings of: ISINP, Como, 28 July 2019 - 02 August 2019. (2019), .

Bioinformatics

M. Castelluzzo, A. Perinelli, S. Detassis, M. A. Denti and L. Ricci,
MiRNA-QC-and-Diagnosis: An R package for diagnosis based on MiRNA expression,
SoftwareX **12** (2020), 100569, doi:10.1016/j.softx.2020.100569

The possibility of using microRNA (miRNA) levels as diagnostic and prognostic tools to detect different pathologies requires the implementation of reliable classifiers, whose training and use call for quality control of data corresponding to miRNA expression. In this work we present the MiRNA-QC-and-Diagnosis package. The package provides a set of functions for the R environment that implement the required quality control steps and thereupon allow to train, use and optimize a Bayesian classifier for diagnosis based on the measured miRNA expressions. The package thus makes up a complete and dedicated analytical toolbox for miRNA-based diagnosis.

Experimental design & measurement

N. Pace, L. Ricci, M. Scotoni, A. Perinelli and J. Jovicich,
Characterization of time-varying magnetic fields and temperature of helium gas exit during a quench of a human magnetic resonance system,
Biomedical Physics & Engineering Express **5** (2019), 045021, doi:10.1088/2057-1976/ab2300

The quench of a human magnetic resonance imaging system is a critical event that may occur spontaneously, as an accident or purposely in response to an emergency. Although a magnet's quench presents its own risks, little experimental data is available in this respect. In this study, the programmed quench of a human MRI scanner was used to measure the induced time varying magnetic fields ($\mathrm{d}B/\mathrm{d}t$*) inside the bore in order to evaluate cardiac stimulation risks during a quench. Additionally, we measured the exit temperature of the helium gas, to evaluate potential implications in quench pipe design. The maximum* $\mathrm{d}B\mathrm{d}t$ *was* $360\ mTs^{1}$ *at the center of the magnet, far below the cardiac stimulation threshold (*$20\ Ts^{-1}$*. The helium exit temperature reached* $35\ K$*, perhaps implying further considerations about quench pipe designs. Replication of similar experiments on programmed quenches, specially in high-field MRI systems, will be useful to further characterize quench risks.*

A. Perinelli, F. Finotti, A. M. Tonelli, L. Ricci and R. Albatici,
Experimental apparatus for the determination of thermal conductivity and humidity in building materials by means of electrical permittivity measurements,
Tema **5** (2019), 29–41, doi:10.17410/tema.v5i2.226

Relationship between thermal conductivity and electrical permittivity for different building materials is defined. A novel test apparatus for laboratory activity has been designed and realized based on a special capacitive sensor where the material sample acts as dielectric. The first measurement approach has been refined by considering parallel measurements consisting in the comparison of the frequency of two oscillators. Electrical permittivity shows to be well correlated to the moisture content and variation of the material under investigation.

Acknowledgements

My profound gratitude goes to my supervisor Prof. Leonardo Ricci, for his unrelenting guidance and his instructive advices. I can undoubtedly say that most of what I know, I learnt from him: I am not just talking about *knowledge*, but also about a *view* of what Physics is. I am also grateful to him for the inclusion in many parallel research projects, which allowed me to glimpse at different topics, and—not least—for involving me into teaching.

I wish to express my gratitude to the members (present and former) of the NSE laboratory, Michele, Diana and Davide, whom I have pleasurably and fruitfully worked with, sharing both successes and failures. I also want to thank people from other laboratories with whom I shared countless lunches and enjoyable moments: above all, Matteo, Alessandro, Cesare, Luca Matteo, and Damiano.

Among the friends and relatives belonging to the world outside the University, there are a few that I want to mention here. First and foremost, my parents, who I thank for their inestimable support. Pietro and Davide, for the delightful moments of foolery and science debates (often mixed together). And finally, a wholehearted thank to Fabio and Giulia, who know my darkest and brightest secrets.