



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DEPARTMENT OF STATISTICAL SCIENCES
Second Cycle Degree in Statistical Sciences

**Chaos Detection in Human EEGs:
A Lyapunov Exponents Approach**

Defended by:
Stefano Cartabia

Supervisor:
Alessandra Luati
Co-supervisor:
Simone Giannerini

II Session
Academic Year 2023/2024

Abstract

The high dimensionality and nonlinearity of human brain dynamics raise the question of whether electroencephalograms (EEGs) exhibit chaos. Since the mid-1980s, research has focused on the estimation of the Lyapunov exponent, an attractor invariant, that measures a fundamental property of chaos: sensitivity to initial conditions. However, the conflicting needs for long, stationary, and noise-free EEGs, together with a lack of statistically rigorous estimation methods, have led to inconclusive findings.

This dissertation investigates the existence of global chaotic dynamics and/or local instability in human EEGs by consistently estimating the maximum Lyapunov exponent through an indirect nonparametric neural network approach, which supports a statistical chaos test. The analysis of the UKB dataset, provided by the Epilepsy Center at the University of Bonn, reveals that, unlike normal brain activity, both seizures and interictal phases are associated with a significant reduction in system complexity. Moreover, global chaotic dynamics are identified during epileptic episodes, corroborating a theory previously proposed in the literature but not conclusively proven.

Key words: Chaos theory, Nonlinear time series, Lyapunov exponent, Electroencephalogram (EEG), Epilepsy

Contents

1	Introduction	1
2	Lyapunov exponents: A characterisation of initial state sensitivity	3
2.1	Deterministic Chaotic systems	4
2.1.1	Lyapunov exponents	5
2.1.2	Maximum Lyapunov exponent estimation	7
2.1.3	Local Lyapunov exponents	15
2.2	Stochastic Dynamical Systems	16
2.2.1	Initial value sensitivity	18
2.2.2	Noise amplification and non-monotonicity of prediction accuracy	21
3	Nonlinear Brain Dynamics: A literature review	23
3.1	Nonlinear EEG analysis	23
3.2	UKB dataset and related literature	28
4	Chaos detection in human EEGs: A Lyapunov exponents Approach	32
4.1	Preliminary Analysis	32
4.2	Nonlinear Time Series Analysis	35
4.2.1	R Software Implementation	37
4.2.2	System complexity	38
4.2.3	Global chaotic dynamics	38
4.2.4	Local chaotic dynamics	41
4.3	Conclusion	43

Chapter 1

Introduction

The striking idea that deterministic systems can produce highly erratic and never-repeating oscillating solutions has fundamentally reshaped our perception of concepts like randomness and chance. The discovery that nonlinear dynamical systems with at least three degrees of freedom can exhibit completely unpredictable behaviour traces back to the late 1880s, with the gravitational studies of the French mathematician Henri Poincaré. However, it was only with the meteorologist Edward Lorenz, and his 1963 nonlinear model for atmospheric convection, that the chaos theory revolution actually began, deeply influencing many fields of science [C. J. Stam, 2005].

The study of chaotic systems has revealed surprising connections with the mathematical frontier, in particular with fractal geometry [Mandelbrot & Wheeler, 1983; Ruelle & Takens, 1971], and has highlighted two fundamental characteristics of chaos. First, sensitivity to initial conditions, or the butterfly effect, which makes even tiny errors grow exponentially fast and prevents any long-term prediction [Eckmann & Ruelle, 1985]. Second, chaotic dynamics reveal that determinism and randomness can coexist, inevitably invoking the philosophical question of whether randomness actually exists as a distinct paradigm.

The human brain, with its vast interconnected network of about 80 billion neurons, embodies one of the most complex systems in nature and represents an emblematic field of application of chaos theory. Indeed, the high dimensionality and nonlinearity of brain dynamics, even at a neural level, raise the question of whether brain electrical activity recordings — the electroencephalograms, EEGs — actually exhibit chaotic behaviours. Moreover, since the mid-1980s, the limitations of both parametric and nonparametric linear signal processing tools in handling intricate nonlinear dynamics have spurred a growing interest in applying chaos theory methods to EEG analysis, revealing a considerable reduction of the system complexity for various mental states and neurological conditions, such as eye closure, alcohol consumption, schizophrenia, Parkinson's, and epilepsy [Babloyantz & Destexhe, 1986; Lee et al., 2001; Müller et al., 2001]. Nonetheless, the conflicting desires for long, stationary, and noise-free EEG signals, together with a lack of statistically appropriate methods, have led to conflicting and inconclusive results, especially for what concerns the presence of global chaotic dynamics [Palus, 1996; Prichard & Theiler, 1994].

In this dissertation, it is first provided a theoretical introduction to the emergence of chaos and the estimation methods of the Lyapunov exponent, an attractor invariant that measures a fundamental characteristic of chaos: sensitivity to initial conditions. Then, in the second chapter, a comprehensive literature review outlines the past re-

search in the field of nonlinear EEG analysis, especially in the context of epilepsy, where a reduction of the system complexity has been associated with the seizure onset, as a result of a widespread brain region hyper-synchronization [Lehnertz & Elger, 1995]. However, although chaotic dynamics have been hypothesized, their presence has not been conclusively proven and various studies [Feucht et al., 1998; Friedrich & Uhl, 1996] suggested instead that ictal stage EEGs behave as noisy limit cycle regimes. For this reason, in the last chapter, we explore the existence of global chaotic dynamic and/or local instability in the human electroencephalograms (EEGs) of the UBK dataset — made available online by the Epilepsy Center at the University of Bonn [Andrzejak et al., 2001] — by consistently estimating the Lyapunov exponent through an indirect neural network approach [Dechert & Gencay, 1992; McCaffrey et al., 1992]. This state-of-the-art method uses a nonparametric regression to first estimate the underlying dynamics and then assess the average rate of divergence of initially nearby trajectories, with the advantages of not requiring a previous attractor reconstruction step and allowing the presence of dynamic noise. Moreover, its consistency properties [Whang & Linton, 1999] allow the development of a proper statistical chaos test, based on the maximum Lyapunov characteristic exponent (MLCE) [Shintani & Linton, 2004]. To preview the results, the overall analysis reveals two main findings. First, it shows a gradual reduction of brain complexity — in terms of embedding dimension — beginning in the healthy brain state, diminishing further in the interictal phase, and reaching its lowest level during the ictal phase. Second, by classifying EEGs into chaotic, non-chaotic, and ambiguous cases, according to the MLCE and the associated p-value, it emerges that ictal stage signals display low-dimensional chaotic regimes. No clear indications have emerged from the detection of chaotic dynamics in normal brain activity, possibly because such dynamics do exist but cannot be detected with the Lyapunov exponents in such high-dimensional state spaces. In the end, local Lyapunov exponents (LLEs) have been estimated, revealing the great complexity and heterogeneity of human brain dynamics.

Chapter 2

Lyapunov exponents: A characterisation of initial state sensitivity

Sensitive dependence on initial conditions is a fundamental property of a wide range of natural phenomena which expresses how infinitesimal perturbations of the initial conditions are exponentially amplified by the high nonlinearity characterising the system. A dramatic implication of this peculiarity, often popularised as the butterfly effect, is that long-term forecasts are completely unreliable, as even tiny numerical approximation errors are amplified enormously fast. Despite there being no universally accepted definition of what a chaotic system is, most definitions agree on the deterministic, aperiodic, and bounded nature of chaotic motion, but above all on its sensitivity to initial conditions [Eckmann & Ruelle, 1985]. Indeed, none of the first three features is actually unique to chaotic systems, even though they provide interesting characterisations of such processes. This is the case of aperiodicity, shared with quasi-periodic motions, which explains how each point — state — is only crossed once and how the system trajectories never repeat, without converging to a fixed point or a periodic orbit.

The dual essence of chaos, as both deterministic and random, has raised the empirical challenge of distinguishing chaos from chance [Theiler et al., 1992], inevitably invoking a deep philosophical question on whether randomness actually exists as a distinct paradigm.

Moreover, the study of chaotic systems has revealed surprising connections with the mathematical scientific frontier, particularly the fact that most, though not all, chaotic attractors exhibit a fractal structure [Mandelbrot & Wheeler, 1983; Ruelle & Takens, 1971]. However, the existence of strange nonchaotic attractors [Grebogi et al., 1984] has revealed the fundamental importance of initial value sensitivity in detecting global chaotic dynamics. A quantitative indication of the degree of exponential divergence of initially nearby trajectories is provided by a positive estimate of the Lyapunov exponent, which will be the object of this dissertation and especially of the present chapter. The present chapter offers a theoretical introduction to the genesis of chaos in deterministic systems, focusing on the definition of the Lyapunov exponent, methods for estimating both global and local Lyapunov exponents, as well as their extension to stochastic dynamical systems.

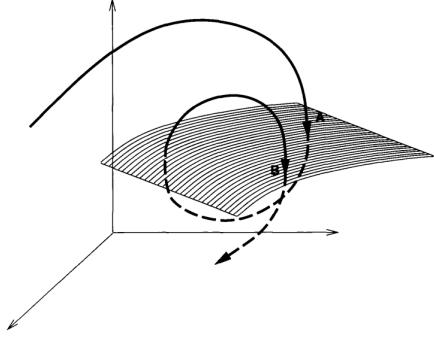


Figure 2.1: Example of the Poincaré Section for a 3d flow. Image credit: [1994]

2.1 Deterministic Chaotic systems

From an algebraic perspective, chaos may either arise from an m -dimensional nonlinear map, as exemplified by the famous logistic map, or from an autonomous system of m first-order ordinary differential equations (ODEs), to which any higher-order nonlinear problem can be reduced to, as the historically groundbreaking Lorenz's system of equations for atmospheric convection [Kantz & Schreiber, 2004]. The former case will then be regarded as a discrete dynamical system

$$x_{n+1} = F(x_n) \quad F : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad n \in \mathbb{Z} \quad (2.1)$$

and it can be rewritten as a difference equation of the form $\Delta x_n = g(x_n)$ by subtracting on both sides the state parameter x_n . In contrast, the latter case describes a dynamical system that is said to be continuous, as a consequence of the time derivative involvement in its definition

$$\dot{x}(t) = f(x(t)) \quad f : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad t \in \mathbb{R} \quad (2.2)$$

and we often refer to this system as a flow.

For the sake of convenience, we will refer from now on to the discrete case. However, this does not involve any loss of generality since any map of type (2.1) can be regarded as the projection — known as the Poincaré section ⁽ⁱ⁾ — of the flow (2.2) on a $(m-1)$ -hyperplane, depicted in Figure 2.1 [Baker & Gollub, 1996]. In addition, another crucial aspect, which will be discussed in greater detail in Section 2.2, is that (2.1) can be interpreted as the deterministic counterpart of a Markov process, whose transition kernel $P(x_{t+1}|x_t)$ is the Dirac delta function and in which the future state is fully determined by the current one, through the nonlinear function $F(x_t)$ [Chatterjee & Yilmaz, 1992].

The mathematical assessment of the trajectory stability is a widely studied topic in the field of dynamical systems and quantitative information can often be derived by studying the linearised system near the orbits, as well as by applying topological and analysis tools to explore the state space \mathcal{M} — essentially, the Euclidean space \mathbb{R}^m or the m -dim manifold containing the set of all possible states assumed by the dynamical system⁽ⁱⁱ⁾. Of particular interest is a specific region of the state space known as the

⁽ⁱ⁾Conversely, through integration, one can reconstruct the flow $x_{n+1} = x_n + \int_t^{t+T_n} F(x(t))dt$, where T_n is the returning time to the hyperplane at the n -th map iteration [Liu, 2010].

⁽ⁱⁱ⁾The motion of a particle in three dimensions requires a six-dimensional space, three for the positions and three for the velocity in each direction [Baker & Gollub, 1996]

attractor, $\mathcal{A} \subset \mathcal{M}$, which is a bounded set satisfying for a deterministic dynamical system of type 2.1 or 2.2 the following conditions: i) $\forall X \in \mathcal{M}$ then $F(X) \in \mathcal{A}$, so the set is invariant; ii) $\forall X_0$ sufficiently closed to \mathcal{A} then $F^t(X_0) \rightarrow \mathcal{A}$ as $t \rightarrow \infty$. An attractor can take the form of a fixed point, limit cycle, torus or fractal structure⁽ⁱⁱⁱ⁾ and it describes the asymptotic behaviour of a dynamical system, generally after an initial transient regime.

In real-world applications, however, the empirical determination of system stability is much more intricate as we are generally unaware of the equations governing a process, and the only information comes from time series recordings — discrete measurements, often in a scalar form. Therefore, the study of the attractor invariants becomes crucial and, for this purpose, a preliminary step is the so-called phase space embedding, which essentially transforms the time series into a collection of ordered state-space vectors that reproduce the process dynamics. Technicalities about the embedding method are unfortunately far beyond the scope of this chapter, and actually of the whole thesis. Nevertheless, it is worth understanding that since we can only interact with the original attractor, defined in the original — physical — state space, through its squashed projections — the measured signal — the aim is to derive a topologically fair copy of the attractor in a newly defined state space — known as reconstructed state space [Sauer, 2006]. Fortunately, two important theorems come to the aid. First, Whitney's theorem guarantees that any smooth d -dim manifold (\mathcal{M}) can be unfolded into a higher dimensional Euclidean space \mathbb{R}^{2d+1} through a diffeomorphism — a one-to-one C^1 map, with full rank Jacobian — that avoids the trajectory crossing and preserves the tangent space structure — but not the geometric shape. Second, Takens' embedding theorem states that a strange attractor — a subset of the state space manifold, $\mathcal{A} \subset \mathcal{M}$ — can be embedded without the intersection of the reconstructed trajectories in a d -dimensional Euclidean space, using the time-delayed versions of the measured signal $[y(t), y(t - \tau), y(t - 2\tau), \dots, y(t - 2d\tau)]$, where τ is the time delay and d the embedding dimension such that $d \geq 2d_a$, with d_a denoting the box-counting dimension [Sauer et al., 1991]. As a whole, if the proper embedding space is chosen, studying the reconstructed attractor^(iv) will be completely equivalent to studying the original one, as invariants are preserved with respect to diffeomorphisms.

Despite the lack of a universally accepted definition of a chaotic system, nonlinear continuous dynamical systems with at least three degrees of freedom generally exhibit chaos [Schuster, 1984], while discrete dynamical systems may present chaos at any dimensionality, as in the case of the logistic maps. However, from an operational perspective, it is worth referring to the definition of [Eckmann & Ruelle, 1985] which states that a chaotic system is a bounded system — converging to an attractor — with a positive Lyapunov exponent, and so displaying sensitive dependence on initial conditions, which will be the focus of the following subsection.

2.1.1 Lyapunov exponents

Up to this point, we have briefly introduced the paradigm of chaos and the empirical challenge of inferring notions about a dynamical system based on time series data.

⁽ⁱⁱⁱ⁾Chaotic attractors, also known as strange attractors [Ruelle & Takens, 1971], exhibit a fractal structure, meaning they have a non-integer dimension, which is a measure of the system complexity.

^(iv)For the right choice of the time delay τ and the embedding dimension d , the reconstructed dynamics are diffeomorphic to the original one, meaning they preserve the same topology.

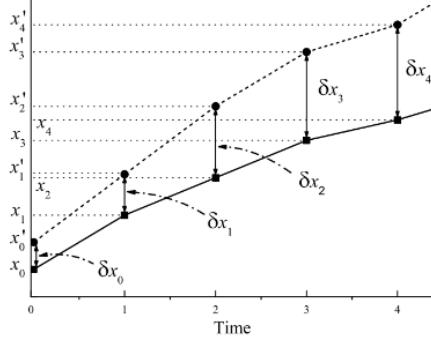


Figure 2.2: Exponential divergence of nearby trajectories. Image credit: [2018]

Now, our interest is devoted to the study of a particularly striking feature of chaotic processes: the initial value sensitivity, which explains the exponential divergence of trajectories subject to an initial small perturbation.

Let's consider the dynamical system $x_t = F(x_{t-1})$ with $x = (x_1, \dots, x_m) \in \mathbb{R}^d$ being the generic state vector lying in the phase-space \mathcal{M} , and $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a real-value function of the class C^2 — so continuous up the second partial derivatives. Then, given a small perturbation, δ_0 , applied on the initial conditions, $x_0 \in \mathbb{R}^m$, we can study the stability of the two initially closed orbits by looking at the linearised dynamics. Indeed, according to the Taylor series expansion, after one step

$$\begin{aligned} x_1 + \delta x_1 &= F(x_0 + \delta x_0) \\ &\approx DF(x_0)\delta x_0 + F(x_0) \end{aligned} \tag{2.3}$$

where $DF(x)_{a,b} = \frac{\partial F_a(x)}{\partial x_b}$ is the $m \times m$ Jacobian associated to the orbit described by F . By applying the chain rule, we get that after a generic number of n steps ahead

$$\begin{aligned} x_n + \delta x_n &= F^n(x_0 + \delta x_0) \\ &\approx DF^n(x_0)\delta x_0 + F^n(x_0) \end{aligned} \tag{2.4}$$

where F^n is the n -fold composition, and the degree of separation between the orbits amounts to

$$\begin{aligned} \delta x_n &= F^n(x_0) - F^n(x_0 + \delta x_0) \\ &\approx DF(x_{n-1}) \times DF(x_{n-2}) \times \cdots \times DF(x_0)\delta x_0 \\ &\approx DF^n(x_0)\delta x_0 \end{aligned} \tag{2.5}$$

From this last equation, we can already intuit that the rate of divergence of two initially nearby trajectories depends on the Jacobian matrix $DF^n(x_0)$ of the map F and if there exist one — or even more — positive eigenvalue then there will be an exponential growth of the initial perturbation in the associated eigendirection — at a rate of $e^{n\lambda}$ with $\lambda > 0$ [H. Abarbanel, 2012]. More rigorously, with this regard, the multiplicative ergodic theorem — also known as the Oseledec theorem — states that given

$$|\delta x_n|^2 = \delta^T \times (DF^n(x))^T \times DF^n(x) \times \delta \tag{2.6}$$

$$O(x_0, n) = (DF^n(x_0) \times DF^n(x_0))^{\frac{1}{2n}} \quad (\text{Oseledec matrix})$$

it not only exists the limit of $O(x_0, n)$ for $n \rightarrow \infty$ and for almost all initial conditions with respect to the natural measure p but if p is ergodic, then this limit is also independent of x_0 p -almost everywhere — so for almost every initial vector in the basin of

attraction^(v) [Eckmann & Ruelle, 1985]. As a result, we can define the m -dim vector of the average rates of compression in the stable manifold directions and of stretching in the unstable manifold as follows:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \log |DF^n \delta x(0)| \quad (2.7)$$

The elements in this vector correspond to the logarithmic eigenvalues of the asymptotic Oseledec matrix [Giannerini & Rosa, 2002], which are the global Lyapunov exponents and the highest of these exponents is known as the maximum Lyapunov characteristic exponent (MLCE). The collection of these exponents is often called the Lyapunov spectrum, and it is invariant with respect to coordinate transformations (e.g. rescaling or shifting the data). Moreover, the Lyapunov spectrum satisfies certain properties, in the case of a chaotic dissipative system^(vi) the sum of the exponents must be negative — $\sum_{i=1}^m \lambda_i < 0$, in accordance with the general loss of energy in the system and the state space measure reduction — and there should be at least one positive exponent. It may now be surprising to notice that whenever the spectrum exhibits at least two positive Lyapunov exponents, there emerge even more intricate dynamics known as hyperchaotic, which are characterised by instability in multiple directions — the minimum dimension for such systems is four [Rossler, 1979]. Unfortunately, even if interesting questions may arise, especially in relation to the concept of randomness, we prefer instead to conclude this subsection with a deeper visual understanding of the Lyapunov spectrum in dissipative systems; see [Chatterjee & Yilmaz, 1992] for a more detailed explanation.

Let's consider $\{\mathbf{x} : \|\mathbf{x}\|_2 \leq P_0\} \subset \mathbb{R}^m$, a hypersphere of initial conditions with radius P^0 and belonging to the phase space. After n iterations of the map, the hypersphere will be shaped into an ellipsoid whose i -th axis has length P_i^n while its hypervolume $\det(DF^n)$ will be strictly smaller than the one of the initial hypersphere, in line with the system's overall energy dissipation.

Then, by assuming P_i^n/P^0 to be the proportion of change in the length of the i -th axis, the associated i -th Lyapunov exponent will precisely be the mean of the logarithmic rates of change

$$\lambda_i = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{P_i^n}{P_0} \right)$$

which in limit explains the compression or stretching force — depending on the sign — exerted over the hypersphere in such a direction.

2.1.2 Maximum Lyapunov exponent estimation

As previously anticipated, the stability analysis and the estimation of the Lyapunov spectrum for a dynamical system with known motion equations is fairly straightforward [Benettin et al., 1980; Shimada & Nagashima, 1979]. Unfortunately, when dealing with experimental data, this task becomes much more intricate as the system

^(v)The *basin of attraction* associated to a given attracting set is the collection of all the initial conditions in the phase space whose trajectories asymptotically converge to the attracting set [Nusse & Yorke, 1998; Ott, 2006]

^(vi)*Dissipative systems* are in contrast to *conservative system*, for which the Liouville's theorem states that the volume of their phase space \mathcal{M} is preserved over time [Eckmann & Ruelle, 1985]

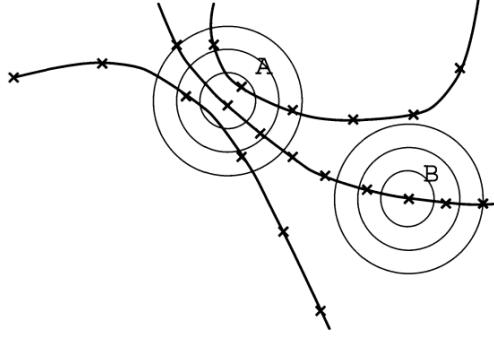


Figure 2.3: The evaluation of the exponential divergence of nearby trajectories should exclude points within the time neighbourhood $|i - j| < n_{min}$ to avoid temporal correlation. For point B, all neighbouring points are direct images or pre-images of the starting point, potentially biasing estimates towards lower values. Image credit: [1994]

dynamics cannot be studied in the original phase space but rather in the reconstructed one. As a consequence, any algorithm for estimating the Lyapunov exponents requires a phase-space reconstruction procedure — which can either take place before or simultaneously with the Lyapunov exponents estimation. In this regard, it is important to note that, due to the higher dimensionality of the reconstructed state space ($m < d$), the cardinality of Lyapunov spectrum associated with the newly defined state space will be larger than that of the spectrum associated with the unknown physical attractor. These additional exponents are called spurious, and there exist various techniques for their identification [Hegger et al., 1999]. Unfortunately, according to the various techniques proposed so far, we are limited to estimating the largest exponent of the Lyapunov spectrum, the MLCE. We will now proceed to introduce the two major classes of algorithms for MLCE estimation.

2.1.2.1 Direct Methods

From a historical perspective, [Wolf et al., 1985] provides the first MLCE estimation procedure for time series data, which is based on tracking the separation of infinitesimally close trajectories along the previously reconstructed attractor. More precisely, Wolf's method works by choosing two closed initial conditions x_0 and x_{e0} , such that $\|x_{e0} - x_0\| = \epsilon_0 < \epsilon_{max}$, and observing how their resulting orbits drift apart. When eventually the distance between these two orbits exceeds a predefined threshold ϵ_{max} , let that distance be denoted as ϵ'_0 and let T_1 represent the number of time steps required. Next, by choosing a new point x_{e1} nearby x_1 , which is the evolution of x_0 on the fiducial trajectory, such that $\|x_{e1} - x_1\| = \epsilon_1 < \epsilon_{max}$, we repeat the last step M times [Kutepov et al., 2020]; see Figure 2.4 for a graphical illustration.

The estimated maximum Lyapunov exponents will then be

$$\lambda \approx \frac{1}{\sum_{k=1}^M T_k} \sum_{k=0}^{M-1} \ln \left(\frac{\epsilon'_k}{\epsilon_k} \right).$$

By considering areas, volumes or hyper-volumes rather than distances, it is theoretically possible to estimate the overall Lyapunov spectrum. Unfortunately, though, even in its basic distance form, Wolf's method is not particularly precise, especially for small

datasets. A reason for this is provided by [Rosenstein et al., 1993] which suggests that Wolf's method focuses on constructing a single fiducial trajectory by considering one nearest neighbour at the time, thus avoiding leveraging all available data. Moreover, the Gram-Schmidt procedure required for determining a new neighbour x_{ϵ_1} that preserved the original orientation brings significant computational burdens. Nevertheless, Wolf's method has definitively paved the way for further improvement of direct estimation approaches [Kantz, 1994; Rosenstein et al., 1993; Sato et al., 1987]. Indeed, Wolf's method suggests that even though in experimental data we cannot perturb the initial conditions, we can still treat time-separated segments of the single trajectory, forming the reconstructed attractor, as distinct trajectories [Frank et al., 1990].

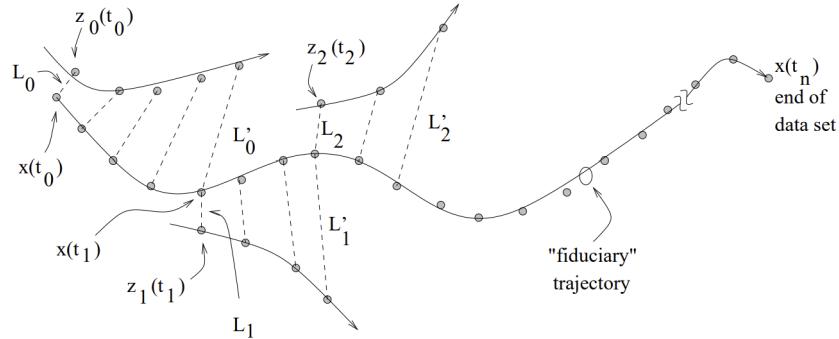


Figure 2.4: Illustration of Wolf's Algorithm for estimating the Lyapunov Exponents from time series data. Image credit: [L. Bradley, Chaotic Dynamics Course Materials]

Rosenstein and Kantz's methods have garnered greater scientific approval and, differently from Wolf's method, they provide a time-dependent estimator — a function of time — instead of a proper scalar estimate of the MLCE. These two algorithms share their general framework with [Sato et al., 1987], and, by relying on the ergodic assumption, they allow choosing any arbitrary direction in the phase-space, since two initial points will then diverge with probability 1 along the most unstable manifold direction, at a rate equal to the maximum Lyapunov exponent. However, it is worth noticing that the differences between these two methods are modest and they essentially reduce to the fact that Kantz's method considers multiple reference points within a neighbourhood, averaging out the orbit divergences.

More specifically, Rosenstein's method is constructed by iteratively computing the average logarithmic difference between the nearest neighbour along a single trajectory such that at each i -th time step

$$\lambda_1(i) = \frac{1}{i \cdot \Delta t} \cdot \frac{1}{(N-i)} \sum_{j=1}^{N-i} \ln \frac{d_j(i)}{d_j(0)}, \text{ with } d_j(0) = \min_{\mathbf{X}_j} \|\mathbf{X}_j - \mathbf{X}_{\hat{j}}\|_2 \quad (2.8)$$

where $\|\cdot\|_2$ is the Euclidean norm, Δt is the sampling period, and N is the number of points in the reconstructed embedding space.

Kantz's method, on the other hand, works by computing the logarithmic average distance between orbits starting within the spherical neighbourhood $\mathcal{U}(x_{n_0}) = \{x_i \mid \|x_i - x_{n_0}\| \leq \epsilon\}$ of the reference point x_{n_0} . Moreover, to thoroughly explore the diverging behaviour

across the entire attractor, this procedure is replicated using N different initial reference points. The general divergence is then

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{|\mathcal{U}(x_{n_0})|} \sum_{x_n \in \mathcal{U}(x_{n_0})} |x_{n_0+\Delta n} - x_{n+\Delta n}| \right) \quad (2.9)$$

where Δn ranges from 1 to a predefined number of steps ahead, N . Much care should be paid to the choice of the spherical neighbour size, indeed a common issue — already highlighted in [Theiler, 1986] — concerns the fact that the estimation procedures, based on the comparison of state vectors (Takens' vectors), may be biased because of temporal correlation. Takens' vectors can in fact be neighbours as a mere consequence of a temporal closeness, rather than a spatial one. Therefore, points belonging to each spherical neighbourhood should be closer than ϵ to x_{n_0} but also be adequately temporally separated. In this regard, [Provenzale et al., 1992] introduced the space-time separation plot, which provides a heuristic for the choice of such a parameter, generally known as Theiler's window. For more technical details, please refer to [Hegger et al., 1999; Kantz & Schreiber, 2004], while for a visual aid Figure 2.3 may be helpful.

An important advantage of considering logarithmic distances is that if the distance between two points grows exponentially $\delta_i(\Delta n) \sim C_i \exp(\lambda \Delta n)$, then the slope of the linear portion in the semilogarithmic scale, $S(\Delta n) = \log \delta_i(\Delta n) \sim C_i \lambda \Delta n$, is precisely an estimator of the MLCE [Huffaker et al., 2018]. Unfortunately, as also occurs in Rosenstein's method, selecting the linear growth region is somewhat arbitrary (Figure 2.5).

In conclusion, direct methods offer some remarkable advantages, such as the fact that they do not require any modelling and the computation burdens are also quite limited. Unfortunately, though, they are limited by the absence of any theoretical result concerning the estimator consistency and asymptotic distribution. Furthermore, they have unsatisfactory performance in detecting nonlinearity for moderately short time series and they are not even robust to measurement noise^(vii), indeed, the presence of any noise term tends to be associated with the detection of global chaotic dynamics.

2.1.2.2 Indirect Methods

There is also a second class of estimators for the MLCE called indirect or Jacobian-based methods. Originally introduced by [Eckmann & Ruelle, 1985; Eckmann et al., 1986], they consistently estimate the exponential divergence of nearby trajectories by first computing the Jacobian matrix along the system orbit and then extracting the Lyapunov exponents as the eigenvalues of the Oseledec matrix, in accordance with the multiplicative ergodic theorem introduced in subsection 2.1.1. Moreover, since indirect methods rely on an intermediary nonparametric estimation, it is worth noticing that as discussed in [Chan & Tong, 2001; Giannerini & Rosa, 2004; Giannerini et al., 2020], the time series embedding can be seen as a statistical subset selection problem in the context of nonparametric regression. Compared to traditional heuristic methods used prior to the actual estimation, this approach thus yields a consistent estimate

^(vii) *Measurement noise* refers to the distortion of the observed values based on an error term that is independent of the system dynamics. On the contrary, *dynamical noise* consists of an actual perturbation which interacts at each time step with the system dynamics, introducing a higher degree of complexity [Kantz & Schreiber, 2004]

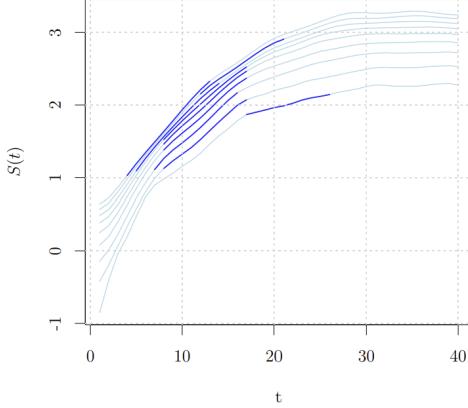


Figure 2.5: Plot on the semilogarithmic scale showing the exponential rate of divergence of nearby trajectories computed using Kantz’s method for the Lorenz attractor time series. The analysis was performed with the `tseriesChaos` R library [Narzo, 2019] for an embedding dimension between 2 to 10, and 40 steps ahead.

of the embedding dimension and the time delay. In addition, unlike direct methods, significant consistency results for the MLCE estimator have been derived for some of the methods belonging to this class, such as the thin plate spline and ANN approaches. However, since in a purely deterministic system there is no way of evaluating the uncertainty surrounding a Lyapunov exponent point estimate, it is included an iid stochastic disturbance to the deterministic skeleton in 2.1 — the proper model will be explained in Section 2.2 — which has allowed — under the knowledge that a positive estimate still indicates sensitive dependence on initial conditions — to first prove the consistency of the nonparametric estimator [McCaffrey et al., 1992; Nychka et al., 1992] and then find its asymptotic distribution [Whang & Linton, 1999]. We prefer, however, to postpone this discussion to Subsection 2.1.3, where the notion of local Lyapunov exponents will be introduced, and a statistical chaos test will be presented [Shintani & Linton, 2004].

Eckmann’s indirect method

The groundbreaking algorithm that has paved the way to the class of indirect methods was proposed by [Eckmann & Ruelle, 1985; Eckmann et al., 1986] and it employs a local linear nonparametric regression approach to estimate the positive exponents. The procedure requires the Takens’ vectors $x(i)$ to be already defined in the reconstructed state space and it provides an estimate of the system Jacobian matrix, $T_{x(i)}$, by solving the following least squares problem:

$$T_{x(i)}[x(j) - x(i)] \approx x(j + p) - x(i + p) \stackrel{\text{def}}{=} \arg \min_T \sum_{i=1}^N (x(i + p) - Tx(i))^2$$

where for a generic number of p -step ahead (often equal to 1), it holds that

$$d[x(i), x(j)] \leq \epsilon \quad \text{and} \quad d[x(i + p), x(i + p)] \leq \epsilon$$

with ϵ being fixed by trial and error, according to the fitting quality.

The difficulty of Eckmann’s indirect method in extracting the non-positive exponents

is addressed by [Brown et al., 1991] to the fractal structure of the chaotic attractors. Indeed, since the attractor does not densely fill the phase space, there may be very few points in the stable manifold directions, associated with negative exponents. This lack of information about the system behaviour produces large inaccuracies in the Jacobian, that are further magnified via the composition $(T_{x(i)}^1)^t$. Nonetheless, soon after Eckmann's method proposal, [Sano & Sawada, 1985] suggested a refinement that mitigates the non-positivity issue by first least squares estimating the map A_j locally, rather than the Jacobian, and then computing the Lyapunov spectrum based on a set of bases e_i , that describes the tangent space at x_i , and whose orthogonality is preserved over time by applying the Gram-Schmidt procedure. Unfortunately, even though this last method theoretically allows the estimation of a wider portion of the Lyapunov spectrum, the linear approximation often lacks precision, and, for this reason, there have been developed more accurate nonlinear approaches.

Polynomial and Thin plate spline regression approach

The attempts to overcome local linear approximation methods have been essentially addressed to develop, on the one hand, local higher order polynomial methods [Brown et al., 1991; Park & Whang, 2012], and, on the other hand, local thin plate spline approaches [McCaffrey et al., 1992; Nychka et al., 1992].

In the former, it is derived a n -th order Taylor expansion to model the divergence between r initially neighbouring trajectories. Namely,

$$\begin{aligned} z^r(n; T) &= y^r(n; T) - y(n + T) \\ &= F(y(n) + z^r(n; 0)) - F(y(n)) \end{aligned}$$

Then, by considering the α -th component of the initial displacement $z^r(n; T)$, we will get that

$$z_\alpha^r(n; T) = DF_{\alpha\beta}(n)z_\beta^r(n; 0) + DF_{\alpha\beta\gamma}^{(2)}(n)z_\beta^r(n; 0)z_\gamma^r(n; 0) + \dots \quad (2.10)$$

where $DF_{\alpha\beta} = \frac{\partial y^\beta}{\partial F^\alpha}$ is an element of the Jacobian, while $D^2F_{\alpha\beta\gamma} = \frac{\partial^2 F^\alpha}{\partial y^\beta \partial y^\gamma}$ one of the Hessian. By truncating the series in 2.10 at a given order, we balance between the accuracy and the computational complexity of the local approximation, knowing that the first-order approximation coincides with Eckmann's indirect method.

An alternative solution to accurately and efficiently approximate nonlinear functions is represented by spline functions. The formal definition of the m -th-order thin plate spline for f_d is the result of the following minimisation problem:

$$\begin{aligned} \mathcal{L}(h) &= \frac{1}{N} \sum_{t=1}^N \{Y_t - h(X_t)\}^2 + \rho \mathcal{J}_{m,d}(h) \quad \text{s.t. } \rho > 0 \\ h : \mathbb{R}^d &\rightarrow \mathbb{R}, \quad \partial^\alpha h = \frac{\partial^\alpha h}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_d^{\alpha_d}} \in L^2(\mathbb{R}^d) \quad \forall \alpha \leq m \\ \text{and } \mathcal{J}_{m,d}(h) &= \sum_{\alpha_1+\dots+\alpha_d=m} \binom{m}{\alpha_1 \dots \alpha_d} \int_{\mathbb{R}^d} \left(\frac{\partial^m h(X)}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right)^2 dX < \infty \end{aligned}$$

More specifically, $\mathcal{J}_{m,d}(h)$ is the rotational invariant regularisation term, that mitigates

— through the associated non-negative coefficient ρ — between bias and variance, in such a way of determining the degree of wiggleness of the fitted map $\hat{f}_d(X_1)$. In a situation with little or no noise ρ tends to be small, while as $\rho \rightarrow \infty$, the estimated map will converge to the interpolating polynomial of degree $m - 1$, whose coefficients are derived through the least square method.

In addition, even though the actual choice of the regularisation coefficient ρ is in principle arbitrary, a generally adopted procedure is the generalised cross-validation (GCV)

$$\rho^* = \arg \min_{\rho} = \frac{1}{N} \frac{\|(I - A(\rho))Y\|^2}{(1 - \text{trace}\{A(\rho)\}/N)^2}$$

which tends to minimise the expected average squared error (EASE). Unfortunately, in some noisy cases, the GCV may undersmooth the solution, bringing ρ close to zero. For this reason, [1992] proposes a modified version which assigns a higher weight to larger values of ρ .

Overall, whereas from a geometrical perspective, the minimisation solution is a set of surfaces, algebraically, it is a linear combination of $\binom{d+m-1}{d}$ monomials of degree $m - 1$, N radial basis functions and a set of coefficients that are linearly related to the observed variable vector $Y = (y_1, \dots, y_N)^T$. Thus, the estimated function can be expressed as $(\hat{f}_\rho(X_1), \hat{f}_\rho(X_2), \dots, \hat{f}_\rho(X_N))^T = A(\rho)Y$, where $A(\rho)$ is $N \times N$ matrix which depends on: i) the regularisation coefficient ρ ; ii) the order m of the derivatives in the regularisation term; iii) the input vector $X_t = (y_t, \dots, y_{t-d+1})$ and its dimensionality d — that is exactly the embedding dimension.

In conclusion, we would like to end this subsection by introducing the most advanced indirect method, which exploits an artificial neural network for the estimation of the underlying nonlinear map. For further technical details and real-world application of the thin plate spline, we suggest the previously mentioned papers and Section 5.2 of the book [Chan & Tong, 2001]

Nonparametric Regression with ANN

In recent years, neural network architectures have increasingly grown in popularity, and besides their wide application, such as in language models or image processing, artificial neural networks (ANN) find also their place in the field of nonlinear dynamics and chaos. Concerning the topics of this chapter, from the independent work of [Dechert & Gencay, 1992] and [McCaffrey et al., 1992], based on an original study of [M. Casdagli, 1989], it has emerged that ANNs provide a state-of-the-art method for the estimation of the multidimensional map underlying a given dynamical system, guaranteeing a consistent method for the MLCE estimation.

The main advantages of this method come from the fact that ANNs theoretically provide the best approximation method for whatever continuous function — universal approximation theorem — and also for its first order derivatives [Shintani & Linton, 2004]. Moreover, the ANN-based method is more robust as it filters out dynamic noise [Giannerini & Rosa, 2004]

The architecture that has been studied the most in this context is the single-hidden-layer feed-forward network (SLFN), which has the following functional form

$$\hat{f}_d(Z_1, \dots, Z_d) = \beta_0 + \sum_{j=1}^q \beta_j \psi \left(\sum_{i=1}^d W_{ij} Z_i + W_{0j} \right)$$

s.t. $(z_1, \dots, z_n)^T \in \mathbb{R}^d$ and ψ is a given activation function.

Among the various proposals for the activation function ψ that have been made, it is worth mentioning:

- Logistic [McCaffrey et al., 1992] $\psi(x) = \frac{e^x}{1 + e^x}$
- Sigmoid [Dechert & Gencay, 1992] $\psi(x) = \frac{1}{1 + e^{-x}}$
- [Nychka et al., 1992] $\psi(x) = \frac{x(1 + |x|/2)}{2 + |x| + x^2/2}$

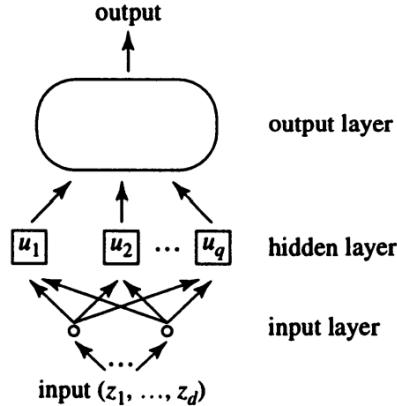


Figure 2.6: Single-hidden-layer feed-forward network (SLFN) architecture. Image credit: [Chan and Tong, 2001]

The overall number of parameters $(q, d, \beta_0, \dots, \beta_q, \omega_{01}, \dots, \omega_{1q}, \dots, \omega_{d1}, \dots, \omega_{dq})$ to be estimated is $q(d+2)+3$ and they respectively are the number of q units $u_j = \sum_{i=1}^n W_{ij} Z_i$ in the hidden layer, the embedding dimension d , the hidden-to-output weights β_j and the input-to-hidden unit strengths ω_{ij} (also called connections). As far as parameter estimation is concerned, an often applied criterium is the generalised cross-validation (GCV) [Chan and Tong, 2001]. Nonetheless, [Nychka et al., 1992] decides to adopt the Bayesian information criterion (BIC), by assuming the gaussianity of the error terms, which may be reasonable for a very large sample. Formally,

$$\text{BIC} = 2[1 + \log(2\pi) + \log(\text{RSS}) + P \log(N)/N]$$

where $\text{RSS} = \frac{1}{N} \sum \hat{\epsilon}^2 = \frac{1}{N} \sum (X_t - \hat{f}_d(X_{t-1}, \dots, X_{t-d}))^2$ is the residual sum of squares, P is the number of parameters, and N is the number of observations. Unfortunately, since the optimisation procedure has to scan through all the parameter combinations, in case of a large parameter grid, the training procedure can be time-consuming.

In conclusion, indirect methods generally outperform direct ones for both clean and noisy data. Moreover, besides providing consistent estimators for the MLCE, indirect methods are considerably less arbitrary as they neither involve a prior time series embedding with heuristic methods nor require the identification of any linear portion, as in the case of Rosenstein's and Kantz's methods. Unfortunately, computational

burdens can arise when being applied to highly-dimensional dynamical systems and also, unless in the identical noise realisation case, there is no natural extension to the stochastic dynamical systems framework, which will be the topic of the following section [Giannerini & Rosa, 2004].

2.1.3 Local Lyapunov exponents

As discussed so far, the global Lyapunov exponent provides an insight into the initial value sensitivity and, hence, the overall system unpredictability. Nonetheless, even if a system is associated with a negative global Lyapunov exponent, there may still be regions characterised by finite-time instability [Bailey et al., 1997]. Conversely, global chaotic time series can include regions of short-term trajectory convergence and higher predictability. It is now crucial to emphasise that chaos is solely a global property of a system, and the concept of local chaos is not meaningful and should be avoided. For this reason, it is useful to introduce the notion of local Lyapunov exponent, which allows the detection of local highly nonlinear dynamics and provides evidence about the system dynamics heterogeneity [Bailey, 1996].

From a historical perspective, the concept of a local criterion for the stability of classical mechanical Hamiltonian systems was first introduced by [Kosloff & Rice, 1981] but only with [Abarbanel et al., 1991; 1992] it was properly formalised and given the name local Lyapunov exponent (LLE). As already anticipated, the crucial idea behind LLEs is to capture the local rate of divergence of initially nearby trajectories and, practically, this is achieved by considering a discrete number of steps ahead, m , in equation 2.7. The local Lyapunov spectrum, also called m -step local Lyapunov vector, is then the set of eigenvalues of the m -step-ahead Oseledec matrix, with the largest corresponding to the LLE. In brief, given a small initial perturbation U_0 , the LLE will be defined as

$$\lambda_m(t) = \frac{1}{m} \ln \|J_{m+t-1} J_{m+t-2} \cdots J_t U_0\| \quad (2.11)$$

where J is the Jacobian of the map F at X_t and with $\lambda_m(t)$ converging almost surely to the global maximum Lyapunov exponent λ as $m \rightarrow \infty$.

Similarly to the global maximum Lyapunov exponent (MLCE), the computation of the local one can be performed with either direct or indirect methods, as done in [Abarbanel et al., 1991], where the underlying dynamics is locally approximated by a nonparametric polynomial regression.

The application of nonparametric indirect methods, together with the formalisation of the LLE within the stochastic dynamical systems framework, has allowed proving the consistency [McCaffrey et al., 1992; Nychka et al., 1992] and then deriving the asymptotic distribution of the indirect estimator of the MLCE [Whang & Linton, 1999]. Indeed, assuming n to represent the number of state vectors and m the number of steps ahead, then, given the true unknown m -fold Jacobian $J_m = DF^m(X_0) = DF(x_{m-1}) \times DF(x_{m-2}) \times \cdots \times DF(x_0)$ and its nonparametric estimate \hat{J}_m , it holds that, under the assumptions:

1. There exists a sequence $\beta_n \rightarrow 0$ s.t $\sup_{1 \leq t \leq n} \|\hat{J}_t - J_t\| = O_p(\beta_n)$ as $n \rightarrow \infty$.
2. The system X_t is ergodic.
3. (a) X_t has finite λ , i.e. $\liminf_{t \rightarrow \infty} \frac{1}{t} \log \|J_t J_{t-1} \cdots J_0\| > -\infty$ a.s.

(b) The n -fold Jacobian is bounded over X_t , i.e. $\sup_{t \geq 0} \|J_t\| < \infty$ a.s.

$$\hat{\lambda}_m - \lambda_m = O_p(\beta_n^{1/m}) \text{ as } m, n \rightarrow \infty \text{ and } m\beta_n \rightarrow 0$$

In plain words, this amounts to saying that the accuracy of $\hat{\lambda}$, the MLCE, is directly related to the convergence rate of the Jacobian matrix estimation. This consistency result is remarkable as it provides the basis for deriving the asymptotic distribution of the MLCE indirect estimator, computed with the nonparametric ANN approach. More specifically, based on the proof of the CLT theorem for the LLE derived by [Bailey, 1996], [Shintani & Linton, 2004] shows that under suitable on the system dynamics X_t and on the single-hidden-layer feed-forward network architecture, it is true that

$$\sqrt{m}(\lambda_m - \lambda) \rightarrow N(0, \Phi) \quad \text{as } m \rightarrow \infty$$

Leveraging the asymptotic normality of the MLCE estimator, it has been possible to derive a statistical chaos test based on the sign of maximum Lyapunov exponent [2004]. This one-sided test takes the form

$$\begin{aligned} H_0 : \lambda &\leq 0 \\ H_1 : \lambda &> 0 \end{aligned}$$

and it has $\hat{t} = \frac{\lambda_m}{\sqrt{\Phi \hat{\lambda}_m}}$ as test statistic, where $\hat{\Phi}$ is a consistent estimator of the variance.

The null hypothesis is then rejected if $\hat{t} \leq -z_\alpha$ where z_α is the critical value $Pr[Z \geq z_\alpha]$ with Z corresponding to a standard normal random variable.

2.2 Stochastic Dynamical Systems

In the previous section, we introduced the study of nonlinear dynamics, focusing on how purely deterministic systems can result in random behaviours. In real-world applications, though, it is unrealistic to assume that observational data are error-free, whether due to measurement or rounding errors. For this reason, it is crucial to introduce the stochastic dynamical systems framework, which includes a dynamic noise component [Longtin, 2010] and treats each state as random. Moreover, as discussed in Section 2.1, chaos can both arise from certain nonlinear maps and flows. Analogously to the deterministic case, there are on the one hand continuous stochastic dynamical systems, generally defined as stochastic differential equations (SDE), and on the other, discrete stochastic dynamical systems, which take the form of stochastic difference equations. Formally, a stochastic difference equation is defined as

$$X_{t+1} = F(X_t, e_{t+1}), \quad t \in T \subseteq \mathbb{Z}^+ \tag{2.12}$$

where that the transition probability $F : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ is a real-value function of class C^2 and $\{e_t\}$ is a sequence of d -dim iid random vectors to account the dynamic noise, not the measurement error. The future state vector X_{t+1} derives its randomness from the dynamic noise and depends on all previously assumed states X_s (for $0 \leq s \leq t$), contained in the σ -algebra \mathcal{J}_s . However, by imposing the simplifying Markov property, which states that

$$F(X_{t+1} \in B \mid \mathcal{J}_t) = F(X_{t+1} \in B \mid X_t) \quad \forall t \in T \subseteq \mathbb{Z}^+, \forall B \in \mathcal{B}(\mathbb{R})$$

the future system state depends entirely on the current one. Then, by assuming an additive noise, we end up with the following first-order Markov chain:

$$X_{t+1} = F(X_t) + e_{t+1} \quad t \in T \subseteq \mathbb{Z}^+ \quad (2.13)$$

which is defined on the continuous state space $S \sim \aleph_1$ and for an at most countable indexing set $T \sim \aleph_0$. The connection between this new general framework and experimental data, in the form of univariate time series, arises from the fact that equation (2.13) is essentially the state-space representation of the following nonlinear autoregressive process of order d :

$$\begin{aligned} y_t &= f_d(X_{t-1}) + \epsilon_t \quad t \in T \subseteq \mathbb{Z}^+, f : \mathbb{R}^d \rightarrow \mathbb{R} \quad \text{s.t.} \\ X_t &= (y_t, y_{t-\tau}, \dots, y_{t-d\tau+1})^T \\ F(X_{t-1}) &= (f(X_{t-1}), y_{t-\tau}, \dots, y_{t-d\tau+1})^T \\ e_t &= (\epsilon_t, 0, \dots, 0)^T \end{aligned} \quad (2.14)$$

where the first component represents the deterministic skeleton of the stochastic difference equation, while the second is the proper source of noise, whose "admissible size" depends on the geometry of both the attractor and the domain of attraction [K. S. Chan & Tong, 1994]. In this regard, it is worth highlighting that, no matter the framework being involved, the nonlinear analysis of the empirical signals always requires the state-phase reconstruction, either with consistent or heuristic methods.

2.2.0.1 Markovian Properties of Stochastic Difference equations

Even though it may often be reasonable to relax the independence assumption by imposing that the dynamic noise satisfies $E[\epsilon_t | \mathcal{F}_{t-1}] = 0$, in the case where $\{\epsilon_t\}$ is iid, equation 2.13 becomes a discrete-time Markov chain defined on a continuous state-space. Consequently, remarkable properties of the discrete random dynamical systems can be derived by extending the canonical framework of countable state space Markov chains; see [Chan & Tong, 2001] for a detailed discussion.

In particular, given the initial condition $X_0 = x$, the subsequent state of the chain is represented by the random variable $X_1 = F(x, e_1)$, distributed according to the first-step transition probability $P(x, A) = P(x_1 \in A | X_0 = x)$ for whatever initial condition x and any borelian $A \in \mathcal{B}$. Analogously, the n -th state will be $X_n \sim P^n(x, A) = P(X_n \in A | X_0 = x)$. As a consequence, the state space or absorbing set — known as the minimal region over which the Markov chain has the same asymptotic behaviour independently from the initial conditions — is defined as the subset $S \in \mathbb{R}^d$ such that $P(x, S) = 1 \forall x \in S$. Nevertheless, the asymptotic behaviour of the chain is actually independent of the initial condition — so the initial state — if the Markov chain is irreducible. More specifically, this condition is satisfied in a continuous state space if it exists some $n > 0$ such that $P^n(x, A) > 0 \quad \forall x \in S$ and for any ϕ -non-null $A \subseteq S$, where ϕ is a positive measure known as irreducible measure, which can be for simplicity the Lebesgue's one.

A fundamental characteristic of chaotic deterministic systems concerns the aperiodicity of the orbits and for such a reason it is worth extending this property to the noisy chaos framework. Indeed, given a ϕ -irreducible Markov chain whose state space S can be partitioned into a finite collection of ϕ -non-null sets, E_1, E_2, \dots, E_d , such that:

1. For $i = 1, 2, \dots, d - 1$, $P(x, E_{i+1}) = 1$ for all $x \in E_i$, and $P(x, E_1) = 1$ for all $x \in E_d$, meaning that the Markov chain cycles through E_1, \dots, E_d ;
2. The state space S is almost entirely covered by the sets E_i , so that $\phi(S \setminus \bigcup_{i=1}^d E_i) = 0$;
3. d is the largest integer for which conditions (1) and (2) are satisfied;

then the Markov chain is called aperiodic if $d = 1$ and periodic with period d otherwise. Finally, as in the deterministic framework, a crucial property is ergodicity, which allows describing the long-term behaviour of a random dynamical system — without dominating noise — regardless of initial conditions. Formally, a continuous state space Markov chain is said to be ergodic if it is irreducible, aperiodic, and admits a stationary distribution — which, due to irreducibility, will be unique^(viii). From the operational standpoint, the ergodicity of the stochastic difference equation is closely related to the stability of its deterministic skeleton, and it allows deriving statistical quantities, including the rate of divergence of stochastic orbits, which would otherwise be unknown.

2.2.1 Initial value sensitivity

2.2.1.1 Identical noise realisation hypothesis

A convenient situation for studying the initial value sensitivity of a stochastic difference equation occurs when it is assumed that the trajectories, whose exponential divergence we aim to measure, start from two nearby initial conditions — say x_0 and $x_0 + \delta$ — and are subject to the same dynamic noise e_t . In fact, in this case — as also highlighted in [Nychka et al., 1992] — the noise effects will then cancel out when considering the distance between trajectories and the overall problem reduces to the deterministic case analysed in Section 2.1. Thus, by considering for simplicity $d=1$, we can see that

$$\begin{aligned} |X_{t+1}(x + \delta) - X_{t+1}(x)| &= |(F_t(x + \delta) + e_t) - (F_t(x) + e_t)| \\ &= |F_t(x + \delta) - F_t(x)| \\ &\approx \exp\{m\lambda\}|\delta| \end{aligned}$$

where $\lambda = \lim_{m \rightarrow \infty} \frac{1}{m} \log |\prod_{i=0}^{n-1} \dot{F}(X_i(x))|$ is the Lyapunov exponent.

Unfortunately, the inclusion of an additive dynamic noise prevents the Lyapunov exponent from remaining invariant with respect to one-to-one transformations, such as coordinate transformations, and it is thereby necessary a modification of the Lyapunov exponent definition 2.7. Moreover, the identical noise realisation is, in general, a quite strong assumption that in many applications cannot be justified. For this reason, the following sub-section will introduce a flexible scheme for extending the notion of Lyapunov exponents to stochastic dynamical systems framework.

^(viii)In this context, Birkhoff's Ergodic theorem [Strogatz, 2018] plays a crucial role, by guaranteeing that asymptotically the temporal mean coincides with the spatial mean, computed with respect to the natural measure, which describes the fraction of time that an orbit spends in a given region. In the case of irreducible chains such distribution is stationary unique, meaning that asymptotically the temporal and spatial mean are actually equal to the natural measure [Grimmett & Stirzaker, 2001]

2.2.1.2 Conditional Distribution Approach

A substantially different approach for measuring the exponential divergence of initially nearby stochastic orbits has been introduced by [Yao & Tong, 1994b; 1994a] and it aims to study the sensitivity of the conditional distribution $g(X_m | X_0 = x)$ with respect to different initial conditions, according to some metric or dissimilarity measure. Alternatively, there may also be considered the sensitivity to some summarising characteristics, such as the conditional mean, which provides an interesting connection with the noise-free case.

Formally, let $F_m(x) = \mathbb{E}(X_m | X_0 = x)$ be the expected value assumed by the m -th state of the chain, conditionally to a generic vector of initial condition $x \in \mathbb{R}^d$. Then, given a perturbation $\delta \in \mathbb{R}^d$, the divergence between two stochastic orbits after m time-steps is given by:

$$F_m(x + \delta) - F_m(x) = \Lambda_m(x)\delta + o(\|\delta\|)$$

where $\Lambda_m(x) = \frac{dF_m(x)}{dx} = \mathbb{E} \left\{ \prod_{k=1}^m \dot{F}(X_{k-1}) \mid X_0 = x \right\}$ is a $d \times d$ matrix, and the squared largest eigenvalue ν_m of the matrix $\Lambda_m(x)^T \Lambda_m(x)$ coincides with the m -step maximum Lyapunov-like index. In the specific sub-case where $d = 1$, the Lyapunov-like index is defined as

$$\kappa(x) = \lim_{m \rightarrow \infty} \frac{1}{m} \log \left| \frac{d}{dx} F_m(x) \right| \quad (2.15)$$

$$= \lim_{m \rightarrow \infty} \frac{1}{m} \log \left| \mathbb{E} \left\{ \prod_{i=0}^{m-1} \frac{d}{dx} F(X_i) \mid X_0 = x \right\} \right| \quad (2.16)$$

$$= \lim_{m \rightarrow \infty} \frac{1}{m} \log |\lambda_m(x)| \quad (2.17)$$

and the estimate will converge to the Lyapunov exponent of a purely deterministic system as the dynamic noise goes to zero, so if the system has a long memory.

From an operational perspective, a more informative sensitivity measure, derived from the conditional distribution approach, is the Kullback-Leibler information which quantifies the sensitivity to the initial conditions by measuring the amount of information of an initial value x_0 retained in a generic future state y ; see [1995] for a comparison with the conditional distribution approach. In formal terms, for two nearby initial conditions $x, x + \delta \in \mathbb{R}^d$ and a sufficiently smooth conditional distribution $g(\cdot|x)$, the n -step Kullback-Leibler information is defined as

$$K_m(x; \delta) = \int \{g_m(y|x + \delta) - g_m(y|x)\} \log \left\{ \frac{g_m(y|x + \delta)}{g_m(y|x)} \right\} dy$$

which for a small δ can be approximated to

$$K_m(x; \delta) = \delta^T I_m(x)\delta + o(\|\delta\|^2)$$

where

$$I_{1,m}(x) = \int \frac{\dot{g}_m(y|x)\dot{g}_m^T(y|x)}{g_m(y|x)} dy \quad (2.18)$$

is the Fisher Information matrix, which explains how the conditional distribution of a generic future state, $g(y|x)$, is sensitive to the initial condition. Intuitively, the more

information a future state, y , retains about the initial condition x_0 , the more sensitive the system will be. Therefore, high values of $I_{1,m}(x)$ indicate that small changes in the initial conditions deeply affect the future states X_m of the system. Nonetheless, since the sensitivity measure $I_{1,m}$ depends on the time step m , it then converges to $\lim_{m \rightarrow \infty} I_{1,m} = 0$, reflecting the long-term memorylessness of the system due to the dynamic noise accumulation.

From a computation standpoint, the estimation of the initial value sensitivity $I_{1,m}$ requires the prior calculation of $g(\cdot|x)$ which can be derived through a nonparametric locally quadratic regression. Indeed, starting from the nonparametric density function $K(\cdot)$, whose smoothness depends on the bandwidth parameter h , we have that

$$E(K_{h_2}(Y_m - y) | X_0 = x) \approx g_m(y | x) \quad \text{as } h_2 \rightarrow 0$$

where the left-hand side of can be seen as a regression on the process values X_m . Then, by applying the second-order Taylor's expansion we get that

$$\begin{aligned} E(K_{h_2}(Y_m - y) | X_0 = z) &\approx g_m(y|z) \\ &\approx g_m(y|x) + \dot{g}_m(y|x)^T(z - x) + \frac{1}{2}(z - x)^T \ddot{g}_m(y|x)(z - x) \\ &\equiv \beta_0 + \beta_1^T(z - x) + \beta_2^T \text{vec}\{(z - x)(z - x)^T\} \end{aligned}$$

where $\beta_0 = g_m(y|x)$, $\beta_1 = \dot{g}_m(y|x) = \frac{\partial g_m(y|x)}{\partial x^T}$, and $\beta_2 = \ddot{g}_m(y|x)$ is the Hessian matrix of $\beta_0 = g_m(y|x)$. Eventually, the estimation of $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ is possible by the following least square minimisation:

$$\arg \min_{\beta_0, \beta_1, \beta_2} \sum_{t=1}^{n-m} (K_{h_2}(Y_{t+m} - y) - \beta_0 - \beta_1^T(X_t - x) - \beta_2^T \text{vec}\{(X_t - x)(X_t - x)^T\})^2 W_{h_1}(X_t - x)$$

in which $W = \text{diag}(W_{h_1}(X_1 - x), \dots, W_{h_1}(X_{n-m} - x))$ is a non-negative kernel, whereas h_1 and h_2 are respectively the bandwidth for W and K , chosen through a modified cross-validation technique that allows tackling data dependencies [Chan & Tong, 2001; Tong et al., 1995; Yao & Tong, 1995]. In addition, the resulting estimators of β_0 and β_1 will be asymptotically normal. In the univariate case, the sensitivity measure $I_{1,m}$ can be either derived from 2.18 or alternatively from

$$I_{1,m}(x) = 4 \int \left\{ \frac{d\sqrt{g_m(y|x)}}{dx} \right\}^2 dy = 4 \int \left\{ \frac{d\sqrt{\hat{\beta}_1}}{dx} \right\}^2 dy$$

which produces a much smoother and easier to simplify solution. Nevertheless, in both cases, the least squares estimation reduces to

$$\hat{\beta}_j(x, y) = h_1^{-1} \sum_{t=1}^{n-m} W_j^n(t) K_{h_2}(Y_{t+m} - y), \quad j = 0, 1, 2$$

$$\text{with } W_j^n(t) = e_j^T S_n^{-1} (1, h_1 t, h_2^2 t^2)^T \times W(t), \quad (e_j)_i = \delta_{i,j+1}^{(ix)},$$

$$\text{the scale matrix } S_n = \begin{pmatrix} s_{n,0} & s_{n,1} & s_{n,2} \\ s_{n,1} & s_{n,2} & s_{n,3} \\ s_{n,2} & s_{n,3} & s_{n,4} \end{pmatrix} \quad \text{and } s_{n,j} = \sum_{t=1}^{n-m} (Y_t - x)^j W_{h_1}(Y_t - x).$$

^(ix)Here, $\delta_{i,j}$ is the Kronecker delta function, which equals 1 when $i = j$ and 0 otherwise.

An alternative sensitivity measure proposed by [Fan & Gijbels, 1995] is the L_2 norm, which shares with the Kullback-Leibler divergence the computational framework, the asymptotic convergence to zero — $\lim_{m \rightarrow \infty} I_{2,m} = 0$ — and the capability of capturing the conditional heteroscedasticity. Nonetheless, regardless of the sensitivity measure being applied, as in the deterministic system, the m -step largest Lyapunov-like index will be invariant under one-to-one transformations.

2.2.2 Noise amplification and non-monotonicity of prediction accuracy

Besides the exponential divergence of nearby trajectories, there are other peculiar characteristics associated with highly nonlinear dynamics, such as the noise amplification and the non-monotonicity of prediction accuracy — both of which are actually strictly related to the initial value sensitivity. In particular, given model 2.13 and assuming

$$\mathbb{E}[e_t | X_{t-1}, X_{t-2}, \dots] = 0, \quad \text{Var}[e_t | X_{t-1}, X_{t-2}, \dots] = \Sigma \quad \text{and} \quad \|e_t\| \leq \xi \quad \text{a.s.}$$

it is possible to get, through the Taylor's expansion, that for a generic number of steps-ahead $m \geq 1$

$$X_m = F^{(m)}(x) + e_m + \Lambda[F^{(m-1)}(x)]e_{m-1} + \dots + \left\{ \prod_{k=1}^{m-1} \Lambda[F^k(x)] \right\} e_1 + O_p(\xi^2) \quad (2.19)$$

whose conditional variance $\Sigma_m(x) = \text{Var}(X_m | X_0 = x)$ is

$$\Sigma_m(x) = \Sigma + \sum_{j=1}^{m-1} \left(\prod_{k=j}^{m-1} \Lambda\{F^{(k)}\}(x) \right) \left(\sum_{j=1}^{m-1} \prod_{k=j}^{m-1} \Lambda\{F^{(k)}\}(x) \right)^T + O(\xi^3) \quad (2.20)$$

$$= \Sigma_0 \mu_m(x) + O(\xi^3), \quad \mu_m(x) = 1 + \sum_{j=1}^{m-1} \left\{ \prod_{k=j}^{m-1} \Lambda\{F^{(k)}(x)\} \right\}^2 \quad (2.21)$$

In the sub-case where $d = 1$, equation 2.20 reduces to

$$\sigma_m^2(x) = \sigma^2 \left\{ 1 + \sum_{j=1}^{m-1} \left(\prod_{k=j}^{m-1} \lambda f^k(x) \right)^2 \right\} + O(\xi^3) \quad (2.22)$$

which shows that, depending on the initial conditions, large values of the Lyapunov exponent can significantly influence the conditional variances, bringing to the apparent strange result of the non-monotonicity of the conditional variance, namely, $\sigma_{t+1}^2 \leq \sigma_t^2$ when $\{\lambda f^{(m)}(x)\}^2 \leq 1 - 1/\mu_m(x)$.

The understanding of the conditional variance behaviour is particularly helpful when studying the system m -step point prediction, $f_m(x) = \mathbb{E}(Y_m | X_0 = x)$. Indeed, the decomposition theorem derived by [Yao & Tong, 1994a] states that given the least squares predictor $\hat{f}_m(x)$, if $\mathbb{E}\{[f_m(x) - \hat{f}_m(x)]^2 | X_n\} \rightarrow 0$ a.s., then

$$\lim_{n \rightarrow \infty} \mathbb{E}\{[(Y_{n+m} - f_m(x))]^2 | X_n = x + \delta\} = \sigma_m^2(x + \delta) + (\delta^T \lambda_m(x))^2 + o(\|\delta\|^2) \quad \text{a.s.}$$

In plain words, the mean-square prediction error for a stochastic dynamical system subject to a tiny initial shift, δ , depends on the conditional variance but also on the

exponential divergence generated by δ and driven by the Lyapunov exponent. In conclusion, as a result of equation 2.22, one may see that the point prediction accuracy depends on the initial condition, as well as on the initial value sensitivity — through both decomposition theorem terms — implying that the $(m+1)$ -step ahead prediction can more accurate than the m -step ahead one.

Chapter 3

Nonlinear Brain Dynamics: A literature review

The human brain is considered one the most complex systems in the universe, and its dynamics are still largely unknown, representing the subject of extensive multidisciplinary research [Hadeaghi, 2023]. Among the first non-invasive techniques that have contributed to a better understanding of human brain activity are electroencephalograms (EEGs), developed by the German psychiatrist Hans Berger at the beginning of the 1930s. Even in those early recordings, it became evident that substantial changes in the brain signals waveform occurred when the eyes were closed or depending on the brain's functional status (e.g. relax or focus states). Nonetheless, despite significant bio-medical advancements, EEGs — together with magnetoencephalography (MAG) — have remained a crucial tool for tracking fast temporal changes in brain dynamics, allowing the diagnosis of seizures, strokes and also brain damage or tumours.

In essence, EEGs record brain electrical activity by measuring and amplifying the electrical potential between two adjacent bipolar electrodes, producing signals that can either assume positive or negative values, generally expressed in microVolts (μV , or thousands of a Volt) [Cohen, 2014]. Specifically, EEG signals can either be registered in a completely non-invasive way by placing electrodes on the scalp (generally according to the International 10-20 system) or with intracranial electrodes, which are more sensitive and can capture activities for much smaller cortical areas, without a severe averaging effect between brain current sources.

Considerable challenges in the study of EEGs are posed by the nonstationary, nonlinear and noisy nature — often regarded as the 3N — of bio-signals, which are severely affected by distortions or unwanted signals — known as artifacts — due to technical and physiological issues, such as eyes blinking or muscle activity. These difficulties have significantly impacted the analysis carried out with traditional linear methods and, as we will see in the present chapter, they have promoted the study of brain dynamics through the paradigm and tools of chaos theory.

3.1 Nonlinear EEG analysis

Human brain activity depends on the active interplay of tens of billions of neurons, which results in virtually infinite-dimensional dynamics [Frank et al., 1990]. In addition to that, nonlinearity is introduced even on a cellular level, since the behaviour of individual neurons is governed by threshold and saturation mechanisms. Besides the

Table 3.1: Brain frequency bands and associated activities.

Brain Frequency Band	Associated Activity
Delta (0.5-4 Hz)	Associated with drowsiness and deep relaxed states
Theta (4-8 Hz)	Common during unfocused and distracted phases, but also during creative activities
Alpha (8-13 Hz)	Present during relaxed and conscious wakefulness, especially with closed eyes
Beta (13-30 Hz)	Linked to active thinking, alertness, and agitation
Gamma (30-40 Hz)	Associated with high-level cognitive functions and information processing

caveat that a collection of nonlinear elements does not necessarily behave so [Lehnertz, 2008], it is still particularly convincing that complex, and sometimes chaotic, dynamics may arise. Unfortunately, traditional linear signal processing tools — both parametric or not — provide poor results in the presence of intricate nonlinear dynamics. Indeed, the system’s intrinsic nonlinearity is not detected and information contained in the signal is missed — as noted by [Casdagli et al., 1997], for instance, given any power spectrum, there exists a linear process with such spectral properties.

Contemporaneously with the theoretical advancement described in the previous chapter, since the 1980s a substantial scientific effort has been directed towards detecting chaos in cerebral electrical activity signals (EEG); see [C. J. Stam, 2005] and [Pritchard & Duke, 1995] for a review. Interestingly, as noted by [Rapp, 1992] chaos theory tools have eventually provided a deeper comprehension of brain activity than the concept of chaos itself, for which less conclusive results have been obtained.

Anticipated by [Rapp et al., 1985], which studies the neural activity of a single neuron in the motor cortex of a monkey, the first pioneering studies in the field of human brain nonlinear analysis are those of [Babloyantz et al., 1985] and [Babloyantz & Destexhe, 1986], which focus on peculiar neuronal conditions, such as the sleep cycle and epilepsy, to constrain the problem into a lower dimensional state space compared to the one that would have been necessary for normal brain activity. More specifically, estimates of the correlation dimension⁽ⁱ⁾ have suggested a lower brain activity complexity in sleep stages II — associated with the lightest sleep — and IV — deepest sleep — compared to REM and awake phases for which no estimates could have been provided, but because of the applied saturation point heuristic, this inevitably suggests a higher dimensionality of the reconstructed attractors. For as much as it concerns this first breakthrough publication, [Babloyantz et al., 1985] also provides estimates of the global Lyapunov exponent based on Wolf’s method, developed just a few years earlier and implemented in Fortran. More specifically, global chaotic dynamics are detected in both sleep stages II and IV, with estimates of the Largest Lyapunov exponent particularly high, respectively between 0.4-0.8 and 0.3-0.6.

Of particular interest for the scope of this thesis, [Babloyantz & Destexhe, 1986] discovers a lower complexity dynamics in human EEGs during absence seizures (also called *petit mal*, a specific type of generalised seizure), corroborating the medical intuition

⁽ⁱ⁾The *correlation dimension* [Grassberger & Procaccia, 1983] provides an upper-bound estimate of the fractal dimension

of epilepsy as a time frame of brain regions hyper-synchronisation [Lehnertz & Elger, 1995]. In fact, this hyper-synchronization is associated with a reduction in the system's degrees of freedom⁽ⁱⁱ⁾, and it results in more regular oscillating signals — known in the medical field as *sharp waves*. Moreover, indications of global chaotic dynamics have been inferred from a particularly large — and to some extent also too large — Lyapunov exponent, $\lambda = 2.9 \pm 0.6$.

The enthusiasm for these first publications has led to the flourishing of a vast literature interested in describing a multiplicity of mental states through the lenses of deterministic chaos. [Albano et al., 1986] and [Layne et al., 1986], to name a few, focus for instance on the brain complexity reduction due to eyes closing, corroborating the medical indication of an increase of the alpha waves in the posterior cortex, which are associated with a relaxed state. [Gregson et al., 1990] has instead tried to explain the rise in brain complexity as a result of performing various mental tasks.

Unfortunately, the conflicting requirement of long and stationary series, together with the awareness that methods for the correlation dimension and the largest Lyapunov exponent may lead to a spurious indication of chaos, by detecting linear scaling regions in the presence of noise contamination [Osborne & Provenzale, 1989], led to a general rethinking of the early claims for chaos in the human brain.

Criticisms on the presence of low-dimensional chaotic behaviours have become more severe with the advent of surrogate data testing, which was designed for checking the reliability of the found results but ultimately led to a rejection of the chaotic hypothesis [Prichard & Theiler, 1994], [Palus, 1996]. Introduced by [Theiler et al., 1992], the surrogate data technique is a statistical test that compares through a nonlinear discriminant measure the original time series data and the surrogate one, generated — often with Monte Carlo methods — under a given null hypothesis, which aims to disrupt the deterministic nature of the original signal. Random shuffle, block shuffle — including the spike-and-waves technique as a special case [Theiler, 1995] — and Fourier phase shuffle — together with the Amplitude Adjusted Fourier Transform that prevent the replicated data from appearing Gaussian — are some of the most applied replication schemes. Unfortunately, they all suffer from a reduced sensitivity due to a low variance in the test statistic associated with the replicated data. The narrow distribution of the discriminant statistics may lead to incorrect conclusions; [Giannerini et al., 2015; C. J. Stam, 2005] and reference therein.

While the question of whether the brain is intrinsically deterministic or not has been set aside, a new wave of brain studies has arisen, which are more focused on interpreting the nonlinear measures as relative indices to describe changes in neuronal dynamics — in contrast to the strict initial bit/sec interpretation — and also, more recently, on developing statistically robust methods to be applied with noisy data. In the sub-field of brain dynamics during sleep, various studies — [Achermann et al., 1994] and [Fell et al., 1993] among the others — have challenged the previous assumptions about two positive relations: one between deeper sleep stages and an increase in brain complexity, while the another suggesting a reduction of the Lyapunov exponent. Instead, these recent studies highlight that, although highly nonlinear dynamics are present, there is no evidence of deterministic chaos.

Besides sleep, which perhaps provides a prototypical example of the scientific attitude change, a large variety of psychophysical conditions has been considered. Several

⁽ⁱⁱ⁾Conceptually, the hyper-synchronisation leads to a redundancy in the equations underlying a process, leading to a reduction of their number.

studies have, for instance, shown how brain complexity increases at an early age [Meyer-Lindenberg, 1996] but that tends to decrease in elderly people [King-Hang et al., 2021], as a consequence of reduced interactions between brain regions. Furthermore, the effects of neurodegenerative diseases, such as schizophrenia [Lee et al., 2001], Parkinson's [Müller et al., 2001] or Creutzfeldt-Jakob diseases [C. Stam et al., 1997], may also be associated with a complexity loss and a reduction of nonlinearity.

In the remaining space, we would now like to focus on a particular brain chronic disorder, epilepsy, which has been extensively studied in the field of NLTS and will be at the centre of the analysis carried out in the next chapter.

Epilepsy is the second most common neurological disorder, affecting about 1% of the global population [World Health Organization, 2024]. It is characterised by the occurrence of epileptic seizures, which, according to [Fisher et al., 2005], are "transient occurrences of signs and/or symptoms due to the abnormal excessive or synchronous neural activity in the brain". Seizures can either arise in both hemispheres — *generalised seizures*, which are then subdivided according to the effects on the body, such as the previously mentioned absence seizure — or in just one of them — *focal seizures*⁽ⁱⁱⁱ⁾, 60% of which are temporal lobe epilepsy (TLE). The epilepsy evolution is characterised by 3 main phases: *pre-ictal*, during the seizure onset, which is associated with nausea, déjà vu, headache and sensory confusion — odd smells, sounds, or tastes; *ictal*, the proper epileptic phase, whose effects depend on the specific kind of seizures but that generally involves a loss of awareness, racing heart but also body stiffening or convulsion; *post-ictal*, the time period following a seizure but preceding the reset of the brain's normal activity, during which confusion, lack of consciousness and weakness in the body part are experienced [Epilepsy Foundation Australia, 2024]. The sudden and mostly unpredictable nature of seizures may lead to dangerous accidents and injuries, which contribute to a threefold increase in the risk of premature death for the epilepsy population. Furthermore, severe pathological conditions, such as anxiety and depression, are often experienced by epilepsy subjects. Unfortunately, besides the great scientific effort, much remains unknown about when and why seizures occur and currently nearly one out of four epilepsy subjects cannot be treated with drug or invasive surgery procedures, causing significant quality of life limitations. The belief in complex and highly nonlinear dynamics underlying epileptic seizures, together with the poor results obtained by traditional linear methods [Rogowski et al., 1981], has motivated the application of chaos theory tools to the study of epileptic EEGs. Unfortunately, though, Babloyantz and Dexstele's intuitions were not immediately accepted, sparking a large, and still open, scientific debate. Indeed, although the *complexity loss hypothesis* has been widely corroborated by the vast majority of scientific publications, also considering a plurality of different kinds of generalised seizures [Frank et al., 1990], the issue of chaoticity has remained much more controversial. [Iasemidis et al., 1990] discovered that the Lyapunov exponents computed on the various brain sites reach their maximum during the post-ictal phase, a lower value in the pre-ictal phase and their minimum in the proper ictal phase. Based on this result, a chaos-to-order-to-chaos transition was proposed, according to which, epilepsy, compared to normal brain activity, is still characterised by a chaotic but less irregular dynamic. Moreover, various surrogate data analyses — [Elger & Lehnertz, 1998; Lehnertz & Elger, 1995] among

⁽ⁱⁱⁱ⁾Focal seizures — previously called partial — may remain localised, causing mild psycho-physic symptoms, or they can spread to the other hemisphere, giving rise to severe secondarily generalised seizures [Iasemidis, 2011]

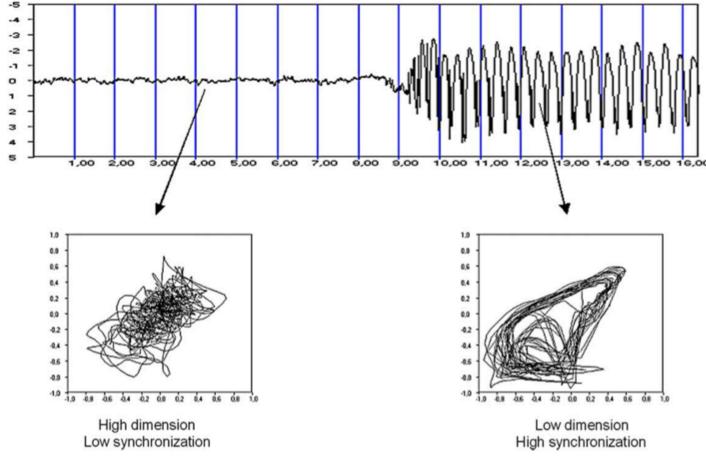


Figure 3.1: The interictal and pre-ictal stages are marked by high dimensionality and low degree of synchronisation, while the ictal stage is characterised by low dimensionality and high synchronisation. This regime change leads to a significant increase in the magnitude of EEG signal values and more regular oscillating patterns. Image credit: [2005].

the others — have, on the one hand, suggested the presence of highly nonlinear but low dimensional dynamics during the ictal phase, while, on the other hand, reveal highly dimensional and less nonlinear dynamics in the interictal phase — 3.1. However, as noted in the multivariate EEG analysis [Rombouts et al., 1995], the evidence of nonlinearity is not accompanied by a proof of deterministic chaos. Additionally, [Feucht et al., 1998; Friedrich & Uhl, 1996] have hypothesised that ictal phase EEGs behave as noisy limit cycles, and in particular, as reported by [C. J. Stam, 2005], [Theiler, 1995] reaches as first this conclusion with a surrogate-based study performed on the same dataset under analysis by [Frank et al., 1990].

Part of the difficulty in effectively detecting chaos is addressed by [Casdagli et al., 1997] to the overly high-dimensional nature of brain dynamics and to the previously described limitations of the surrogate data method, which is not sensitive enough to small variations of the discriminant statistics. As a result, the research has shifted away from proving the presence of chaos, limiting the usage of the Lyapunov exponent as a tentative measure of brain dynamics changes.

Besides the scientific debate, it is worth considering that nonlinear time series tools have provided great contributions to the study of epileptic dynamics, in particular for what concerns the localisation of the primary epileptogenic area. Indeed, a correct identification of the epilepsy focus is necessary not only to understand the type of epilepsy but also to detect the resection area — when the removal of the region is required, approximately in 8% of epilepsy cases — and to develop prevention techniques that can stop seizures by the excretion of fast-acting anticonvulsive substances or by resetting the brain dynamics through electrical stimulus [Morrell, 2006]. Unfortunately, thus far no extracranial recording-based and three-dimensional localisation technique of the epileptogenic area has been developed. However, it is possible to determine which intracranial electrode is associated with the pathological neuronal activity, circumscribing the epileptogenic area as early as the interictal phase, with a much greater precision than with spectral tools [Osterhage & Lehnertz, 2007]. More

specifically, in the case of temporal lobe epilepsy (TLE), it is possible to observe that in both hemispheres, and especially in the focal one, during the pre-ictal phase, and more vividly during the ictal one, there is a reduction of the correlation dimension near the focal area, while minor modifications occur at remote locations. In this respect, an often applied measure to capture the temporal variations in the dimensionality of the brain dynamics is the *complexity loss* L^* , which is essentially the integral over a given time frame of the correlation dimension profile (see Figure 3.3) and that allows the identification of spatiotemporal modifications [Lehnertz & Elger, 1995]. The same measure has also been widely used for studying the effect of antiepileptic drugs. For instance, in [1997] where it is highlighted the negative relation between L^* and Carbamazepine, an anticonvulsant medication which inhibits the neurons from firing at abnormal frequencies. Furthermore, the *complexity loss* measure has been employed for developing prediction methods that can warn the patient in advance, avoiding dangerous situations, which are actually the leading cause of epilepsy-related death. The earliest prediction method, CPLXMON, was introduced by the Department of Epileptology at University Hospital Bonn in the paper [Widman et al., 1995] and, as later explained by [Lehnertz, 2008], it exploits the just discussed complexity loss measure L^* , anticipating seizures even 25 minutes earlier. A similar prediction technique was also developed in [Martinerie et al., 1998] but here only a 2-6 minute interval of advance was attested. Additionally, also the Lyapunov exponent has been exploited to construct prediction methods, as in [Iasemidis, 2011] where the convergence of the various brain site Lyapunov exponents during the ictal phase allows predicting the seizure of about 10 minutes. However, despite the potential of these methods, severe criticisms have been made — [McSharry et al., 2003] — and the presence of such wide temporal gaps inevitably undermines the reliability of these methods, which have gradually been set aside.

3.2 UKB dataset and related literature

As previously discussed, the lack of firm conclusions in the field of nonlinear EEG analysis, and more specifically of epileptology, does not seem to be merely related to the variety of measures and estimation procedures being involved. Part of the issue should indeed be addressed to the employment of different datasets, especially in a context where tiny differences in terms of stationarity or signal length can produce large impacts on the estimated results. For such a reason, from the early 2000s, numerous nonlinear time series-based studies in epileptology studies have applied the UKB dataset, made available online by Dr. Ralph Andrzek of the Epilepsy Center at the University of Bonn (Universitätsklinikum Bonn), where renowned scientists such as Elger and Lehnertz were actively working^(iv).

This dataset, which has become a benchmark over time and that will be employed for our analysis, is composed of 5 groups (A-E) of 100 single-channel EEG segments of 23.6-sec duration, extracted from continuous multichannel EEG recordings, in a manner that guarantees weak stationarity and avoids artifacts — distortions or unwanted signals. On one hand, subsets A and B contain extracranial EEGs from five healthy awake volunteers, recorded according to the international 10–20 electrode placement

^(iv)The UKB can be freely accessed at the following website: <https://www.ukbonn.de/epileptologie/arbeitsgruppen/ag-lehnertz-neurophysik/downloads/>.

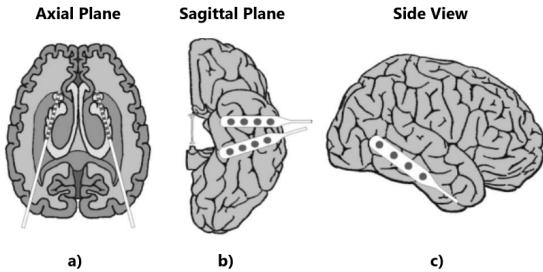


Figure 3.2: UKB dataset intracranial recording scheme. Interictal EEGs (C and D) are recorded with two deep electrodes symmetrically implanted in the hippocampi, targeting the epileptic and contralateral regions (a). Ictal EEGs involve both these hippocampal electrodes (a) and strip electrodes placed in the lateral (b) and basal neocortex (c)

scheme, with eyes open and closed, respectively. On the other hand, subsets C, D, and E are derived from the intracranial presurgical analysis recordings of five epilepsy subjects suffering from temporal lobe epilepsy (TLE), with the epileptogenic focus located in the hippocampal formation, that was later resected. More specifically, EEGs from subsets C and D were recorded using depth electrodes placed in the hippocampal formations, with subset D capturing activity from the epileptogenic zone, while subset C was recorded from the contralateral hippocampal formation, which was not involved in seizure generation (Figure 3.2 a). Moreover, EEG signals belonging to these two classes only refer to neural activity recordings during seizure-free intervals — interictal phases. In contrast, EEG signals belonging to subset E exclusively concern the ictal phase and they were recorded using both the depth electrodes, placed in the hippocampal formations, and the strip electrodes, on the lateral and basal neocortex (Figure 3.2 a, b and c).

As a whole, all EEG signals were recorded in analog form using a 128-channel amplifier system and later digitised with a 12-bit Analog-to-Digital (A/D) conversion, enabling to capture even small signal intensity variations with high precision. The chosen sampling rate is 173.61 Hz, which is consistent with other epileptogenic EEG recordings. For instance, in [Lehnertz & Elger, 1998], a sampling rate of 173 Hz was chosen with the same 12-bit signal quantization as used for the UKB dataset, while [Casdagli et al., 1997] employed a sampling rate of 200 Hz with a 10-bit analog-to-digital converter (ADC).

Unfortunately, some limitations are associated with this dataset, inevitably precluding certain analyses and also, as we will see in the next chapter, diminishing the interpretative power of the results. In particular, EEG signals belonging to each subset are randomised, resulting in the loss of temporal indications that make it impossible to determine the time interval from either the epileptic phase or, rather, in the case of subset E, from the seizure onset. Furthermore, no patient identification information is available, and all patients suffer from the same type of focal epilepsy, limiting broader epilepsy studies as well as the development of classification techniques for recognising different types of epilepsy. Additionally, no extracranial EEGs from either the ictal or interictal phases were recorded.

Despite the described limitations, since its release, the UKB dataset has been extensively used for a considerable number of research projects on brain nonlinear time

Table 3.2: Summary of the structure of UKB electroencephalogram dataset, Epilepsy Center at the University of Bonn.

Group	A	B	C	D	E
Subjects	Five healthy volunteers	Five healthy volunteers	Five epilepsy patients	Five epilepsy patients	Five epilepsy patients
Subject status	Awake, eyes open	Awake, eyes closed	Seizure-free interval (Interictal phase)	Seizure-free interval (Interictal phase)	Ictal phase
EEG type	Extracranial EEG (10–20 system)	Extracranial EEG (10–20 system)	Intracranial EEG, Contralateral hippocampus	Intracranial EEG, Epileptogenic hippocampus	Intracranial EEG, Epileptogenic hippocampus

series, and more recently also in the field of machine learning where classification based on chaos theory measures [Brari & Belghith, 2022] or simply spectral tools — as the spectral density entropy in [Mirzaei et al., 2010], [2011] and [Tsipouras, 2019] — have reached up to 90% accuracy in distinguishing epileptic from healthy subjects. In the hope of providing a more complete perspective on the data analysed in the following chapter, in the remaining space, we would like to propose a brief overview of the main publications analysing the UKB dataset.

The UKB dataset is presented for the first time in [Andrzejak et al., 2001], where the surrogate data technique is employed to investigate the presence of deterministic dynamics and/or a low-dimensional chaotic structure in human EEGs. Statistically speaking, this is carried out through a hypothesis test which assumes as null hypothesis that the signals are the results of stationary Gaussian linear stochastic processes, passed through a static nonlinear measurement function. Based on this hypothesis, 39 surrogate time series are generated for each of the original segments. Two discriminant statistics were applied: the Grassberger-Procaccia correlation dimension $D_{2,eff}$, and the nonlinear prediction error P , which instead offers indications on the degree of randomness and the possible deterministic structure of the phenomenon under analysis. A nonparametric Wilcoxon signed-rank test then revealed that subsets D and E have the highest rejection rates for both statistics, indicating a significant deviation from the null hypothesis. Subsets B and C showed instead high rejection rates only when considering the prediction error, while subset A was consistent with the null hypothesis for both statistics. The intuition behind these results is essentially that ictal phase EEGs (E), along with interictal phase EEGs originating in the epileptogenic hemisphere (D), are compatible with a nonlinear deterministic system. Closed eyes EEGs (B) and interictal phase EEGs originating in the hemisphere opposite the epileptogenic one (D) are a sort of middle ground, as their signals may both appear as nonlinear deterministic or stochastic, depending on the criterion being employed. Finally, open-eye recordings (A) are consistent with a stochastic behaviour. Besides similar surrogate data analyses have reached the same conclusions [Kunhimangalam et al., 2007], these studies are not

exempt from criticism, especially concerning the consistency of the estimation methods being applied — for instance, the Lyapunov exponent is generally derived according to Wolf's algorithm.

Overall these findings offer valuable insights into the nature of randomness, suggesting how it can both arise from stochastic and deterministic systems, but that eventually, these two paradigms may appear indistinguishable in the context of high-dimensional problems — especially by applying the surrogate data technique. Indeed, coherently with the broader literature on nonlinear time series in epilepsy, it has emerged from various studies on the UKB dataset — [Kannathal, Acharya, et al., 2005], [Kannathal, Choo, et al., 2005] and [Adeli et al., 2007]^(v) among the others — that normal brain activity EEGs (A and B) compared to the ictal phase one (E), not only exhibit a higher complexity but also more chaotic behaviours — through higher estimates of the maximum Lyapunov exponents. However, since this chaoticity operationally resembles a purely stochastic process, the chaotic hypothesis has then been often rejected.

As a final consideration, it is worth noticing though that most of the cited literature in this chapter, does not involve statistically rigorous estimation methods and it is precisely this reason to leads us to the following chapter in which we propose a statistical-based chaotic time series analysis of human EEGs, which we hope will give us a better comprehension of the chaos paradigm.

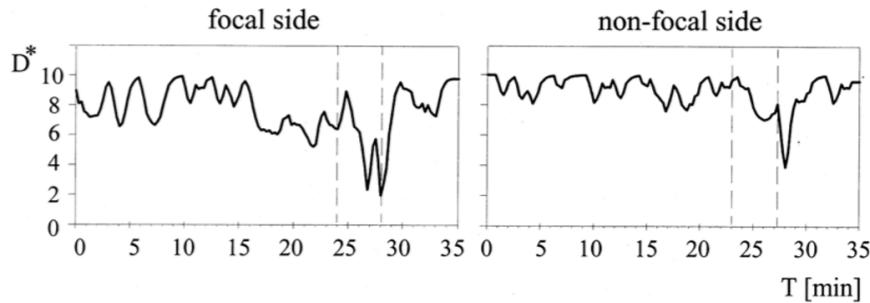


Figure 3.3: The correlation dimension profiles for both the focal and non-focal sides exhibit a remarkable decrease in the seizure onset neighbourhood. This distinct pattern is observed in the frontal, lateral, and medial temporal lobes, with a pronounced effect in the hippocampal formation, the one being here displayed. The integral of such a correlation dimension profile corresponds to the *complexity loss* measure L^* . Image credit: [1995]

^(v)After applying the wavelet decomposition, it emerges that the largest Lyapunov exponents associated with the beta and gamma sub-bands make the healthy group's signals (A and B) distinguishable from ictal phase EEG (E), which are instead characterised by higher exponents for the alpha sub-band.

Chapter 4

Chaos detection in human EEGs: A Lyapunov exponents Approach

In the previous chapters, statistically rigorous methods for the characterisation of chaos — as the sensitivity to initial conditions — have been introduced. Moreover, it has also become evident that the limited application of such methods in much of the past literature hinders conclusive indications regarding the presence of nonlinear and chaotic dynamics in the human brain. In this light, the substantial differences in terms of attractor invariants that have been observed for similar analysis, even with the same dataset, suggest the inappropriateness for EEG signals of traditional nonlinear time series heuristics and tools, since they are deeply affected by both noise contamination [Osborne & Provenzale, 1989; Ruelle, 1990] and the time series embedding [Das et al., 2001].

In the present chapter, following a preliminary analysis in both the time and frequency domains, it will be carried out a consistent estimation of the maximum Lyapunov exponent, applying to the UKB dataset the indirect ANN-based method described in the first chapter. The ultimate purpose of the analysis is to assess the presence of global chaotic dynamics and/or local instability in human EEGs, aiming to provide insights into the main differences between epileptic and healthy subjects' brain dynamics.

4.1 Preliminary Analysis

Epilepsy, as discussed in the second chapter, involves significant changes in the EEG waveform, in particular, large amplitude oscillations during the ictal stage, as a result of a widespread hyper-synchronization of brain regions. In the data we use for our analysis, the UKB dataset, even a preliminary visual inspection reveals that the epileptic

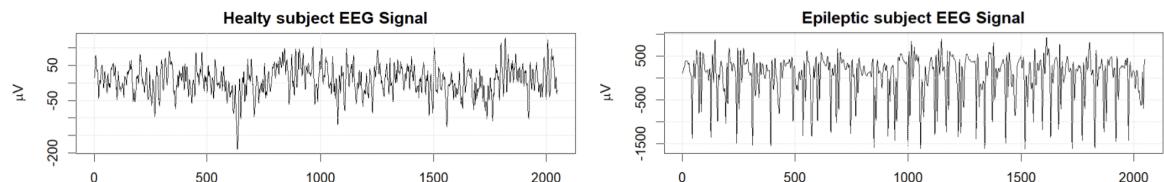


Figure 4.1: Time series of a healthy subject EEG (*A_1*) and an epileptic subject EEG (*E_1*). In each plot, the first 2048 observations are displayed, representing half of the original signal length.

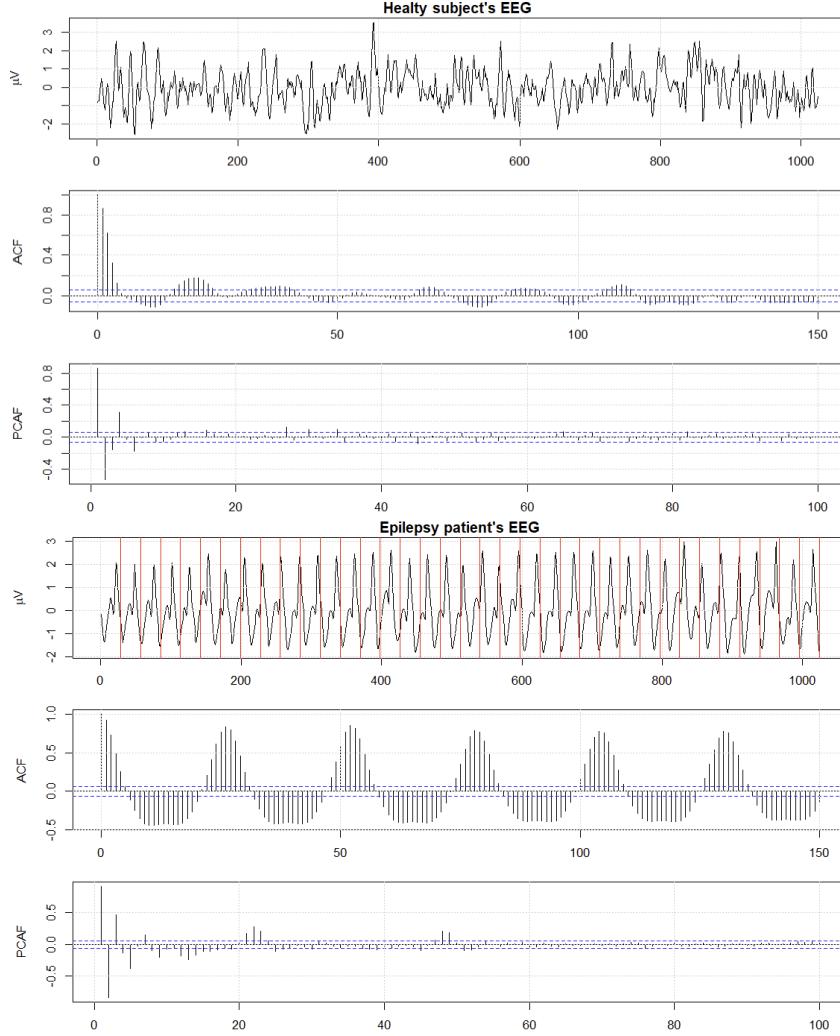


Figure 4.2: Standardized time series, autocorrelation function (ACF), and partial autocorrelation function (PACF) for a healthy subject’s EEG (*A*_4) and an ictal stage EEG (*E*_20). The vertical red lines, displayed at a frequency corresponding to the maximum peak in the associated spectrogram, emphasize the pronounced regularity in the oscillatory pattern of ictal phase EEGs.

EEGs are characterised by more regularly occurring spikes, which lead to a substantial increase of both the overall range and the interquartile range (IQR) of the time series value distributions. Indeed, the IQR goes from an average of 53.86 for open-eye EEGs (A) to around 391.64 for ictal stage recordings (E).

Another distinct feature of EEGs is that despite exhibiting highly random behaviours, these signals retain a strong memory, which does not monotonically vanish. This fact can be better appreciated by both observing the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, which show remarkable time dependences (Figure 4.2). In particular, the systematic presence of peaks in the PACF for even large time lags appears incompatible with traditional linear models — AR, MA, ARMA — in which PACF values tend to monotonically decrease. Instead, these peculiar wavy patterns in the ACF and PACF, characterising all the groups, particularly the epilepsy patients’ EEGs (E), seem more consistent with some seasonal stochastic processes exhibiting remarkable periodicity.

Table 4.1: Mean and variance of the Power Spectral Density across sub-bands for different EEGs: open eyes (A), closed eyes (B), interictal (C and D), and ictal (E).

Frequency Band	Open eyes (A)	Closed eyes (B)	Interictal (C)	Interictal (D)	Ictal (E)
Delta (0.5-4 Hz)	8.66 (3.6)	4.53 (4.3)	14.84 (4.6)	13.55 (10.2)	6.21 (25.6)
Theta (4-8 Hz)	3.50 (1.2)	2.68 (0.7)	4.16 (2.2)	5.10 (6.2)	8.54 (24.1)
Alpha (8-13 Hz)	3.52 (8.2e-1)	7.63 (12.5)	1.15 (3.3e-1)	1.67 (1.4)	3.02 (1.6)
Beta (13-30 Hz)	9.11e-1 (1.4e-1)	1.02 (3.8e-1)	1.98e-1 (2.4e-2)	1.99e-1 (1.2e-2)	8.59e-2 (5.4e-1)
Gamma (30-60 Hz)	3.31e-2 (7.7e-4)	2.84e-2 (7.7e-4)	9.42e-3 (6.1e-5)	7.78e-3 (3.2e-5)	1.14e-3 (2.0e-4)
Noise (> 60 Hz)	9.42e-3 (1.5e-4)	3.30e-3 (1.6e-5)	4.54e-3 (1.4e-5)	3.76e-3 (1.2e-5)	1.44e-3 (1.4e-5)

Spectral Analysis

Besides the limitations of linear methods for EEG signals discussed in the previous chapter, interesting information may still arise from the spectral analysis, especially for what concerns periodicities. Moreover, due to the choice of continuing the analysis with standardised data and the fact that EEG segments were cut to be weakly stationary, the time series are all zero-mean stationary and so the fundamental requirements to carry on a statistically appropriate spectral analysis are satisfied [Brockwell & Davis, 2009]. For this reason, a nonparametric estimate of the spectral density is performed and then smoothed by applying the Daniell kernel method, which applies the moving average to the periodogram values over a 50-point window. Due to the 173.61 Hz sampling rate, the Nyquist sampling theorem establishes that the upper useful frequency is fixed to 86.81 Hz, so half the sampling rate [Belzile, 2019].

Table 4.1 summarises the mean power spectral density associated with each brain frequency band, where, as per convention, frequencies higher than 60 Hz are considered as noise [Adeli et al., 2007; Mirzaei et al., 2011]. At first glance, it is possible to notice how open-eyes (A) and closed-eyes (B) EEGs, on the one hand, and interictal recordings (C and D), on the other, tend to behave similarly. On the contrary, ictal stage EEGs appear to be more erratic, and except for the sub-band *alpha* and *noise*, such recordings are characterised by a much higher within-group variance; see Figure 4.3 for a graphical response.

By reviewing the various frequencies in more detail, we can observe how interictal stage EEGs (C and D) have a substantially higher *delta waves* (0.5-4 Hz) component compared to both healthy subjects' EEGs (A and B) and ictal phase recordings (E). Besides large within-group variances, the one of the ictal stage is particularly high and prevents a proper characterisation of the group. For *theta waves* (4-8 Hz), the distinction between healthy subjects' recordings (A and B) and interictal stage ones (C and D) is less clear but still possible, indeed, seizure-free recordings (C and D) have generally a higher power spectral density. A peculiar behaviour for seizure stage EEGs (E) also occurs at the *theta* sub-band, where, even though the variation range includes the spectral densities of all other groups, the boxplot as a whole suggests a large impact of these frequencies on ictal stage signals. In the *alpha waves* sub-band (8-13 Hz), it stands out the striking behaviour of closed-eyes signals (B), which, coherently with the existing literature, reach here a peak as a consequence of a calm and relaxed state [Adeli et al., 2007]. In addition, a larger degree of between-group separation is obtained for this sub-band and, in particular, we can see how such frequencies have a modest impact on interictal recordings (C and D) while a more significant one for healthy (A and B) and ictal stage signals (E). *Beta* (13-30 Hz) and *gamma* (30-60 Hz) waves have less interpretable results, and a weak degree of separation between seizure-free (C and

D) and healthy recordings (A and B) is visible, blurred by a certain degree of within-group variability. Finally, from the last frequency sub-band (≥ 60 Hz), we can both notice that open-eyes recordings (A) are effectively much more noisy than closed-eyes ones (B) and also that the amplitude burst produced by the seizure onset suppresses the noise, increasing the signal-to-noise ratio.

A more careful comparative study between open (A) and closed-eyes (B) recordings, provided in Figure 4.4, allows capturing some of the intuition so far discussed, in particular, the lower profile of the power spectral density associated with open-eyes (A) signals for all sub-bands —especially, the one of *delta*, *gamma* and *noise* sub-band. Notable differences with the spectral density of ictal stage recordings may also be appreciated, especially the presence of the highest spectrum peak in the *delta* (0.5-4 Hz) sub-band (the light blue region in Figure 4.4) which is coherent with the ~ 5 Hz spike-and-wave discharges described in the literature [Emmady & Anilkumar, 2023; Jiruska et al., 2010] and that produces a higher degree of synchronisation compared to healthy subjects' EEGs.

As a final remark, a weird peak precisely at 50 Hz arises in all EEG spectrograms (Figure 4.4), except those referring to the ictal stage (E). This kind of peak is not extensively discussed in the literature [Osterhage & Lehnertz, 2007], but there is evidence suggesting that it may be due to electromagnetic interference from the power line system. In the United States, such interference is typically recorded at 60 Hz, while in Europe, where the EEGs were recorded, the electric grid operates at exactly 50 Hz, and so synchronously with this observed peak. Unfortunately, none of the papers applying the UKB dataset reports this anomaly and in some of them, especially the one introducing the dataset [Andrzejak et al., 2001], the interference may be ruled out by the application of a 40 Hz band-pass filter.

4.2 Nonlinear Time Series Analysis

As mentioned earlier, the present analysis aims at providing a statistically sound characterisation of human brain chaotic dynamics through the estimation of the maximum Lyapunov exponent. To achieve this, we employ the indirect ANN-based method described in section 2.1.2.2, which uses a nonparametric regression to first estimate the underlying dynamics and then to assess the average rate of divergence of initially nearby trajectories by computing the Jacobian matrix. [McCaffrey et al., 1992; Nychka et al., 1992; Shintani & Linton, 2004].

Since the time series embedding can be seen as a statistical subset selection problem in the context of nonparametric regression [Chan & Tong, 2001; Giannerini & Rosa, 2004; Giannerini et al., 2020], the ANN-based method does not require a previous attractor reconstruction phase, based on statistically inconsistent — and sometimes tedious to apply — heuristics. Moreover, as anticipated in the previous chapter and further discussed in [Pritchard & Duke, 1995], some non-trivial issues are often encountered in nonlinear brain analysis, such as the presence of noise, nonstationarity, or insufficiently long time series, which may lead to spurious indications of chaos. Fortunately, these requirements are fulfilled in the present analysis, as the nonparametric regression in the ANN-based method allows for filtering out the dynamic noise and the UKB dataset only includes weakly stationary EEG segments, each 4,096 observations long.

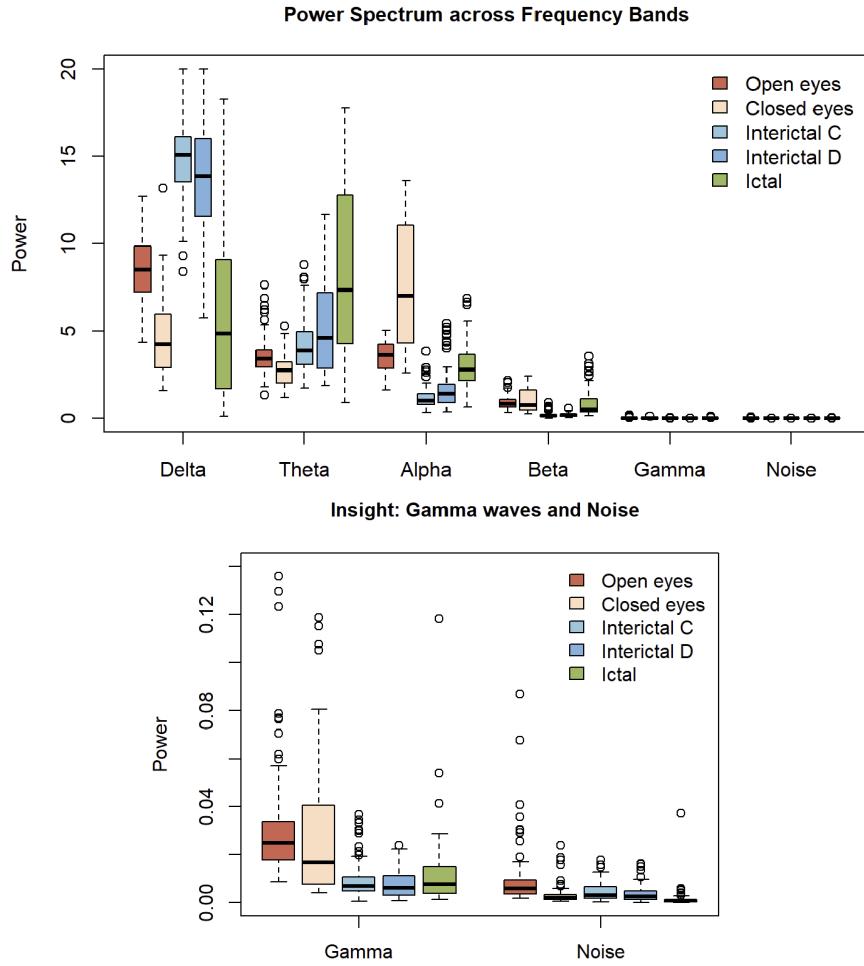


Figure 4.3: Boxplots representing the power spectral density for each sub-band computed from 100 EEGs associated with the five different recording groups in the UKB dataset: open eyes (A), closed eyes (B), interictal (C and D), and ictal (E).

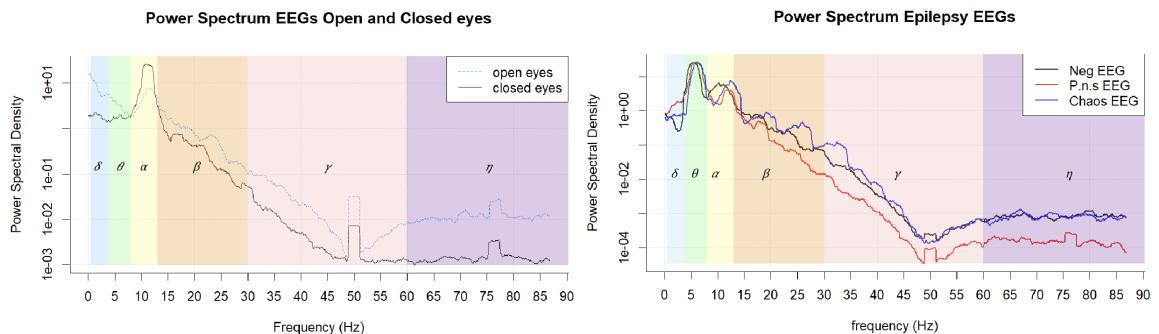


Figure 4.4: The left plot shows the spectrograms of both an open eyes (A_1) and a closed eyes (B_1) EEG recorded from healthy subjects. The right plot depicts the spectrograms of three ictal stage EEGs, classified respectively as Chaos, Neg, and P.n.s. The power spectral density estimates are obtained using a non-parametric technique and smoothed with the Daniell kernel method (50-point window).

Series	m	d	Size	BIC	λ	p -value
A_1_std	35	1	2	-14815.23	-0.003	0.995
B_1_std	24	1	2	-12816.71	-0.029	1.000
C_1_std	4	1	3	-18609.64	-0.144	1.000
D_1_std	6	1	2	-13839.14	-0.099	1.000
E_1_std	10	1	11	-19714.42	0.0349	3.660e-17

Table 4.2: Example of outputs produced by the `netfit` function for the first time series of each of the five recording groups: open eyes (A), closed eyes (B), interictal (C and D), and ictal (E). The Table reports the consistent estimate of the embedding dimension (m), time delay (d), size, BIC, Lyapunov exponent (λ), and the associated p -value.

4.2.1 R Software Implementation

From a practical perspective, this kind of analysis has been made possible by the R implementation of the ANN-based indirect method developed in the private library `tseriesChaos`⁽ⁱ⁾, in which the function `netfit` selects the best model architecture, according to the minimisation of the BIC criterion, within a predefined range parameters — embedding dimensions m , time delays d , and the number of hidden units $size$.

An exploratory analysis reveals that ictal stage (E) and seizure-free (C and D) signals generally are associated with an embedding dimension higher than 8, whereas signals from healthy subjects tend to require a dimensionality of the reconstructed attractor greater than 25. Interestingly, no solution has ever been found with a time delay d different from 1. For as much as it concerns the number of hidden ANN units, it emerges that healthy subjects' and interictal stage EEGs (A, B, C and D) require a modest number of units, generally below 5, while for the ictal stage, this number is around 8-9. Based on the evidence of the preliminary analysis and to avoid sub-optimal minimum solutions, we chose to train the single-hidden-layer ANN for each of the time series on the following grid of parameters:

Parameter	Grid
Embedding dimension (m)	2–45
Time delay (d)	1–5
Number of hidden units ($size$)	1–15

The BIC minimisation yields a consistent estimate of the embedding dimension — which, coherently with the Takens theorem, is a proxy of the system complexity — and also of the maximum Lyapunov exponents, which is accompanied by the p -value resulting from the significance test contained in the `lyap_nnet.test` function; see Section 2.1.3 for a discussion. Table 4.2 proposes an example of the optimal ANN architecture and the MLCE estimate associated with the first EEG segment of each subset, respectively open eyes (A), closed eyes (B), interictal (C and D), and ictal (E).

Due to the significant computational burden associated with estimating a single MLCE over such a large parameter grid, it was not feasible to perform a local analysis on all

⁽ⁱ⁾The indirect ANN-based method is also currently implemented in the open-access Dchaos package [Sandubete & Escot, 2021]

500 time series. Fortunately, we had the opportunity to rely on the computational resources provided by the *Newton server* at the Department of Statistics, University of Bologna, utilizing up to 100 cores at the time through code parallelisation. More specifically, this coding technique allows to break down a complex problem into a series of instructions that are later simultaneously executed on multiple processors, greatly reducing computation time [Laboratory, n.d.]. In our study, the overall optimization of the 500 ANNs using parallelisation took 38.23 hours, compared to at least 120 days that would have been required on a laptop.

4.2.2 System complexity

As a first general result of our analysis, evidence in favour of the "complexity loss hypothesis", which posits that human brain dynamics during the ictal stage (E) exhibit a significantly lower dimensionality compared to the one of healthy subjects (A and B), has emerged. More specifically, as it may be seen from Table 4.3, there is a progressive reduction in the embedding dimension, beginning with the interictal phase (C and D) and further declining during the ictal phase (E). Evidence of such a characterising behaviour are well documented [Lehnertz & Elger, 1995; C. J. Stam, 2005] and since the end of the '90s, as we reported in the previous chapter, it has been thought of exploiting such a dynamic to anticipate seizures — at least with wide time confidence intervals, see [Lehnertz & Elger, 1998; Martiner et al., 1998]. Unfortunately, since the EEGs in the UKB dataset are temporally shuffled and the fact that interictal groups (C and D) both include pre- and post-ictal signals, it has not been possible to properly see the dynamic evolution of how the system dimensionality decreases during seizure onset and how it increases soon after the proper epileptic phase, reaching an even higher level than the one of the pre-ictal phase. Nonetheless, from a static perspective, we can still notice that the dynamics of a healthy brain are associated with the highest mean embedding dimensions, respectively 29.25 for subset A and 27.19 for B. Interictal phase EEGs exhibit an intermediate level of complexity, while the lowest mean embedding dimension, 10.36, is associated with epileptic EEGs (E). Moreover, it is particularly curious and reassuring to see that the sense of sight contributes to the greater complexity of an open-eye brain. Unfortunately, no clear indications about the mean embedding dimension difference for the interictal stages (C and D) have been found, especially in view of the fact that in the literature it has been attested a slightly higher correlation dimension for non-focal side interictal EEGs [Lehnertz & Elger, 1995]. Nonetheless, the difference we have found may actually depend on the fact that the focal side subset (D) has more post-ictal recordings and also on the fact that we merely estimate the dimensionality in which the attractor is reconstructed and not the fractal dimensionality of the attractor itself, which will be necessarily lower for chaotic attractors. In conclusion, we suggest seeing the boxplots on the left-hand side of Figure 4.5 for graphical feedback and we postpone a deeper discussion on this topic for the following section where we will study these differences in relation to the estimates of the Lyapunov exponent.

4.2.3 Global chaotic dynamics

Based on the MLCE estimate and the associated p-value, which results from the chaos test $H_0 : \lambda \leq 0$ [Shintani & Linton, 2004], discussed in Section 2.1.3 and provided by

Table 4.3: Summary of estimation statistics for the EEG recording groups: open eyes (A), closed eyes (B), interictal (C and D), and ictal (E)

Recording Group	Mean Embedding Dim.	Mean MLCE (Mean p-value)	Mean ANN Size
Healthy Open Eyes (A)	29.3	-2.692e-2 (9.46e-1)	2.02
Healthy Closed Eyes (B)	27.2	-1.46e-2 (9.31e-1)	2.45
Non-focal Interictal (C)	15.6	-7.862e-2 (9.86e-1)	2.20
Focal Interictal (D)	17.7	-4.640e-2 (8.93e-1)	4.10
Ictal (E)	10.4	3.810e-3 (4.76e-1)	10.55

Table 4.4: Summary of estimation statistics for the various EEG recording groups and MLCE classes: *Chaos* ($\lambda > 0$, $p\text{-value} < 0.05$), *P.n.s* ($\lambda > 0$, $p\text{-value} > 0.05$), and *Neg* ($\lambda < 0$, $p\text{-value} > 0.05$)

Group	Label	Count	Mean Embedding Dim.	Mean ANN Units	Mean MLCE
A	P.n.s	5	38.0	2.0	2.332e-3
	Neg	95	28.8	2.0	-2.846e-2
B	Chaos	4	16.3	7.8	2.743e-2
	P.n.s	1	24.0	5.0	1.949e-3
	Neg	95	27.7	2.2	-1.651e-2
C	P.n.s	1	44.0	2.0	6.328e-3
	Neg	99	15.3	2.2	-7.948e-2
D	Chaos	9	17.8	12.6	5.627e-2
	P.n.s	2	33.5	3.5	4.312e-3
	Neg	89	17.2	3.3	-5.792
E	Chaos	47	11.1	12.9	4.873e-2
	P.n.s	6	11.3	10.0	4.803e-3
	Neg	47	9.5	8.3	-4.124e-2

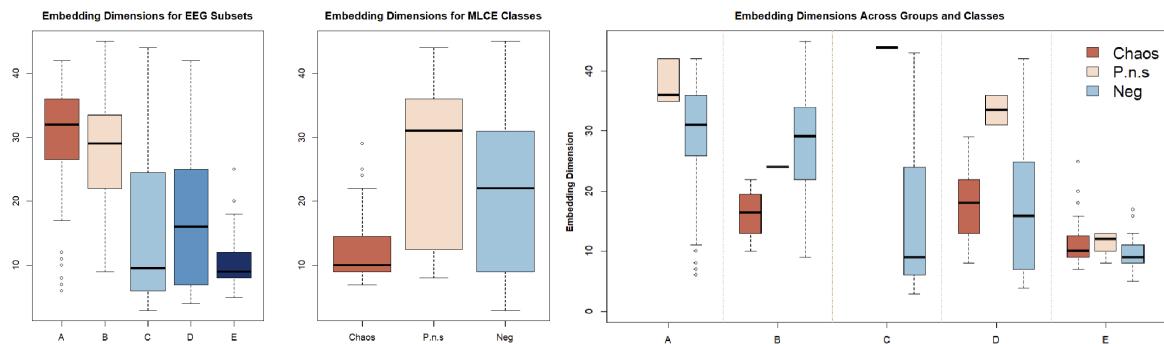


Figure 4.5: The first two panels from the left present respectively the boxplots of the embedding dimensions for each of the EEG recording groups — open eyes (A), closed eyes (B), interictal (C and D), and ictal (E) — and for each of the MLCE classes (Chaos, P.n.s, and Neg). The right plot displays the boxplots of the MLCE for the same group and classes.

the `lyap_nnet.test` function, each of the 500 EEGs is assigned to one of the following three classes:

- i *Chaos*, in case of a statistically significant positive MLCE estimate with 5% confidence level ($p\text{-value} < 0.05$)
- ii *P.n.s* (acronym of positive non-significant estimate) when a MLCE estimate larger than zero is associated with a $p\text{-value}$ larger than 0.05
- iii *Neg*, in the context of a significant negative MLCE estimate ($p\text{-value} \geq 0.05$)

Based on this classification, further information about brain complexity can now be retrieved. In particular, we can notice that *Chaos* labelled signals tend to have a lower embedding dimension (mean value of 12.43) compared to both *Neg* (20.84) and *P.n.s* (26.2) classified EEGs, which also have much larger variabilities — please refer to the central plot in Figure 4.5 for a visual representation. However, by examining the embedding dimension associated with each recording group (A, B, C, D, and E) and MLCE class (*Chaos*, *P.n.s*, and *Neg*), as shown on the right side of Figure 4.5 and in Table 4.4, we see that the attractors of ictal stage EEGs (E) are all reconstructed in Euclidean spaces having approximately the same dimensionality, with *Neg* labelled ictal recordings actually exhibiting a lower embedding dimension compared to chaotic ictal series. Furthermore, it emerges that the dimensionality of the reconstructed state space is negatively correlated with the optimal —actually, sub-optimal — number of units in the estimated ANN architecture; see Table 4.3, Table 4.4, and the panel above in Figure 4.6. Unfortunately, no clear explanation for this has been found yet.

Overall, out of the 500 EEGs analysed, 60 are classified as belonging to the *Chaos* class, with 47 of these being ictal recordings (E). This fact represents the second important result of our analysis, proving that around half of the seizure EEGs (E) we have studied display global chaotic behaviours. In this light, it seems reasonable to object the hypothesis, described in the previous chapter, of ictal stage EEGs as the result of noisy limit cycle dynamics [Feucht et al., 1998; Friedrich & Uhl, 1996]. Moreover, in agreement with [Osterhage & Lehnertz, 2007], we can conclude that the difficulty in detecting global chaotic behaviours in previous works may have actually been due to the lack of statistically rigorous methods being applied.

Of the remaining signals, 425 are assigned to the *Neg* class while 15 to *P.n.s*, which is a sort of borderline set composed of signals with an average MLCE of 0.00382 ($p\text{-value}$ of 0.2199) and displaying a certain degree of periodicity. Larger interpretation uncertainties come instead from the class *Neg*. Indeed, if, on the one hand, the negative MLCE suggests a rejection of the hypothesis of chaotic motion, especially in healthy subjects' EEGs, on the other hand, these signals compared to both *Chaos* and *P.n.s* labelled signals seem to display a significantly higher degree of randomness, probably incompatible with quasi-periodic motion. Elements of apparent strangeness have already emerged in the spectral analysis, where we see that, coherently with the existing literature, ictal stage recordings follow more regular patterns (Figure 4.2). By looking at the significantly lower average embedding dimension of the *Chaos* labelled series, a possible reason for this behaviour may be related to the system dimensionality and to the fact that chaotic dynamics do exist in healthy subjects' EEGs but that they cannot be detected with the Lyapunov exponents in such high-dimensional spaces. Indeed, this hypothesis does not only justify the random waveform of the EEGs but it is also coherent with the existing literature which attests — through direct estimation

approaches — the presence of larger Lyapunov exponents in healthy subjects' EEGs but rejects the global chaotic nature as the EEGs do not significantly differ from a stochastic signal. As a personal note, it seems reasonable to believe that the desire to obtain statistically reliable results and to situate brain dynamics within either the stochastic or the deterministic paradigm, by exploiting the surrogate data method, may have shifted the focus from the actual research on the existence of chaotic or hyperchaotic [Gallez & Babloyantz, 1991] dynamics within the human brain.

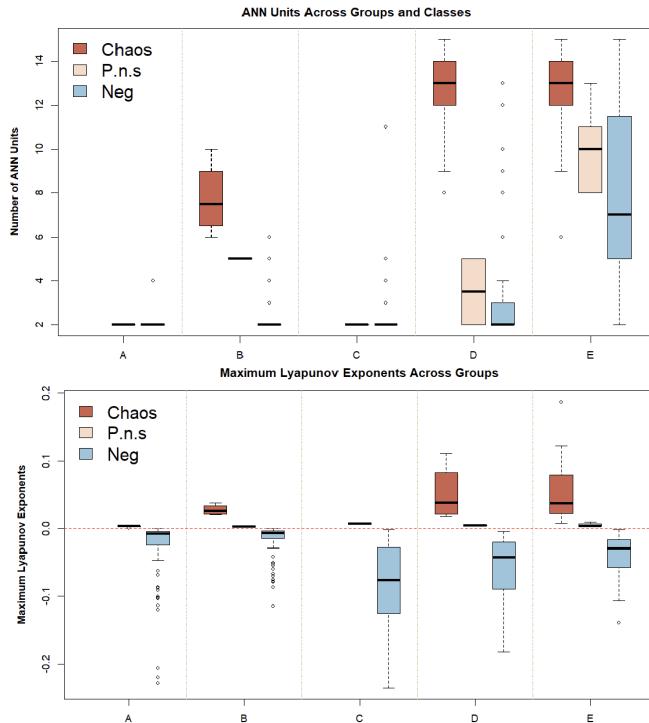


Figure 4.6: The left panel illustrates the embedding dimensions for the various groups — open eyes (A), closed eyes (B), interictal (C and D), and ictal (E) — and MLCE classes (Chaos, P.n.s, and Neg), while the right panel shows the optimal ANN architecture sizes, in terms of units, for the same groups and classes.

4.2.4 Local chaotic dynamics

As discussed in the theoretical Section 2.1.3, the global Lyapunov exponent provides an insight into the initial value sensitivity and, consequently, on the overall chaotic nature of the system. Nonetheless, beyond the scientific interest of determining whether an EEG signal exhibits chaotic dynamics, further information about the system may arise from the local Lyapunov exponent (LLE), also called M-step Lyapunov exponents. Indeed, even a system with a negative MLCE may still present regions characterised by local high nonlinearity and finite-time horizon instability [Bailey, 1996]. Moreover, as highlighted in [Giannerini & Rosa, 2004], specific nonlinear patterns — either at regular periodicity or in the form of peaks in the ascending/descending sections of the time series — could characterise a global non-chaotic time series, explaining much of the underlying process.

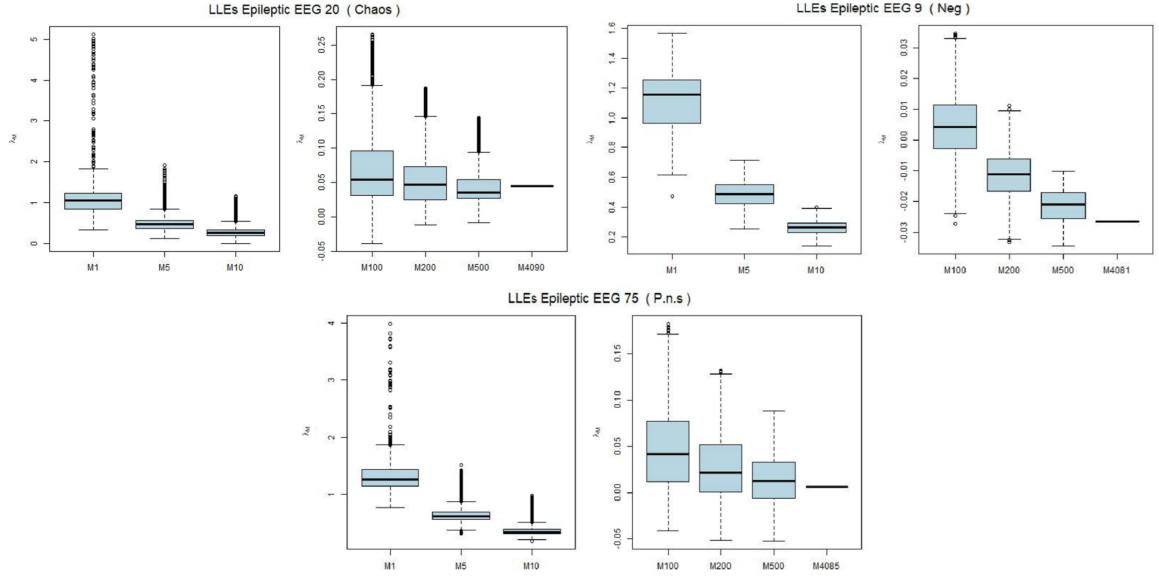


Figure 4.7: Boxplots of the LLEs for three epileptic signals classified as Chaos ($E_{\cdot 20}$), Neg ($E_{\cdot 9}$), and P.n.s ($E_{\cdot 75}$)

Thanks to the `lyap_nnet` function, it has been possible to estimate the LLEs, specifying the desired number of steps ahead M . The boxplots of the LLEs in Figure 4.7 give a tentative idea of both the local rate of divergence of nearby trajectories — for a number of steps ahead equal to 1, 5, 10, 100, 200 — and even the asymptotic one, by considering M as the number of embedding vectors minus 1 — which will ultimately depend on the characteristics of the optimal reconstructed state space. In particular, we may observe that, for a small number of steps ahead, the initial value sensitivity is not a prerogative of *Chaos*-labelled signals and even for 100 steps ahead, significant tail portions of the boxplots still fall into the positive region. This intuition is also confirmed by Figure 4.8 where we see that a large portion of the states in the reconstructed attractors display local exponential divergence of trajectories even at 200 steps ahead — red dots are associated with positive LLEs. However, this behaviour of the LLEs is not exclusive to epilepsy subjects' EEGs (E) but characterises all the other groups, indicating a general degree of system dynamics heterogeneity and a considerable nonlinearity [Bailey, 1996]. Unfortunately, with very few exceptions (see in Figure 4.9 the 30th EEG signal from subset D), it has not been possible to find either distinctive LLE patterns along the EEG time series or specific regions in the reconstructed attractor displaying local instability; please refer to Figure 4.7. Moreover, no recurring pattern characterises any recording group (A, B, C, D, E), and no remarkable difference distinguishes the various MLCE classes (*Chaos*, *Neg*, *P.n.s*). Indeed, regardless of the number of step-ahead, regions with positive LLEs seem to appear randomly along the EEG signals or within the reconstructed attractor.

This impression is further corroborated by considering the 10th and 90th percentiles of the LLEs distributions, where, once again, no patterns occur within the reconstructed attractor or along the EEG time series (Figure 4.9).

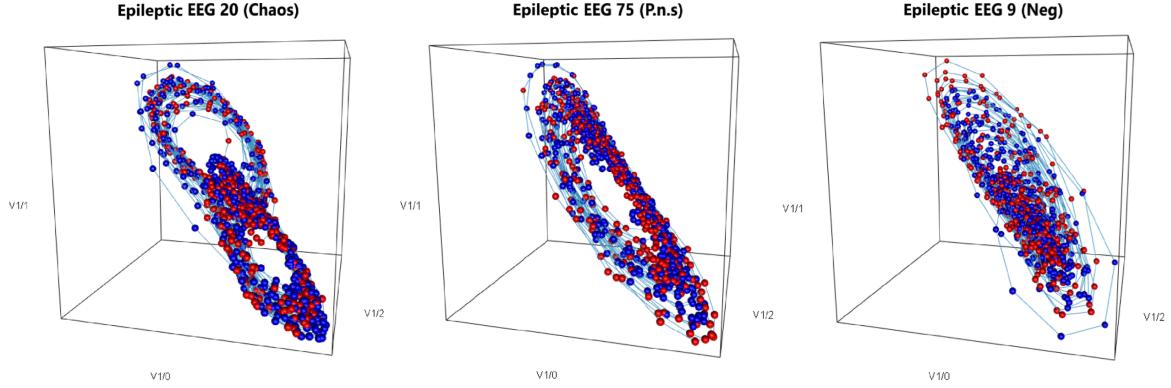


Figure 4.8: 3D representation of the reconstructed attractors for three epileptic signals classified as Chaos (E_{-20}), Neg (E_{-9}), and P.n.s (E_{-75}). Red points indicate states associated with a positive 200-step-ahead LLE.

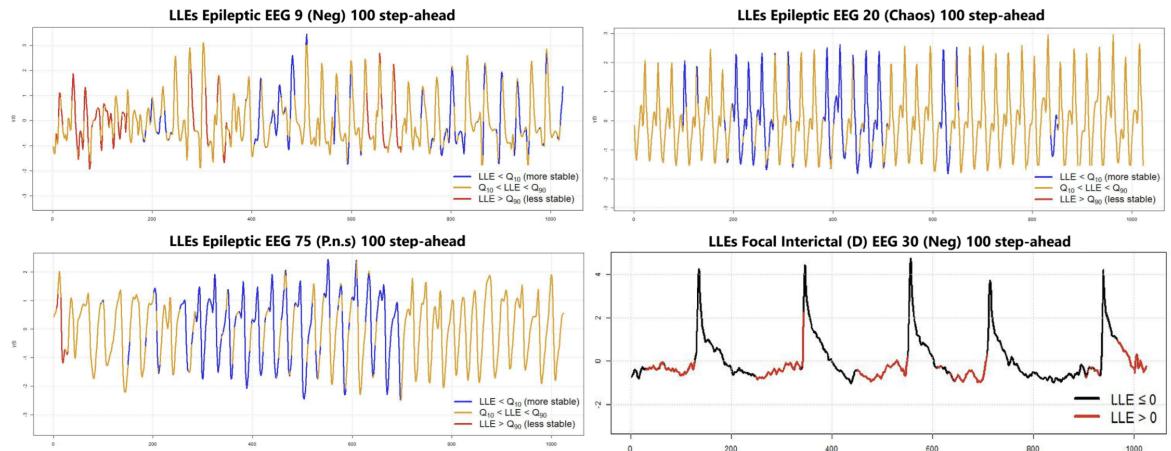


Figure 4.9: Standardised time series for three ictal stage EEGs, each belonging to one of the MLCE classes: Chaos (E_{-20}), Neg (E_{-9}), and P.n.s (E_{-75}). The time series values are coloured according to the associated 100-step LLE: red for the top 10% of the LLEs distribution; blue for the lowest 10%, associated with more stable dynamics and a lower degree of exponential divergence; and yellow for the intermediate 80%. In the lower right, the time series of a focal interictal EEG (E_{-30}) displaying some LLE-pattern is shown. Values are red when associated with a positive 100-step LLE.

As a final consideration, it is worth noting that part of the difficulty in identifying patterns of local instability among EEGs may be addressed to an intrinsic heterogeneity of the signals, which display very different patterns. Unfortunately, the absence of temporal and patient identification, as well as a lack of neuroscientific knowledge of the undersigned, prevent a more in-depth analysis of the results. Nonetheless, also the analysis of the LLEs has revealed a great heterogeneity and an extremely high complexity of human brain dynamics.

4.3 Conclusion

The high dimensionality and the inherent nonlinearity, even at a neural level, of human brain dynamics raise the question of whether brain electrical activity recordings — the

electroencephalograms, EEGs — exhibit chaos. Since the mid-1980s, scientific research has focused on the estimation of the Lyapunov exponent, an attractor invariant, which measures a fundamental property of chaos: sensitivity to initial conditions. Furthermore, the presence of intricate nonlinear dynamics in human EEGs precludes the application of both parametric and nonparametric linear signal processing tools as they provide poor medical interpretations and prevent any system forecasting. Consequently, nonlinear time series and chaos theory tools have found a new field of application in studying a large variety of mental states or conditions, such as mathematical reasoning, eye closure, alcohol consumption, or mental disorders like schizophrenia, Parkinson's, and epilepsy, just to name a few. However, the conflicting needs for long, stationary, and noise-free EEG signals, together with a lack of statistically rigorous estimation methods, have led to contradictory and inconclusive findings, especially regarding the existence of global chaotic dynamics.

This dissertation offers a theoretical overview of the genesis of chaos and a comprehensive review of the literature on nonlinear EEG analysis, with a particular focus on epilepsy. It then investigates the existence of global dynamics and/or local instability in human EEG by estimating the Lyapunov exponent through an indirect nonparametric neural network approach. Advantages in favour of this method lie in the statistically consistent results and the determination of the asymptotic distribution, which enables the definition of a chaos test [Shintani & Linton, 2004].

The analysis developed in the present dissertation draws on the UBK dataset, which was made available online by the Epilepsy Center at the University of Bonn (Universitätsklinikum Bonn), and it uncovers two main findings. First, it has been shown a gradual reduction of brain dynamics complexity — in terms of embedding dimension — which begins in the healthy brain state, decreases further in the interictal phase, and reaches its lowest level during the ictal phase. Second, it has emerged that ictal stage signals do display low-dimensional chaos and it is thus possible to corroborate a theory proposed in the literature but not conclusively proven. In contrast, no clear indications of chaos have arisen from healthy subjects' EEGs, and a possible reason for this may stand in the fact that chaotic dynamics do exist but that they cannot be detected with the Lyapunov exponents in such high-dimensional spaces. The study of the local Lyapunov exponents also reveals a great heterogeneity in the degree of local exponential divergence. However, even if most healthy, interictal and ictal stage EEGs display local instability, no recurring patterns in the time series or distinct regions in the reconstructed attractor are characterised by a local exponential divergence of nearby trajectories, with positive LLE states randomly appearing along the EEG signals.

It is important to note, though, that the limitations related to the UKB dataset have deeply affected the analysis and the comprehension of the underlying process. Notably, the absence of differentiation between pre- and post-ictal EEGs impedes to thoroughly study the dynamic changes in the system complexity and the confirmation of previous findings indicating that the post-ictal EEGs have higher dimensionality than both pre-ictal and ictal stage signals.

We believe that significant work remains to be done, in particular since EEGs — and especially extracranial recordings — contain noise, it may be worth evaluating the Lypaunov-like indexes, discussed in Section 2.2, which are specifically designed to measure initial-value sensitivity within the stochastic dynamical system framework. Furthermore, the employment of datasets providing patient identification for the stud-

ied signals may provide further insights into the seizure development mechanism. In this regard, we can envision a new season of research focused on forecasting and epilepsy prevention techniques based on statistically consistent chaos theory tools.

In conclusion, despite the theoretical challenges still surrounding the field of stochastic dynamical systems, above all the absence of a formal definition of chaos in such a framework, this dissertation has underpinned the necessity of adopting robust and statistically consistent estimation methods to reliably assess the existence of both chaotic dynamics and/or local instability in human EEGs.

Bibliography

- Abarbanel, Brown, R., & Kennel, M. B. (1991). Variation of Lyapunov Exponents on a Strange Attractor. *Journal of Nonlinear Science*, 1(2), 175–199. <https://doi.org/10.1007/bf01209065>
- Abarbanel, Brown, R., & Kennel, M. B. (1992). Local Lyapunov Exponents Computed From Observed Data. *Journal of Nonlinear Science*, 2(3), 343–365. <https://doi.org/10.1007/bf01208929>
- Abarbanel, H. (2012, December). *Analysis of observed chaotic data*. Springer Science Business Media.
- Achermann, P., Hartmann, R., Gunzinger, A., Guggenbüh, W., & Borbély, A. A. (1994). Correlation Dimension of the Human Sleep Electroencephalogram: Cyclic Changes in the Course of the Night. *European Journal of Neuroscience*, 6(3), 497–500. <https://doi.org/10.1111/j.1460-9568.1994.tb00292.x>
- Adeli, H., Ghosh-Dastidar, S., & Dadmehr, N. (2007). A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy. *IEEE Transactions on Biomedical Engineering*, 54(2), 205–211. <https://doi.org/10.1109/TBME.2006.886855>
- Albano, A. M., Abraham, N. B., De Guzman, G. C., Tarroja, M. F. H., Bandy, D. K., Gioggia, R. S., Rapp, P. E., Zimmerman, I. D., Greenbaum, N. N., & Bashore, T. R. (1986, January). *Lasers and Brains: Complex Systems With Low-Dimensional Attractors*. https://doi.org/10.1007/978-3-642-71001-8\{\}_27
- Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E*, 64, 061907. <https://doi.org/10.1103/PhysRevE.64.061907>
- Babloyantz, A., & Destexhe, A. (1986). Low-dimensional Chaos in an Instance of Epilepsy. *Proceedings of the National Academy of Sciences*, 83(10), 3513–3517. <https://doi.org/10.1073/pnas.83.10.3513>
- Babloyantz, A., Salazar, J., & Nicolis, C. (1985). Evidence of Chaotic Dynamics of Brain Activity During the Sleep Cycle. *Physics Letters A*, 111(3), 152–156. [https://doi.org/10.1016/0375-9601\(85\)90444-x](https://doi.org/10.1016/0375-9601(85)90444-x)
- Bailey, B. A., Ellner, S. P., & Nychka, D. W. (1997). Chaos with confidence: Asymptotics and applications of local lyapunov exponents. <https://api.semanticscholar.org/CorpusID:118264670>
- Bailey, B. A. (1996). *Asymptotics and applications of local lyapunov exponents* [PhD thesis]. North Carolina State University.
- Baker, G. L., & Gollub, J. P. (1996, January). *Chaotic dynamics*. Cambridge University Press.
- Belzile, L. (2019). Time series (math 342) [Version of 2019-02-26].

- Benettin, G., Galgani, L., Giorgilli, A., & Strelcyn, J.-M. (1980). Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems - a method for computing all of them. i - theory. ii - numerical application. *Meccanica*, 15, 21–30. <https://doi.org/10.1007/BF02128236>
- Brari, Z., & Belghith, S. (2022). A New Algorithm for Largest Lyapunov Exponent Determination for Noisy Chaotic Signal Studies With Application to Electroencephalographic Signals Analysis for Epilepsy and Epileptic Seizures Detection. *Chaos Solitons Fractals*, 165, 112757. <https://doi.org/10.1016/j.chaos.2022.112757>
- Brockwell, P. J., & Davis, R. A. (2009, May). *Time Series: Theory and Methods*. Springer Science Business Media.
- Brown, R., Bryant, P., & Abarbanel, H. D. I. (1991). Computing the lyapunov spectrum of a dynamical system from an observed time series. *Phys. Rev. A*, 43, 2787–2806. <https://doi.org/10.1103/PhysRevA.43.2787>
- Casdagli, Iasemidis, L. D., Savit, R. S., Gilmore, R. L., Roper, S. N., & Sackellares, J. C. (1997). Non-linearity in Invasive EEG Recordings From Patients With Temporal Lobe Epilepsy. *Electroencephalography and Clinical Neurophysiology*, 102(2), 98–105. [https://doi.org/10.1016/s0921-884x\(96\)95195-4](https://doi.org/10.1016/s0921-884x(96)95195-4)
- Casdagli, M. (1989). Nonlinear Prediction of Chaotic Time Series. *Physica D Nonlinear Phenomena*, 35(3), 335–356. [https://doi.org/10.1016/0167-2789\(89\)90074-2](https://doi.org/10.1016/0167-2789(89)90074-2)
- Chan, K. S., & Tong, H. (1994). A note on noisy chaos. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(2), 301–311. Retrieved October 15, 2024, from <http://www.jstor.org/stable/2345901>
- Chan & Tong. (2001, January). *Chaos: A Statistical Perspective*. <https://doi.org/10.1007/978-1-4757-3464-5>
- Chatterjee, S., & Yilmaz, M. R. (1992). Chaos, Fractals and Statistics. *Statistical Science*, 7(1). <https://doi.org/10.1214/ss/1177011443>
- Cohen, M. X. (2014, January). *Analyzing Neural Time Series Data*. <https://doi.org/10.7551/mitpress/9609.001.0001>
- Das, A., Das, P., & Roy, A. (2001). Applicability of lyapunov exponent in eeg data analysis. *Complexity International [electronic only]*, 9.
- Dechert, W. D., & Gencay, R. (1992). Lyapunov Exponents as a Nonparametric Diagnostic for Stability Analysis. *Journal of Applied Econometrics*, 7(S1), S41–S60. <https://doi.org/10.1002/jae.3950070505>
- Eckmann, J. .-, Kamphorst, S. O., Ruelle, D., & Ciliberto, S. (1986). Liapunov exponents from time series. *Physical Review. A, General Physics*, 34(6), 4971–4979. <https://doi.org/10.1103/physreva.34.4971>
- Eckmann, J. .-, & Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57, 617–656. <https://doi.org/10.1103/RevModPhys.57.617>
- Elger, C. E., & Lehnertz, K. (1998). Seizure Prediction by Non-linear Time Series Analysis of Brain Electrical Activity. *European Journal of Neuroscience*, 10(2), 786–789. <https://doi.org/10.1046/j.1460-9568.1998.00090.x>
- Emmady, P. D., & Anilkumar, A. C. (2023). *Eeg abnormal waveforms* (Updated 2023 May 1) [Available from: StatPearls [Internet]]. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK557655>
- Epilepsy Foundation Australia. (2024). *Seizure phases* [Accessed: 10 October 2024]. <https://epilepsyfoundation.org.au/understanding-epilepsy/seizures/seizure-phases/>

- Fan, J., & Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(2), 371–394. Retrieved October 15, 2024, from <http://www.jstor.org/stable/2345968>
- Fell, J., Röschke, J., & Beckmann, P. (1993). Deterministic Chaos and the First Positive Lyapunov Exponent: A Nonlinear Analysis of the Human Electroencephalogram During Sleep. *Biological Cybernetics*, 69(2), 139–146. <https://doi.org/10.1007/bf00226197>
- Feucht, M., Möller, U., Witte, H., Schiecke, K., Arnold, M., Benninger, F., Steinberger, K., & Friedrich, M. (1998). Nonlinear dynamic of 3 hz spike-and-wave discharges recorded during typical absence seizures in children. *Cerebral cortex (New York, N.Y. : 1991)*, 8, 524–33. <https://doi.org/10.1093/cercor/8.6.524>
- Fisher, R. S., Van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., & Engel, J. (2005). Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4), 470–472. <https://doi.org/10.1111/j.0013-9580.2005.66104.x>
- Frank, G. W., Lookman, T., Nerenberg, M. A. H., Essex, C., Lemieux, J., & Blume, W. (1990). Chaotic Time Series Analyses of Epileptic Seizures. *Physica D Nonlinear Phenomena*, 46(3), 427–438. [https://doi.org/10.1016/0167-2789\(90\)90103-v](https://doi.org/10.1016/0167-2789(90)90103-v)
- Friedrich, R., & Uhl, C. (1996). Spatio-temporal Analysis of Human Electroencephalograms: Petit-mal Epilepsy. *Physica D Nonlinear Phenomena*, 98(1), 171–182. [https://doi.org/10.1016/0167-2789\(96\)00059-0](https://doi.org/10.1016/0167-2789(96)00059-0)
- Gallez, D., & Babloyantz, A. (1991). Predictability of Human EEG: A Dynamical Approach. *Biological Cybernetics*, 64(5), 381–391. <https://doi.org/10.1007/bf00224705>
- Giannerini, S., Maasoumi, E., & Dagum, E. B. (2015). Entropy testing for nonlinear serial dependence in time series. *Biometrika*, 102(3), 661–675. <https://doi.org/10.1093/biomet/asv007>
- Giannerini, S., Rojo, B., Salas de Leon, D., Monreal-Gómez, M., & Cartwright, J. (2020). Chaos and periodicities in a climatic time series of the iberian margin. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30, 063126. <https://doi.org/10.1063/1.5123509>
- Giannerini, S., & Rosa, R. (2002). Caos, statistica e metodi di ricampionamento. *DOAJ (DOAJ: Directory of Open Access Journals)*. <https://doi.org/10.6092/issn.1973-2201/414>
- Giannerini, S., & Rosa, R. (2004). Assessing chaos in time series: Statistical aspects and perspectives. *Studies in Nonlinear Dynamics and Econometrics*, 8(2). <https://doi.org/10.2202/1558-3708.1215>
- Grassberger, P., & Procaccia, I. (1983). Characterization of strange attractors. *Phys. Rev. Lett.*, 50, 346–349. <https://doi.org/10.1103/PhysRevLett.50.346>
- Grebogi, C., Ott, E., Pelikan, S., & Yorke, J. A. (1984). Strange attractors that are not chaotic. *Physica D: Nonlinear Phenomena*, 13(1), 261–268. [https://doi.org/10.1016/0167-2789\(84\)90282-3](https://doi.org/10.1016/0167-2789(84)90282-3)
- Gregson, R. A., Britton, L., Campbell, E. A., & Gates, G. (1990). Comparisons of the nonlinear dynamics of electroencephalograms under various task loading conditions: A preliminary report. *Biological Psychology*, 31(2), 173–191. [https://doi.org/https://doi.org/10.1016/0301-0511\(90\)90016-P](https://doi.org/https://doi.org/10.1016/0301-0511(90)90016-P)

- Grimmett, G., & Stirzaker, D. (2001). *Probability and random processes*. Oxford University Press. http://www.worldcat.org/search?qt=worldcat.org_all&q=9780198572220
- Hadeaghi, F. (2023). Eeg - a brief introduction [Lecture Notes for BSP, Chapter 4, Master Program Data Engineering]. https://ai.rug.nl/minds/uploads/Chapter_4_EEG_Introduction.pdf
- Hegger, R., Kantz, H., & Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The tisean package. *Chaos*, 9(2), 413–435. <https://doi.org/10.1063/1.166424>
- Huffaker, R., Bittelli, M., & Rosa, R. (2018, February). *Nonlinear time series analysis with r*. <https://doi.org/10.1093/oso/9780198782933.001.0001>
- Iasemidis, L. D. (2011). Seizure Prediction and Its Applications. *Neurosurgery Clinics of North America*, 22(4), 489–506. <https://doi.org/10.1016/j.nec.2011.07.004>
- Iasemidis, L. D., Sackellares, J. C., Zaveri, H. P., & Williams, W. J. (1990). Phase Space Topography and the Lyapunov Exponent of Electrocorticograms in Partial Seizures. *Brain Topography*, 2(3), 187–201. <https://doi.org/10.1007/bf01140588>
- Jiruska, P., Finnerty, G. T., Powell, A. D., Lofti, N., Cmejla, R., & Jefferys, J. G. R. (2010). Epileptic high-frequency network activity in a model of non-lesional temporal lobe epilepsy. *Brain*, 133(5), 1380–1390. <https://doi.org/10.1093/brain/awq070>
- Kannathal, N., Acharya, U. R., Lim, C. M., & Sadashivan, P. K. (2005). Characterization of eeg-a comparative study. *Comput. Methods Prog. Biomed.*, 80(1), 17–23. <https://doi.org/10.1016/j.cmpb.2005.06.005>
- Kannathal, N., Choo, M., Acharya, U., & Puthusserypady, S. (2005). Entropies for detection of epilepsy in eeg. *Computer Methods and Programs in Biomedicine*, 80(3), 187–194.
- Kantz, H. (1994). A robust method to estimate the maximal lyapunov exponent of a time series. *Physics Letters. A*, 185(1), 77–87. [https://doi.org/10.1016/0375-9601\(94\)90991-1](https://doi.org/10.1016/0375-9601(94)90991-1)
- Kantz, H., & Schreiber, T. (2004, January). *Nonlinear time series analysis*. Cambridge University Press.
- King-Hang, M., MA, Fong, M. C.-M., Xie, C., Lee, T., Chen, G., & Wang, W. S. (2021). Regularity and Randomness in Ageing: Differences in Resting-state EEG Complexity Measured by Largest Lyapunov Exponent. *Neuroimage Reports*, 1(4), 100054. <https://doi.org/10.1016/j.ynirp.2021.100054>
- Kosloff, R., & Rice, S. A. (1981). Dynamical correlations and chaos in classical hamiltonian systems. *The Journal of Chemical Physics*, 74, 1947–1955.
- Kunhimangalam, R., Joseph, P., & Sujith, O. (2007). Nonlinear Analysis of EEG Signals: Surrogate Data Analysis. *IRBM*, 29(4), 239–244. <https://doi.org/10.1016/j.rbmret.2007.09.006>
- Kutepov, I. E., Dobriyan, V. V., Zhigalov, M. V., Stepanov, M. F., Krysko, A. V., Yakovleva, T. V., & Krysko, V. A. (2020). Eeg analysis in patients with schizophrenia based on lyapunov exponents. *Informatics in Medicine Unlocked*, 18, 100289. <https://doi.org/10.1016/j.imu.2020.100289>
- Laboratory, L. L. N. (n.d.). Introduction to parallel computing tutorial [Accessed: 2024-10-13].

- Layne, S. P., Mayer-Kress, G., & Holzfuss, J. (1986, January). *Problems Associated With Dimensional Analysis of Electroencephalogram Data*. https://doi.org/10.1007/978-3-642-71001-8\{\}_29
- Lee, Y.-J., Zhu, Y.-S., Xu, Y.-H., Shen, M.-F., Zhang, H.-X., & Thakor, N. (2001). Detection of Non-linearity in the EEG of Schizophrenic Patients. *Clinical Neurophysiology*, 112(7), 1288–1294. [https://doi.org/10.1016/s1388-2457\(01\)00544-2](https://doi.org/10.1016/s1388-2457(01)00544-2)
- Lehnertz, K., & Elger, C. E. (1997). Neuronal Complexity Loss in Temporal Lobe Epilepsy: Effects of Carbamazepine on the Dynamics of the Epileptogenic Focus. *Electroencephalography and Clinical Neurophysiology*, 103(3), 376–380. [https://doi.org/10.1016/s0013-4694\(97\)00027-1](https://doi.org/10.1016/s0013-4694(97)00027-1)
- Lehnertz, K., & Elger, C. (1995). Spatio-temporal Dynamics of the Primary Epileptogenic Area in Temporal Lobe Epilepsy Characterized by Neuronal Complexity Loss. *Electroencephalography and Clinical Neurophysiology*, 95(2), 108–117. [https://doi.org/10.1016/0013-4694\(95\)00071-6](https://doi.org/10.1016/0013-4694(95)00071-6)
- Lehnertz, K. (2008). Epilepsy and nonlinear dynamics. *Journal of biological physics*, 34, 253–66. <https://doi.org/10.1007/s10867-008-9090-3>
- Lehnertz, K., & Elger, C. E. (1998). Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity. *Phys. Rev. Lett.*, 80, 5019–5022. <https://doi.org/10.1103/PhysRevLett.80.5019>
- Liu, Z. (2010). Chaotic time series analysis. *Mathematical Problems in Engineering*, 2010(1). <https://doi.org/10.1155/2010/720190>
- Longtin, A. (2010). Stochastic Dynamical Systems. *Scholarpedia*, 5(4), 1619. <https://doi.org/10.4249/scholarpedia.1619>
- Mandelbrot, B. B., & Wheeler, J. A. (1983). The Fractal Geometry of Nature. *American Journal of Physics*, 51(3), 286–287. <https://doi.org/10.1119/1.13295>
- Martinerie, J., Adam, C., Van Quyen, M. L., Baulac, M., Clemenceau, S., Renault, B., & Varela, F. (1998). Epileptic Seizures Can Be Anticipated by Non-linear Analysis. *Nature Medicine*, 4(10), 1173–1176. <https://doi.org/10.1038/2667>
- McCaffrey, D. F., Ellner, S., Gallant, A. R., & Nychka, D. W. (1992). Estimating the Lyapunov Exponent of a Chaotic System With Nonparametric Regression. *Journal of the American Statistical Association*, 87(419), 682. <https://doi.org/10.2307/2290206>
- McSharry, P. E., Smith, L. A., & Tarassenko, L. (2003). Prediction of Epileptic Seizures: Are Nonlinear Methods Relevant? *Nature Medicine*, 9(3), 241–242. <https://doi.org/10.1038/nm0303-241>
- Meyer-Lindenberg, A. (1996). The Evolution of Complexity in Human Brain Development: An EEG Study. *Electroencephalography and Clinical Neurophysiology*, 99(5), 405–411. [https://doi.org/10.1016/s0013-4694\(96\)95699-0](https://doi.org/10.1016/s0013-4694(96)95699-0)
- Mirzaei, A., Ayatollahi, A., Gifani, P., & Salehi, L. (2010). Spectral entropy for epileptic seizures detection. *2010 2nd International Conference on Computational Intelligence, Communication Systems and Networks*, 301–307. <https://doi.org/10.1109/CICSyN.2010.84>
- Mirzaei, A., Ayatollahi, A., & Vavadi, H. (2011). Statistical Analysis of Epileptic Activities Based on Histogram and Wavelet-Spectral Entropy. *Journal of Biomedical Science and Engineering*, 04(03), 207–213. <https://doi.org/10.4236/jbise.2011.43029>

- Morrell, M. (2006). Brain Stimulation for Epilepsy: Can Scheduled or Responsive Neurrostimulation Stop Seizures? *Current Opinion in Neurology*, 19(2), 164–168. <https://doi.org/10.1097/01.wco.0000218233.60217.84>
- Müller, V., Lutzenberger, W., Pulvermüller, F., Mohr, B., & Birbaumer, N. (2001). Investigation of Brain Dynamics in Parkinson's Disease by Methods Derived From Nonlinear Dynamics. *Experimental Brain Research*, 137(1), 103–110. <https://doi.org/10.1007/s002210000638>
- Narzo, A. F. D. (2019). *Tserieschaos: Analysis of nonlinear time series* [R package version 0.1-13.1]. <https://doi.org/10.32614/CRAN.package.tseriesChaos>
- Nusse, H., & Yorke, J. (1998). *Dynamics: Numerical explorations*. Springer New York. <https://books.google.it/books?id=Bb4O8fnh8CUC>
- Nychka, D., Ellner, S., Gallant, A. R., & McCaffrey, D. (1992). Finding chaos in noisy systems. *Journal of the Royal Statistical Society. Series B. Methodological*, 54(2), 399–426. <https://doi.org/10.1111/j.2517-6161.1992.tb01889.x>
- Osborne, A., & Provenzale, A. (1989). Finite Correlation Dimension for Stochastic Systems With Power-law Spectra. *Physica D Nonlinear Phenomena*, 35(3), 357–381. [https://doi.org/10.1016/0167-2789\(89\)90075-4](https://doi.org/10.1016/0167-2789(89)90075-4)
- Osterhage, H., & Lehnertz, K. (2007). NONLINEAR TIME SERIES ANALYSIS IN EPILEPSY. *International Journal of Bifurcation and Chaos*, 17(10), 3305–3323. <https://doi.org/10.1142/s0218127407019081>
- Ott, E. (2006). Basin of attraction [revision #170495]. *Scholarpedia*, 1(8), 1701. <https://doi.org/10.4249/scholarpedia.1701>
- Palus, M. (1996). Nonlinearity in normal human eeg: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biological cybernetics*, 75, 389–96. <https://doi.org/10.1007/s004220050304>
- Park, J. Y., & Whang, Y.-J. (2012). Random walk or chaos: A formal test on the lyapunov exponent. *Journal of Econometrics*, 169(1), 61–74. <https://doi.org/10.1016/j.jeconom.2012.01.012>
- Prichard, D., & Theiler, J. (1994). Generating Surrogate Data for Time Series With Several Simultaneously Measured Variables. *Physical Review Letters*, 73(7), 951–954. <https://doi.org/10.1103/physrevlett.73.951>
- Pritchard, W., & Duke, D. (1995). Measuring chaos in the brain - a tutorial review of eeg dimension estimation. *Brain and Cognition*, 27(3), 353–397. <https://doi.org/https://doi.org/10.1006/brcg.1995.1027>
- Provenzale, A., Smith, L. A., Vio, R., & Murante, G. (1992). Distinguishing Between Low-dimensional Dynamics and Randomness in Measured Time Series. *Physica D Nonlinear Phenomena*, 58(1-4), 31–49. [https://doi.org/10.1016/0167-2789\(92\)90100-2](https://doi.org/10.1016/0167-2789(92)90100-2)
- Rapp, P. (1992). Chaos in the neurosciences: Cautionary tales from the frontier. 40.
- Rapp, P., Zimmerman, I., Albano, A., Deguzman, G., & Greenbaum, N. (1985). Dynamics of spontaneous neural activity in the simian motor cortex: The dimension of chaotic neurons. *Physics Letters A*, 110(6), 335–338. [https://doi.org/https://doi.org/10.1016/0375-9601\(85\)90786-8](https://doi.org/https://doi.org/10.1016/0375-9601(85)90786-8)
- Rogowski, Z., Gath, I., & Bental, E. (1981). On The Prediction of Epileptic Seizures. *Biological Cybernetics*, 42(1), 9–15. <https://doi.org/10.1007/bf00335153>
- Rombouts, S., Keunen, R., & Stam, C. (1995). Investigation of Nonlinear Structure in Multichannel EEG. *Physics Letters A*, 202(5-6), 352–358. [https://doi.org/10.1016/0375-9601\(95\)00335-z](https://doi.org/10.1016/0375-9601(95)00335-z)

- Rosenstein, M. T., Collins, J. J., & De Luca, C. J. (1993). A practical method for calculating largest lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1–2), 117–134. [https://doi.org/10.1016/0167-2789\(93\)90009-p](https://doi.org/10.1016/0167-2789(93)90009-p)
- Rössler, O. (1979). An Equation for Hyperchaos. *Physics Letters A*, 71(2-3), 155–157. [https://doi.org/10.1016/0375-9601\(79\)90150-6](https://doi.org/10.1016/0375-9601(79)90150-6)
- Ruelle, D. (1990). Is There Screening in Turbulence? *Journal of Statistical Physics*, 61(3-4), 865–868. <https://doi.org/10.1007/bf01027304>
- Ruelle, D., & Takens, F. (1971). On The Nature of Turbulence. *Communications in Mathematical Physics*, 20(3), 167–192. <https://doi.org/10.1007/bf01646553>
- Sandubete, J., & Escot, L. (2021). Dchaos: An r package for chaotic time series analysis. *The R Journal*, 13, 232–252. <https://doi.org/10.32614/RJ-2021-036>
- Sano, M., & Sawada, Y. (1985). Measurement of the lyapunov spectrum from a chaotic time series. *Phys. Rev. Lett.*, 55, 1082–1085. <https://doi.org/10.1103/PhysRevLett.55.1082>
- Sato, S., Sano, M., & Sawada, Y. (1987). Practical methods of measuring the generalized dimension and the largest lyapunov exponent in high dimensional chaotic systems. *Progress of Theoretical Physics*, 77, 1–5. <https://api.semanticscholar.org/CorpusID:120974262>
- Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65(3-4), 579–616. <https://doi.org/10.1007/bf01053745>
- Sauer, T. (2006). Attractor Reconstruction. *Scholarpedia*, 1(10), 1727. <https://doi.org/10.4249/scholarpedia.1727>
- Schuster, H. G. (1984, January). *Deterministic Chaos: An Introduction*. <http://ci.nii.ac.jp/ncid/BA03602841>
- Shimada, I., & Nagashima, T. (1979). A numerical approach to ergodic problem of dissipative dynamical systems. *Progress of Theoretical Physics*, 61, 1605–1616. <https://api.semanticscholar.org/CorpusID:122578707>
- Shintani, M., & Linton, O. (2004). Nonparametric Neural Network Estimation of Lyapunov Exponents and a Direct Test for Chaos. *Journal of Econometrics*, 120(1), 1–33. [https://doi.org/10.1016/s0304-4076\(03\)00205-7](https://doi.org/10.1016/s0304-4076(03)00205-7)
- Stam, C., Van Woerkom, T., & Keunen, R. (1997). Non-linear Analysis of the Electroencephalogram in Creutzfeldt-Jakob Disease. *Biological Cybernetics*, 77(4), 247–256. <https://doi.org/10.1007/s004220050385>
- Stam, C. J. (2005). Nonlinear dynamical analysis of eeg and meg: Review of an emerging field. *Clinical Neurophysiology*, 116, 2266–2301. <https://api.semanticscholar.org/CorpusID:15359405>
- Strogatz, S. H. (2018, May). *Nonlinear dynamics and chaos*. <https://doi.org/10.1201/9780429492563>
- Theiler, J. (1986). Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A*, 34, 2427–2432. <https://doi.org/10.1103/PhysRevA.34.2427>
- Theiler, J. (1995). On The Evidence for Low-dimensional Chaos in an Epileptic Electroencephalogram. *Physics Letters A*, 196(5-6), 335–341. [https://doi.org/10.1016/0375-9601\(94\)00856-k](https://doi.org/10.1016/0375-9601(94)00856-k)
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Doyne Farmer, J. (1992). Testing for nonlinearity in time series: The method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1), 77–94. [https://doi.org/https://doi.org/10.1016/0167-2789\(92\)90102-S](https://doi.org/https://doi.org/10.1016/0167-2789(92)90102-S)

- Tong, H., Chan, K. S., Cox, D. R., Cutler, C. D., Guégan, D., Jensen, J. L., Johansen, S., Lawrence, A. J., Lebaron, B., Ozaki, T., Nychka, D. W., Ellner, S., Bailey, B. A., Gallant, A. R., Smith, L. R., Smith, R. L., & Wolff, R. C. L. (1995). A personal overview of non-linear time series analysis from a chaos perspective [with discussion and rejoinder]. *Scandinavian Journal of Statistics*, 22(4), 399–445. Retrieved October 17, 2024, from <http://www.jstor.org/stable/4616372>
- Tsipouras, M. G. (2019). Spectral Information of EEG Signals With Respect to Epilepsy Classification. *EURASIP Journal on Advances in Signal Processing*, 2019(1). <https://doi.org/10.1186/s13634-019-0606-8>
- Whang, Y.-J., & Linton, O. (1999). The Asymptotic Distribution of Nonparametric Estimates of the Lyapunov Exponent for Stochastic Time Series. *Journal of Econometrics*, 91(1), 1–42. [https://doi.org/10.1016/s0304-4076\(98\)00047-5](https://doi.org/10.1016/s0304-4076(98)00047-5)
- Widman, G., Lehnertz, K., & Elger, C. (1995). Cplxmon, a system for real-time, online monitoring of neuronal complexity loss in the ecog of patients with temporal-lobe epilepsy. *Epilepsia*, 36, 106–106.
- Wolf, A., Swift, J., Swinney, H. L., & Vastano, J. (1985). Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3), 285–317. [https://doi.org/10.1016/0167-2789\(85\)90011-9](https://doi.org/10.1016/0167-2789(85)90011-9)
- World Health Organization. (2024). *Epilepsy: Fact sheet* [Accessed: 10 October 2024]. <https://www.who.int/en/news-room/fact-sheets/detail/epilepsy>
- Yao, Q., & Tong, H. (1995). On initial-condition and prediction in nonlinear stochastic systems.
- Yao, Q., & Tong, H. (1994b). On prediction and chaos in stochastic systems. *Philosophical Transactions: Physical Sciences and Engineering*, 348(1688), 357–369.
- Yao, Q., & Tong, H. (1994a). Quantifying the influence of initial values on non-linear prediction. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(4), 701–725. <http://www.jstor.org/stable/2346192>