POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

Nonlinear analysis of neurophysiological time series: methodological evaluation and clinical application in Parkinson's and Alzheimer's diseases

TESI MAGISTRALE IN BIOMEDICAL ENGINEERING

AUTHOR: ROSANNA FERRARA

ADVISOR: ANNA MARIA BIANCHI

ACADEMIC YEAR: 2020-2021

1. Introduction

The need to introduce advanced techniques for a greater understanding of the dynamic properties of complex phenomena has been approached through nonlinear analysis which, in the last twenty years, has seen a growing interest for application to neuro-electrophysiological recordings. These techniques integrated the classical linear techniques with the purpose to assess important information about brain functioning in both physiological and pathological conditions. However, despite the interesting results obtained in literature, such measures are still far from the clinical practice and this is due to the lack of proper guidelines aimed at increasing their knowledge and physiological interpretation.
The presented work aimed to evaluate the performance of a set of selected nonlinear measures, defining specific criteria for their applicability on such signals and for the appropriate input parameter selection in order to increase the repeatability of the algorithms along different datasets. The features analyzed regards the fractal scaling properties, quantified by the power-law exponent (PLE) and the detrended fluctuation analysis (DFA), as well as the regularity of the time series, estimated by entropy indices, namely the Approximate Entropy (ApEn), the Sample Entropy (SampEn) and the multiscale entropy (MSE).
The application of the algorithms and the correct parameters setting were presented on simulated time series, divided in fractional Gaussian noise (fGn) and fractional Brownian motion (fBm) for fractal indices, and white noise and 1/f noise for entropic measures. Moreover, to provide examples of application, two clinical datasets of real neurophysiologic signals are evaluated in terms of test benches for the investigated nonlinear analysis techniques

2. Selected nonlinear methods

The five methods selected in this work were chosen for their clinical and applicative relevance in the most common neurodegenerative disorders, i.e. Parkinson's (PD) and Alzheimer's diseases (AD).

- Power law exponent (PLE)

The power law exponent, also indicated as β-exponent, is defined on the power spectrum density (PSD) of the signal. By considering the inversely proportional relationship between the

PSD and the frequency, the method measures changes in the scale-free dynamic properties of the time series as the slope of the regression line defined in log-log coordinates.
Moreover, in this work, to guarantee that the estimation of the β-values was not compromised by the rhythmic oscillatory components, characterizing real neurophysiological time series, the regression line is identified following a peak removal operation. As suggested by the study of Colombo et al. the peaks of the signals are accurately suppressed before performing the linear regression operation [1].

- Detrended fluctuation analysis (DFA)

The detrended fluctuation analysis, also indicated as α-exponent, estimates the so-called long-range temporal correlation (LRTCs) present in the signal. By defining the relationship between the variability of the signal and the length of the intervals over which the variability is computed, the method measures the level of self-similarity in a time series as the slope of a regression line in log-log coordinates. This procedure is applied on the amplitude envelopes of the oscillatory activity in specific frequency ranges, extracted by bandpass filtering the time series and by applying on the filtered signals the Hilbert transformation [2].

- Entropy

Entropy represents the most classic nonlinear measure used to evaluate the amount of information hidden in the signal. It arises from the idea of quantifying uncertainty and predictability in a time series, assuming high values for signals of great complexity and lower values for more regular and predictable signals. There are different indices of entropy, but those chosen for the current analysis are the approximate entropy (ApEn), sample entropy (SampEn) and multiscale entropy (MSE) [3]–[5]. The main idea of the first two is to calculate the conditional probability that two similar sequences of m points remain similar at the point m+1. The mathematical steps underlying their calculation are similar. The main difference regards the management of the self-counting in the mathematical computations. Although SampEn and ApEn are the traditional algorithms used to quantify the complexity, in some cases they seem to fail and not to correctly quantify the degree of randomness intrinsic in the time series. This happens when the series under analysis presents structures on multiple scales, as the case of neurophysiological time series. For this reason, the need of introducing a new measure of complexity to capture the information at multiple time scales gave rise to the so-defined multiscale entropy (MSE). MSE is estimated in two steps including the construction of the coarse-grained time-scale structure and the estimation of a measure of entropy (e.g., SampEn) at each time scale.

3. Applicability criteria

The intention to identify and verify the conditions of applicability of the different nonlinear methods, described in section 2, represents the main objective of this thesis. It arises from the need to provide guidelines for the correct implementation of the measures and the appropriate selection of the parameters which highly influence the application performances. The analyses are carried out on different sets of 1000 artificial time series in order to investigate both parameters common to all the measures, such as data length, and specific parameters linked to the implementation of each single measure. For sake of clarity, methods have been grouped into two different sections according to which aspects they analyze, i.e. whether the fractal behavior or the irregularity of the neural time series.

3.1. Fractal analysis

The decision to group the power-law exponent (β) and the detrended fluctuation analysis (α) under the section fractal finds confirmation in relations which exists among them:

$$\beta_{fBm} = 1 + 2HE \qquad \alpha_{fBm} = 1 + HE \tag{3.1}$$

$$\beta_{fBm} = 2HE - 1 \qquad \alpha_{fBm} = HE \tag{3.2}$$

where fGn and fBm are respectively abbreviations of fractional gaussian noise and fractional brownian motion, mathematical models introduced by Mandelbrot and Van Ness in 1968 and used to characterize and classify long-range dependent processes [6]. Briefly, the concepts of "fractional gaussian noise" and "fractional Brownian motion" are generalization of two classical processes: fGn constitutes a stationary series of Gaussian random variables with constant variance and mean, whereas fBm is not stationary and represents the ordinary Brownian motion obtained by adding a parameter that assumes

values between 0 and 1 and is indicated by *h*, which it must not be confused with the Hurst exponent, estimator of the self-similarity property of signals [7]. The relations (3.1) and (3.2) are different according to whether the time series under analysis is fGn or fBm. This suggested the possibility to consider them as useful tool to verify the robustness of the fractal algorithms, and identify also the more suitable parameter to classify the time series as fBm or fGn.

Hence, the idea to compare the direct estimation of both β and α exponents with the values indirectly obtained by applying the theoretical relations. For sake of completeness, two comparisons were carried out: the first performed on 1000 fBm and 1000 fGn time series generated by implementing a function that receives as input the HE, while the second focused on similar artificial data but generated starting from known values of PLE (β-exponent). In both cases, as expected, tables 3.1 and 3.2 clearly showed that the measures estimated and the theoretically predicted ones were coherent.

	Theoretical results		Estimated results	
	β-value	α-value	β-value	α-value
fGn (HE=0.25)	-0.5	0.25	-0.57 +- 0.02	0.25 +- 0.03
fGn (HE=0.5)	0	0.5	0.00 +- 0.02	0.49 +- 0.06
fGn (HE=0.75)	0.5	0.75	0.53 +- 0.02	0.74 +- 0.07
fBm (HE=0.25)	1.5	1.25	1.46 +- 0.02	1.25 +- 0.1
fBm (HE=0.5)	2	1.5	1.97 +- 0.02	1.48 +- 0.1
fBm (HE=0.75)	2.5	1.75	2.39 +- 0.03	1.72 +- 0.09

Table 3.1 Measures estimated and theoretically predicted from fGn and fBm time series simulated from values of HE.

	Theoretical results		Estimated results	
	β-value	α-value	β-value	α-value
fGn (β =0)	0	0.5	0.00 +- 0.02	0.50 +- 0.06
fGn (β =0.5)	0.5	0.75	0.50 +- 0.02	0.74 +- 0.04
fGn (β=0.98)	0.98	0.99	0.98 +- 0.02	0.99 +- 0.04
fBm (β=1.02)	1.02	1.01	1.02 +- 0.02	1.01 +- 0.04
fBm (β=1.5)	1.5	1.25	1.48 +- 0.02	1.25 +- 0.04
fBm (β=2)	2	1.5	2.01 +- 0.02	1.48 +- 0.04

Table 3.2 Measures estimated and theoretically predicted from fGn and fBm time series simulated from values of β.

The congruence in the results finds confirmation in literature [8] and allows to identify the more suitable parameter able to classify the time series as fBm or fGn, i.e. the β-exponent. Indeed, it is worth noting that the fact that the two classes are characterized by the same HE, ranging between 0 and 1, makes the index less suitable for discerning them. Besides the analyses carried out to assess the correct functioning of the selected fractal measures, other simulation studies were performed aimed at evaluating the influence of the main parameters that play a key role in each nonlinear algorithm.

- Applicability criteria: PLE

The way in which the frequency resolution affects the power spectrum density (PSD) needs to be evaluated before the application of the power-law exponent. Indeed, an incorrect estimation of PSD may invalidate the calculation of the β-slope of the regression line defined in log-log coordinates.

It is worth to define that, in this work, the PSD is obtained by applying the modified Welch's periodogram.

Thus, to evaluate how β-exponents changes as function of the frequency resolution, the effect of using different windows-length was examined on two sets of 1000 realizations of time series generated from known β-values. i.e. fGn (β=0.5) and fBm (β=1.5). The study is based on the relation:

$$\Delta f = \frac{f_s}{M} \qquad (3.3)$$

that inversely links the frequency resolution (Δf) and the length of the windows expressed in terms of second-length (M) and the results showed how an increase of Δf led to a more stable evaluation of the β-exponent characterized by a reduced dispersion of the values measured across the 1000 surrogate time series.

- Applicability criteria: DFA

The data length (N) and the minimum and the maximum window sizes *n*, in which fluctuations F(n) are estimated to provide the measure of α-exponent, are supposed to influence the DFA-method.

The window sizes (n) along the x axis of the "diffusion plot", on which the α-index is obtained as the slope of the linear regression, are placed following an evenly spaced procedure on the logarithmic scale. The selection of n_{min} and n_{max} were, instead, tested on the same two sets of 1000 realizations defined for the evaluation of the applicability criteria of the power law exponent. Interestingly, the results showed how the accuracy of DFA in estimating the α-exponent appeared not influenced by the minimum window size. On the other hand, as regard the maximum window size, this need to be correctly combined with the data length, since an increase in N requires a larger number of intervals to provide the stability of the measure (thus, a lower value of n_{max}). Starting from

what has been described, changes of N values require further adjustment of the input parameter n_{max}.

- Applicability criteria: ApEn and SampEn

When the embedding dimension (m) and the tolerance (r) are properly set, reliable approximate and sample entropy estimations may be provided [3], [4]. The parameters are strictly related to the data length of the time series (N), and in order to evaluate to what extent this is true, changes of ApEn and SampEn indices as function of different data length were examined on a set of 1000 artificial time series. In this regard, results revealed that not both entropic indices presented the expected trend on Gaussian realizations. Instead of presenting a linear decrease with an increase of tolerance r, in fact, ApEn index assumes a distribution of values that increases up to a certain peak and then decreases.
In addition to this, the selection of values of r which allow to provide more accurate estimation of entropic indices are also evaluated. Such values are found in the range between 0.1 and 0.25 times the standard deviation of the data and were valid both for ApEn and for SampEn for larger values of N. These findings were considered for m=2 but did not change with different values of m, even if more unstable and unreliable entropic measures were found for embedding dimension $m \geq 3$.
As regard the multiscale entropy, the use of different temporal scales τ, supposed to be the influencing factor for the measure, has no considerable effects on the results. The tests, carried out on both white noise (β=0) and 1/f noise (β=1) time series, in fact, showed that with an increase of the scale factor the shape of MSE profiles maintains the same trend. Despite this, however, it is worth to consider that higher values of scale factor τ causes a lower number of samples in the coarse-grained sequence, thus the selection of the most suitable τ_{max} is necessary to avoid unreliable multiscale sample entropy estimations.

4. Application of nonlinear measures

The usefulness of the applicability criteria investigated for the nonlinear measures was demonstrated by two applications to real neuro-physiological signals. The two available datasets are used as a sort of test bench for the selected nonlinear measures. They have been introduced with the sole purpose to help understanding the basic implementation of the measures on real biological signals, since the reduced number of subjects involved and the limited in-depth studies in literature make difficult to evaluate the outcomes from a physiological and clinical point of view.

4.1. LFP analysis in Parkinson's disease

The first case study was designed to assess the activity of the brain in the subthalamic region of Parkinson's disease patients in conditions of pre and post levodopa administration. The aim was to identify the influence of the antiparkinsonian medication in 24 PD subjects, analyzing the dynamic nonlinear behavior of LFP (local field potential) time series. All data were acquired with a sampling rate of 2500 Hz and accurately preprocessed to remove movements and power-line artifacts, prior to the application of the nonlinear measures.
From the fractal point of view, the analysis of LFP time series showed significant differences in both the proposed measures between the conditions of pre and post levodopa administration. In particular, as shown in figure 4.1, β-exponents appeared to be lower in condition of pre levodopa in both the frequency bands (2-45) and (2-156) Hz evaluated.

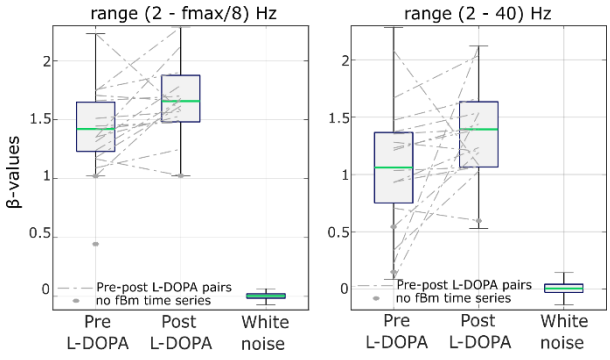


Figure 3.1 Boxplot comparison β scaling exponent quantifying power-law exponent on two frequency bands. The green line is the median value of each distribution while the dashed grey lines link the values referred to pre and post conditions respectively. Each boxplot provides the distribution of β-values estimated in condition of pre and post levodopa administration and β-values estimated on 1000 simulated white noise.

Lower β-values reflected into less steep PSD-slope were also found in the fronto-temporal sites of EEG by the recent study of Mostile et al.[9]. However, the different nature of the

electrophysiological recordings may not provide a reliable comparison. Works based on LFP time series, instead, are mostly concentrated on the demonstration of the link between the excitation-inhibition balance and the nonlinear β-value. In this regard, flatter PSD slopes were related to an increasing excitation/inhibition ratio, that means an increase in inhibition and/or a decrease in excitation [10].
On the other hand, higher α-exponents in terms of long-range temporal correlations were found after the levodopa administration at high frequency range (250-350 Hz). Interestingly, this result finds confirmation in the work of Hohlefeld et al, 2012 that showed how LRTC, being modulated by levodopa, may be proposed LRTC as a possible biomarker for PD [2].
From the entropic point of view, findings suggest that levodopa-related changes in the complexity properties of LFP are not so evident, even if an interesting trend in correspondence of high scale factors ($\tau \geq 7$) seemed to reflect increases in SampEn(2, 0.2) values in condition of pre levodopa.

4.2. EEG analysis in Alzheimer's disease

The second case study of this thesis was designed to assess the activity of the brain in patients affected by four different neurological disorders, i.e. Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), Lewy Body dementia (LBD) and Frontotemporal dementia (FTD). The aim was to verify if the use of parameters, evaluated on the dynamic behavior of EEG time series, might point out possible distinctions among the groups of patients.
The EEG recordings, performed from 19 electrodes positioned according to the International 10-20 System, were acquired at rest in both eyes-open and eyes-closed conditions. All data were recorded with sampling rates of 512 Hz in some cases and 2048 Hz in others and were accurately pre-processed to select 60-seconds stationary epochs, prior to the application of the nonlinear measures. The proposed parameters are calculated for all electrodes and for nine different subcategories of electrodes defined as follow: left frontal (Fp1, F7, F3), right frontal (Fp2, F4, F8), left central (T3, C3), right central (T4, C4), left temporal (T5, P3, O1), right temporal (T6, P4, O2), left fronto-temporal (Fp1, F7, F3, T3), right fronto-temporal (Fp2, F4, F8, T4), z-axis (Fz, Cz, Pz).
From the fractal point of view, the analysis of EEG time series showed lower β-values for MCI in eyes-closed condition and higher β-values for LBD patients in eyes-open conditions, as regard the power-law exponent. This is valid for all the different subcategories of electrodes.
As regard the DFA algorithm, instead, in terms of LRTC, results revealed higher α-exponents for LBD group of patients in eyes-open conditions. This is valid for all subcategories in case of both (8-12 Hz) and (30-40 Hz) frequency bands.
From the entropic point of view, all MSE profiles were characterized by trends that increased up to the maximum SampEn(2, 0.2) value. In line with literature, the maximum value is usually reached on smaller time scales [11], [12], and in the present study such value was found both in eyes-close and eyes-open conditions for all the subcategories of electrodes. Regarding the behavior of the signals at larger scale factors ($\tau \geq 8$), instead, the results suggested that in eyes-closed condition, LBD patients revealed values of SampEn(2, 0.2) approximately constant with increasing τ with respect to the other categories. MCI, AD and FTD, in fact, presented decreasing values of entropy with increasing τ. In eyes-open condition, all the diseases showed decreasing profiles at higher scale factors, even if less pronounced for LBD and steeper values for MCI against the other two categories.

5. Conclusions

The aim of this thesis was to investigate the conditions of applicability of selected nonlinear measures, providing guidelines for the correct setting of the parameters which influence their performance when applied to neuro-physiological signals. There is a clear need to analyze the effects and interactions of changing input parameters, since both fractal and entropic methods appear very sensitive to them [13]–[15]. Power-law exponent, detrended fluctuation analysis and multiscale entropy are the nonlinear techniques widely applied to quantify the information process capacity of the brain in Alzheimer's and Parkinson's disorders. However, even if interesting results were obtained in the literature, there is the suspect that the selection of input parameters for the algorithms was made

empirically without a deep investigation of the applicability criteria in each specific case.

The present thesis evaluated some of methodological issues on the basis of theoretical simulations and demonstrated how concrete advices about the correct choice of nonlinear measures and relative parameters may help to avoid misleading estimations, minimizing errors made by the utilizers. Moreover, these measures were successfully applied on the two real clinical datasets available in this study, even if, in both cases, an increase of the sample sizes may provide more reliable results.

In conclusion, the set of nonlinear tools for studying the neurophysiological signals and applying, if possible, the measures alongside the classical linear ones, could also facilitate clinical use. The idea of developing a graphical interface, which makes the tool easy to use for clinicians, may represent, in fact, a great challenge for research, helping to better answer the still open questions on the clinical interpretation of neural data.

Bibliography

[1] M. A. Colombo *et al.*, "The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine," *NeuroImage*, vol. 189, pp. 631–644, Apr. 2019, doi: 10.1016/j.neuroimage.2019.01.024.

[2] F. U. Hohlefeld *et al.*, "Long-range temporal correlations in the subthalamic nucleus of patients with Parkinson's disease," *European Journal of Neuroscience*, vol. 36, no. 6, pp. 2812–2821, Sep. 2012, doi: 10.1111/j.1460-9568.2012.08198.x.

[3] S. M. Pincus, A. L. Goldberger, and A. L. Goldberger Physiologi-, "Physiological time-series analysis: what does regularity quantify?"

[4] J. S. Richman, J. Randall Moorman, J. Randall, and M. Physi, "Downloaded from journals.physiology.org/journal/ajpheart (131.175.028.198) on," 2000.

[5] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Physical Review Letters*, vol. 89, no. 6, 2002, doi: 10.1103/PhysRevLett.89.068102.

[6] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, Oct. 1968, doi: 10.1137/1010093.

[7] F. J. Molz, H. H. Liu, and J. Szulga, "Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions," *Water Resources Research*, vol. 33, no. 10, pp. 2273–2286, 1997, doi: 10.1029/97WR01982.

[8] A. Eke *et al.*, "Physiological time series: distinguishing fractal noises from motions," *Pflügers Archiv - European Journal of Physiology*, vol. 439, no. 4, pp. 403–415, Feb. 2000, doi: 10.1007/s004249900135.

[9] G. Mostile *et al.*, "Complexity of electrocortical activity as potential biomarker in untreated Parkinson's disease," *Journal of Neural Transmission*, vol. 126, no. 2, pp. 167–172, Feb. 2019, doi: 10.1007/s00702-018-1961-6.

[10] R. Gao, E. J. Peterson, and B. Voytek, "Inferring synaptic excitation/inhibition balance from field potentials," *NeuroImage*, vol. 158, pp. 70–78, Sep. 2017, doi: 10.1016/j.neuroimage.2017.06.078.

[11] T. Mizuno *et al.*, "Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy," *Clinical Neurophysiology*, vol. 121, no. 9, pp. 1438–1446, Sep. 2010, doi: 10.1016/j.clinph.2010.03.025.

[12] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, Nov. 2006, doi: 10.1088/0967-3334/27/11/004.

[13] Z. M. H. Almurad and D. Delignières, "Evenly spacing in Detrended Fluctuation Analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 451, pp. 63–69, Jun. 2016, doi: 10.1016/j.physa.2015.12.155.

[14] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao, "A tutorial introduction to adaptive fractal analysis," *Frontiers in Physiology*, vol. 3 SEP, 2012, doi: 10.3389/fphys.2012.00371.

[15] X. Chen, I. C. Solomon, and K. H. Chon, "Comparison of the use of approximate entropy and sample entropy: Applications to neural respiratory signal," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2005, vol. 7 VOLS, pp. 4212–4215. doi: 10.1109/iembs.2005.1615393.

POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

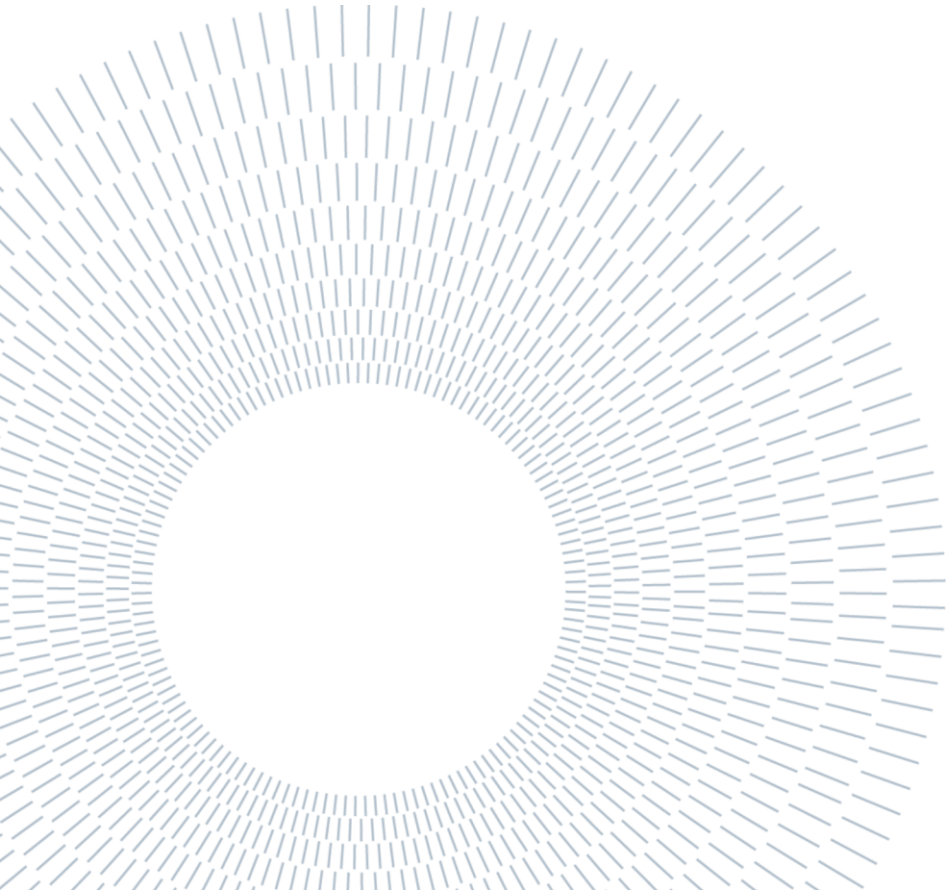
Nonlinear analysis of neurophysiological time series: methodological evaluation and clinical application in Parkinson's and Alzheimer's diseases

TESI DI LAUREA MAGISTRALE IN
BIOMEDICAL ENGINEERING

Author: **Rosanna Ferrara**

Student ID:	928363
Advisor	Prof. Anna Maria Bianchi
Co-advisors:	Prof Alberto Priori PhD Alberto Averna PhD Stefania Coelli
Academic Year:	2020-2021

To my dear mother, for all the love,
understanding and unconditional
support she has been able to give me.
The only one who has always been by my side.
I love you mom, wherever you are.



Abstract

Biomedical signal processing is fundamental in investigating and quantifying the information captured by neuro-electrophysiological recordings, such as local field potential (LFP), electrocorticogram (ECoG), electroencephalography (EEG) and magnetoencephalography (MEG).

These biosignals have been traditionally studied by linear approaches such as descriptive statistics, and the analysis of the frequency content in terms of power spectral density (PSD). However, in the last few decades, concepts and techniques from nonlinear dynamics integrated the classical linear techniques to assess important information about brain functioning in both physiological and pathological conditions.

Moreover, despite the potentiality of nonlinear approach in understanding the complex pathophysiology of neurodegenerative diseases has been shown by several scientific works, the lack of concrete advices about the correct implementation and the appropriate selection of the input parameters which influence the measures, make the applications still far from the clinical practice.

The purpose of this thesis is to evaluate the performance of a set of selected nonlinear processing techniques, defining specific criteria for the appropriate parameter selection in order to increase the repeatability of the algorithms along different signals. In this regard, the fractal scaling properties of the time series is quantified by the power-law exponent (β-exponent) and the detrended fluctuation

analysis (DFA), while predictability and regularity are estimated by different entropy indices, namely the Approximate Entropy (ApEn), the Sample Entropy (SampEn) and the multiscale entropy (MSE).

Insights about the application of the algorithms and the correct parameters setting were presented on simulated time series, but in order to provide examples of the implementation of the measures on real neurophysiologic signals, also two clinical datasets, proposed as test benches, are evaluated. The first one, related to local field potentials recorded from Parkinsonian patients, was analyzed in order to evaluate which is the measure that better enounce the possible effects of the levodopa on patients. The second one, focused on electroencephalograms acquired from different groups of patients, was examined with the purpose to verify if the use of parameters may point out possible distinctions among four neurological diseases, i.e. Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), Lewy Body dementia (LBD) and Frontotemporal dementia (FTD).

The results of this study highlighted the importance of guidelines for a suitable implementation of nonlinear measures on neurophysiological signals. This might provide effective support to the correct choice of nonlinear measures, minimizing the errors made due to an improper selection and enabling future researches to better answer the still open questions in terms of clinical interpretation in neural data.

Key-words: power law exponent, detrended fluctuation analysis, long-range temporal correlation, approximate entropy, sample entropy, multiscale entropy, applicability criteria, local field potential, electroencephalography, Parkinson's disease, Alzheimer's disease.

Abstract in lingua italiana

L'elaborazione dei segnali biomedici è fondamentale per studiare e quantificare le informazioni acquisite da segnali elettrofisiologici, quali il potenziale di campo locale (LFP), l'elettrocorticografia (ECoG), l'elettroencefalografia (EEG) e la magnetoencefalografia (MEG).

Questi segnali sono stati tradizionalmente studiati attraverso approcci lineali come le statistiche descrittive e l'analisi del contenuto in frequenza in termini di densità spettrale di potenza (PSD). Tuttavia, negli ultimi decenni, le tecniche classiche sono state integrate dai concetti e dalle tecniche della dinamica non lineare al fine di valutare importanti informazioni sul funzionamento del cervello in condizioni sia fisiologiche che patologiche.

Inoltre, nonostante la potenzialità dell'approccio non lineare nella comprensione della complessa fisiopatologia dei disturbi neurodegenerativi sia stata dimostrata da diversi lavori scientifici, la mancanza di consigli concreti, in merito alla corretta implementazione e all'appropriata selezione dei parametri di ingresso che influenzano la misura, rende le applicazioni ancora lontane dalla clinica pratica.

Lo scopo di questa tesi è quello di valutare le prestazioni di un insieme di tecniche di elaborazione non lineare selezionate, definendo specifici criteri per un'appropriata selezione dei parametri in modo da aumentare la ripetibilità degli algoritmi su diversi segnali. A questo proposito, la proprietà di ridimensionamento frattale della

serie temporale viene quantificata dal power-law exponent (esponente β) e dalla detrended fluctuation analysis (DFA), mentre la predicibilità e la regolarità vengono stimate attraverso diversi indici entropici, definiti Approximate entropy (ApEn), Sample entropy (SampEn) e Multiscale entropy (MSE).

Gli approfondimenti sull'applicazione degli algoritmi e sulla corretta impostazione dei parametri sono stati presentati su serie temporali simulate, ma anche due dataset clinici, proposti come banco di prova, sono stati valutati al fine di fornire esempi di attuazione delle misure su segnali neurofisiologici reali. Il primo esempio, basato sui segnali LFP acquisiti da pazienti affetti da Parkinson, è stato analizzato in modo da valutare quale sia la misura che meglio evidenzia i possibili effetti della levodopa sui pazienti. Il secondo, centrato su segnali elettroencefalografici acquisiti da diversi gruppi di pazienti, è stato esaminato con l'obiettivo di verificare se l'utilizzo dei parametri possa sottolineare possibili distinzioni tra i quattro gruppi di patologie neurologiche, che sono Alzheimer, deficit cognitivo lieve, demenza frontotemporale, malattia a corpi di Lewi.

I risultati di questo studio hanno sottolineato l'importanza di linee guida per un'implementazione più adatta delle misure non lineari su segnali neurofisiologici. Questo potrebbe fornire un supporto efficace alla scelta corretta delle misure non lineari, minimizzando gli errori commessi a causa di una selezione impropria e consentendo alle ricerche future di rispondere meglio alle domande ancora aperte in termini di interpretazione clinica nei dati neurali.

Parole chiave: power law exponent, detrended fluctuation analysis, long-range temporal correlation, approximate entropy, sample entropy, multiscale entropy, criteri di applicabilità, potenziale di campo locale, elettroencefalografia, malattia di Parkinson, malattia di Alzheimer.

Contents

1. Introduction

Biomedical signal processing is the branch of knowledge that deals with the methods that allow to manage physiological data continuously collected from patients. The great amount of data collected is processed through algorithms that correlate mathematical results to clinical aspects in order to detect useful information about the health state of the patient and/or the natural progression of a particular disease.

In this regard, the simplest and most direct experimental approach implemented is the linear analysis, mainly focused on the estimation of the power spectrum and applied both in time and frequency domain. The approach results quite easy to be interpreted in physiological terms and it is largely implemented in clinical practice still now.

However, the relevant information hidden within biomedical signals is often masked or not well defined, appearing corrupted by interference attributable to the complexity of the underlying generation systems. For this reason, in order to define the notion of complexity and quantify it, researches in the biomedical field have focused their attention on how each measure of complexity is related to particular aspects of the systems, such as regularity or irregularity, randomness, predictability, self -similarity, long-range correlation, etc. [1].

In this perspective, the need to introduce advanced techniques for a greater understanding of the dynamic properties of complex phenomena has been highlighted by the nonlinear analysis which, in the last twenty years, has seen a growing interest for application to electrophysiological recordings.

The use of these measure appears as a great challenge for biomedical engineers, even if further investigations are still required, since until now their criteria of application are not clear. Hence, the need to define a sort of guidelines to increase the knowledge and the physiological interpretation of such measures has emerged, with the purpose of make clinicians more inclined to use them.

Starting from these assumptions, the present thesis aims to evaluate the performances and the conditions of applicability of nonlinear measures of i) irregularity, such as the entropy, and measures of ii) fractal dimension, such as the power-law exponent and the detrended fluctuation analysis. The methodological analysis presented was developed in Matlab (version 9.1 R2021a), and based on selected studies in the literature which are outlined in detail in section 1.3.

i) Entropy represents the most classic nonlinear measure used to evaluate the amount of information hidden in the signal. It arises from the idea of quantifying uncertainty and predictability in a time series, assuming high values for signals of great complexity and lower values for more regular and predictable signals. There are different indices of entropy, from the historical one given by Shannon and Kolmogorov-Sinai to those chosen for the current analysis, i.e. the approximate entropy (ApEn), sample entropy (SampEn) and multiscale entropy (MSE). These measures will be described in sections 2.2 and 2.3.

ii) The fractal analysis aims at evaluating the scaling intrinsic properties of electrophysiological signals. One of the best-known fractal measures is the detrended

fluctuation analysis (DFA) that estimates the so-called long-range temporal correlation (LRTCs) present in the signal. By defining the relationship between the variability of the signal and the length of the intervals over which the variability is computed, the method measures the level of self-similarity in a time series as the slope of a regression line in log-log coordinates. Another important fractal measure is the power law exponent (PLE) that is defined on the power spectrum density (PSD) of the signal. By considering the inversely proportional relationship between the PSD and the frequency, the method measures changes in the scale-free dynamic properties of the time series as the slope of the regression line defined in log-log coordinates. These measures will be described in section 2.1.

In the present work, such nonlinear measures have been studied and employed with the pursue to investigate abnormal changes in two examples of the most common neurodegenerative disorders, i.e. Parkinson's and Alzheimer's diseases. The analysis was carried out first on artificial data and then on neural signals recorded both outside and inside the brain using electroencephalography (EEG) and local field potential (LFP) respectively.

In this regard, section 1.1 describes the neural signals used in this thesis and the corresponding recording techniques; section 1.2 outlines the two pathologies taken into account to evaluate the possible usefulness of the nonlinear approach, i.e. Alzheimer's and Parkinson's diseases; section 1.3 provides a rapid excursus of the studies found in literature about the application of the above-mentioned measures on EEG and LFP, centering the attention on the brain disorders under analysis; section 1.4 shows how the work is organized, summarizing the main aspects of the chapters that make up the thesis.

1.1 Neuro-electrophysiological recording techniques

Neurons are the smallest unit that make up the nervous system and represent its functional activity. In normal condition, the human brain consist of 10^{10} units considered electrically excitable cells organized in a neural network and able to communicate by sending and receiving electrochemical signals.

The flux of information exchanged among neurons is caused by the so-called action potential, a signal generated by the temporary shift in membrane potential and that propagates inside the network. The transmission is mediated by neurotransmitters, like dopamine, that is a chemical substance able to change the permeability of cell membrane favoring the passage of ions and producing variations of potential [2]. But, how this flux of information can be detected? Why it's so important to record the neural activity?

In this regard, it is worth specifying that during the last decade, advanced systems of acquisitions of neural signals have been introduced with the purpose of recording neuronal activity, such as local field potential (LFP), electrocorticogram (ECoG), electroencephalography (EEG) and magnetoencephalography (MEG).

By processing streams of biosignals extracted from patients or healthy subjects through mathematical algorithms, it is possible to understand how the human brain works or how clinical information related to the behavior of the neural system can be detected. These algorithms, which can be distinguished between linear and nonlinear, have been largely discussed in literature in terms of applications on biomedical signals. However, the lack of definition of specific criteria for parameter selection makes difficult the possibility to compare different studies or to increase the measure repeatability, especially for the nonlinear dynamic approach.

For this reason, the present thesis aimed to investigate the effects of changing parameters of nonlinear techniques with the final goal to suggest rules to properly choose them in clinical application. To this end, this study is focused on the neurological EEG or LFP signals, which are distinguished by different spatial resolution and degree of invasiveness of the procedure [3]. Figure 1.1 (a) shows an example of thirty seconds of LFP recorded at 2500 Hz, while figure 1.1 (b) exhibits the EEG data recorded from the sensor C4 at 512 Hz.

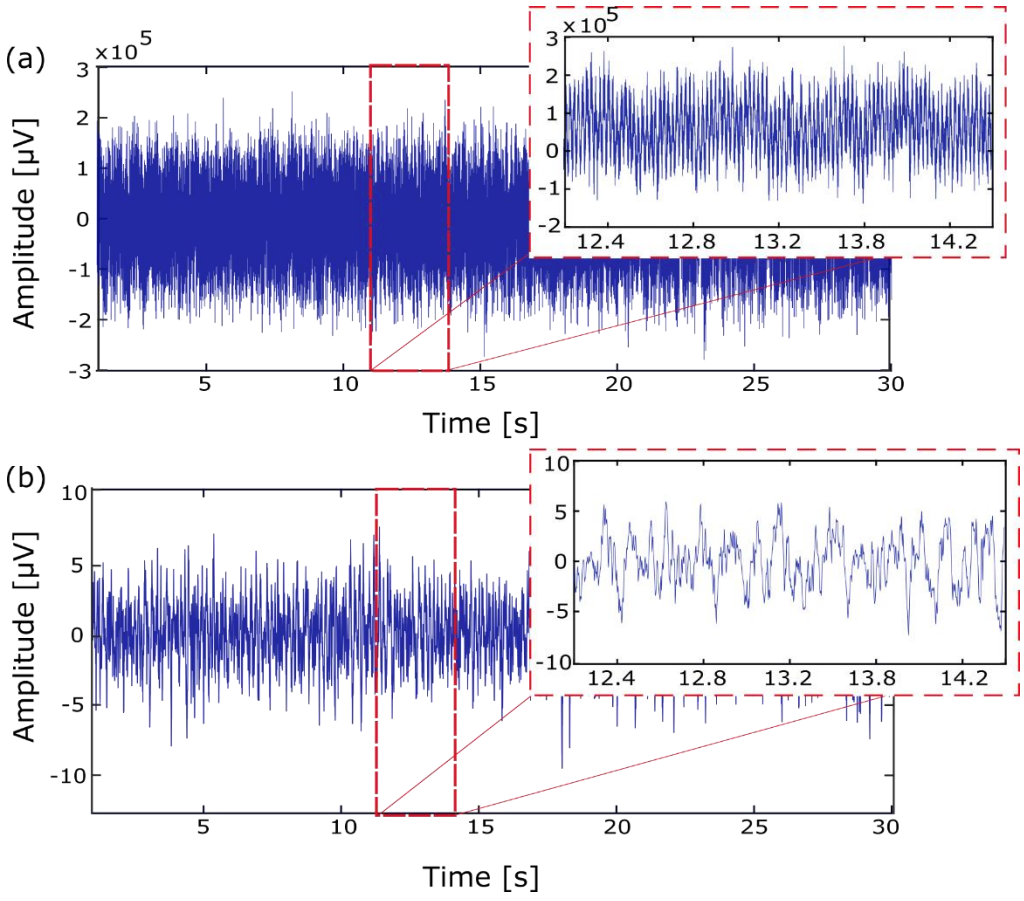


Figure 1.1 (a) Example of thirty seconds of local field potential (LFP) data recorded at 2500 Hz from a representative subject. (b) Example of thirty seconds EEG data recorded from the sensor C4 at 512 Hz.

1.1.1 Local field potential (LFP)

The local field potential represents the electric potential measured through an invasive procedure based on the insertion of macroelectrodes inside the basal ganglia of the patient's brain. It derives from the extracellular signal picked up by the implanted electrode and reflects the behaviors of multiple neurons summed up in a unique recording. This recording can be high-pass filtered (typically > 1000 Hz) originating the time series termed multi-unit-activity (MUA) or low-pass filtered (typically < 500 Hz) giving rise to the time series defined local field potential (LFP).

The main difference between the two time series is the neuronal processes that they reflect in relation to the distance from the recording electrode. Indeed, the first represents the action potentials, identified as spikes and used by neurons placed in proximity to the recording electrode to communicate within the neural network [4]; the second one, instead, reproduces the output signal of more distant neuronal cells reflecting the single action potential in addition to the membrane currents and other charges in brain cells [5]. The latter category of signals, moreover, can be correctly outlined only if the impedance of the recording electrodes is low enough; if not, in fact, the high impedance would act as high pass filter preventing its correct detection [6].

The recording technique for LFPs acquisition takes place during the deep brain stimulation (DBS), a surgical procedure that involves the implantation of electrodes in a specific area of the brain and the application of a high frequency electrical stimulation. This makes the comparison between pathologic and healthy brain impossible in human, despite the high quality and significant spatial and frequency resolution that characterize the signal carried out.

1.1.2 Electroencephalography (EEG)

Electroencephalography (EEG) is the most traditional and ancient technique used to measure the electrical activity produced by pyramidal neurons within the cerebral cortex. Introduced by Hans Berger, German professor of psychology, in 1929 it is still widely used for clinical purpose due to its non-invasive nature and excellent temporal resolution, which allows brain activity to be recorded in real time.

The main advantage of the EEG acquisition is the simplicity of the recording procedure carried out through electrodes organized in a cap that can be easily placed on the human scalp. Each single electrode, accurately positioned and fixed through conductive gel, measures the extracellular currents generated by synchronized neurons and that propagate through the different cortical areas, revealing the mechanisms of information processing implemented in the brain. The positioning of sensors on the subject head is typically implemented in accordance with the suggestion of the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFCN). The adopted method, defined "10-20 system", by associating letters to the brain regions and numbers to the two hemispheres, allows to standardize the electrodes positioning making the recording technique less susceptible to measurement errors [7]

Nevertheless, despite being widely applied to both pathological and healthy human brain, this non-invasive procedure has a particular limitation related to the signal to noise ratio (SNR) which results certainly lower compared to the invasive technique (section 1.1.1). Recording neural activity from the external surface of the brain, in fact, ensures that each electrode captures information not only from neurons directly below it, but also from the surrounding sensors increasing the contribution of error in the measurement.

1.2 Neurodegenerative disorders

Neurodegenerative disorders include all diseases caused by the progressive death of neuronal cells inside the brain. A large number of studies in literature have long described the possibility to make diagnosis through the visual inspection of neurological signals, for example by investigating possible alterations in patient's neural activity with respect to healthy conditions. In this regard, a quantitative analysis, based on the implementation of nonlinear algorithms, may lead to understand specific features related to neurodegenerative disorders adding useful information about both the treatment and the prediction of patient's outcomes.

The present study investigates two of the most common brain disorders worldwide, i.e. the Parkinson's and the Alzheimer's disease, described below.

1.2.1 Parkinson's disease (PD)

Parkinson's Disease (PD), affecting 1-2% of the population over 65 years [8], is related to the loss of dopamine-producing brain cells. Motor impairments (i.e. akinesia, rigidity and tremor) and cognitive decline are the main symptoms that affect PD patients' independence and quality of life [9]

The hallmark sign of the disease is the progressive drop in dopamine production caused by the death of the neurons within the substantia nigra of the basal ganglia. Indeed, in normal condition, the substantia nigra has a high number of dopaminergic neurons, but by the end stages of PD, patients have often lost more than half of those neurons in this region.

The most common treatment for PD involves restoring depleted dopamine levels in the basal ganglia through levodopa, a precursor of dopamine. The reason of the use of this type of treatment is related to the fact that levodopa, unlike dopamine, can cross the blood-brain barrier allowing the brain to synthetize more dopamine.

However, despite this therapy doesn't stop the neurodegeneration that affect PD, it has proved to be very successful in the earlier stages of the disease improving the patients' motor symptoms, but not in the later stages. Indeed, the long-term use of it seems to develop a drug resistance diminishing the treatment's effects and to increase a series of side effects, such as changes from involuntary movements to inability to move [10].

In these later stages, the mobility of the patients is instead treated by deep brain stimulation (DBS), a surgical procedure that, by implanting electrodes in a specific area of the brain and applying high frequency electrical stimulation, appears to manage the motor symptoms which characterize the disease under analysis. In addition to its therapeutic function, the DBS also offers the opportunity to record the electrical activity from the patient's basal ganglia, i.e. LFP signals [11]

1.2.2 Alzheimer's disease (AD)

Alzheimer's disease (AD), affecting more than 47.5 million people worldwide and predicted to exponentially increase in the coming ten years, is caused by the gradually destruction of neuronal cells. Anxiety, depression, apathy, and abnormal behavior, accompanied by persistent memory loss are the main symptoms that affect AD patients[12]

The hallmark signs of the disease are the accumulation of amyloid plaques between neurons and the neurofibrillary tangles formation within the cerebral cortex. These structural changes in the brain lead to neuronal death, neurotransmitter deficits and, as a consequence, the progressive decline in cognitive functions [13].

Unfortunately, there is no possibility to prevent or to stop the development of the disease, but on the other hand, important progresses were made in the development of new drugs aimed to improve the patient's quality of life. In this regard, the main treatment is focused on relieve cognitive impairments and try to slow down the progress of the disease. It involves restoring the cholinergic functions through inhibitors, such as donepezil and galantamine and has proved to be very successful in delaying the decline in cognitive functions.

However, despite these benefits are applicable to mild, moderate, and severe patients, the AD is a chronic disease, thus a regular monitoring of both side and long-term effects of the therapy is necessary to avoid a reduction of treatment performance [14].

1.3 Application of nonlinear analysis in literature

Understanding a complex system such as the human brain has prompted researches to propose different techniques to carry out an in-depth and detailed study of neuronal activities. So far, much of what is known has been provided by the classical linear analysis, usually based on spectral analysis and widely applied to assess the frequency content of oscillatory fluctuations that arise from electrophysiological signals.

Over the years, in fact, works related to the application of such measures have been widely published, but all studies have been limited to observe and interpret the rhythmic oscillatory activity of the brain without taking into account the dynamic properties and the apparent randomness that reflect the brain's functional organization. Hence, the need to introduce nonlinear techniques able to deepening interpret the complex interactions embedded into neural recordings, despite the presence of the nonlinear approach in the clinical application is still far in the future.

Starting from these considerations, the aim of this section is to provide a brief overview of the applications of nonlinear algorithms used in context of LFP signals recorded in conditions of Parkinson's disease and EEG signals related to the conditions of Alzheimer's diseases. For sake of clarity, the reviewed works are divided into fractal and entropic measures.

1.3.1 Application of fractal measures

As regard the power law exponent, usually indicated as β exponent, the work of Vysata et al. revealed that, in resting conditions, an overall EEG decrease of power-law exponent were associated to AD patients and greater variance of the measured parameter were found in patients with respects to healthy subjects. The results were supposed by authors to be related to a reduction of complexity and to the presence of neuroanatomical connectivity changes caused by the brain atrophy, respectively [15].

In case of PD, the recent work of Belova et al., based on the study of LFP signals of patients who had discontinued levodopa treatment, found a lower β-values during movement with respect to the state of rest. This has been considered by authors

reflecting changes in motor processing suggesting a possible correlation between the β exponent measure and the excitation-inhibition ratio [16]. This correlation was already proposed in 2017 by Gao et al, the first to suppose the link between the two, using the parameter to investigate many neural disorders. Since the acquisition of LFP required invasive procedure, they carried out the study on computational models, then validating the results on data recorded from rats. As expected, both artificial and real measures revealed a steeper slope when the inhibitory component prevailed and flatter slope in the opposite case [17]. Higher β-values were also identified in the study conducted by Huang et al. a few years later. Based on the same suppositions, authors demonstrated the link between the excitation-inhibition equilibrium and the nonlinear parameter by considering the slope on LFP signals of PD patients after propofol injection, a drug that induces anesthesia. Surprisingly, result does not change compared to the previously mentioned study; they found an increased β-value after anesthetic induction compared to the awake state [18].

On the other hand, about the detrended fluctuation analysis, a relatively small number of studies have examined the evaluation of the long-range temporal correlations in the selected neurodegenerative disorders. One of these is the work of Stam et al, published in 2005 and based on the study of EEG signals recorded from AD patients. By implementing the DFA algorithm in different frequency bands, authors found higher values of the parameter in the lower alpha (8-10 Hz) and beta bands (13-30 Hz) with respect to the control subjects, suggesting a decrease of spontaneous fluctuations caused by the disease [19]. Instead, another study related to LFP signals, proved the presence of long-range temporal correlations in PD patients before and after levodopa administration. In details, the analysis performed by Hohlefeld et al. shown lower α-values in conditions of pre levodopa medications

in both β (13-35 Hz) and high-frequency (>200 Hz) bands, suggesting more irregular fluctuations after the attenuation of the symptoms of the disease [20].

1.3.2 Application of measures of Entropy

The main application of entropy is to analyze the regularity of the neurophysiological time series. In this regard, the work of Abasolo et al. aimed to analyze changes of the ApEn (m=1, r=0.2) index on EEG signals. The study revealed lower values in AD patients compared to healthy controls, suggesting a reduction in complexity in part of the brains of subjects affected by disease. The authors considered the outcomes caused by the neuronal death and the consequent loss of connectivity of local neural networks, explaining in this way the presence of the deficit in the information processing of the cortex [13]. Similar results were also detected performing entropic measures at different time scales. For example, the analyses carried out by Mizuno et al and Escudero et al. on EEG time series confirmed lower values of MSE entropy in AD patients than control subjects. The findings are related to lower time scales (1-5), indicating the neural activity less complex in patients affected by disease [21][22].

As regard the applications in the context of PD, no many studies used the entropic measures to analyze LFP. This can be related to the fact that no studies can be performed to compare for example healthy and parkinsonian subjects, since the recording procedure, implemented to extract LFP signals, is highly invasive. One of the few works present in literature, in fact, is that of Syrkin-Nikolau et al. that applied the SampEn (m=2, r=0.2) to band-pass filtered LFP signals recorded during DBS implantation. Authors estimated higher entropy in frequency band β (13-30 Hz)

during walking with freezing of gait with respect to walking without freezing of gait, suggesting a less efficient transfer of motor information during the manifestation of the phenomena [23].

1.4 Organization of the thesis

The principal purpose of the current work is to provide insight about the correct implementation and application of nonlinear analysis on electrophysiological recordings. How to manage data and which measure appears more suitable to be used in specific case of study, appears still controversial, despite the growing interest of researches in the use of this approach to detect clinical information related to both physiological and pathological conditions.

The work is organized into two main parts, the first one including all the theoretical and mathematical aspects underlying the five nonlinear algorithms selected in this study and the second one concerning two examples of practical application.

From this perspective, chapter 2 describes the mathematical implementation of the different nonlinear measures distinguishing between fractal and entropic measures. The distinction allows to explore the applicability criteria, outlined in chapter 3, referred to the proper setting of the parameters involved in the algorithms. The analysis is carried out on a large number of artificial time series defined ad hoc to test the outcomes of both fractal and entropic algorithms.

Next, chapter 4 and 5 focused on the two cases of study aimed to apply selected nonlinear methods on real neurophysiological time series. In this regard, the first application aims to quantify the nonlinear parameters on LFP signals before and after

the levodopa administration, in order to evaluate which is the analysis that better enounce the possible effects of the drug on PD patients. The second investigates fractal and entropic features on EEG signals of AD subjects compared to patients affected by other brain disorders, i.e. mild cognitive impairment, lewy body dementia and frontotemporal dementia, with the purpose of verifying if the use of parameters may point out possible distinctions among the groups of subjects. For both studies, chapter 4 presents the study population, data acquisition and pre-processing phase, while chapter 5 summarizes the detailed results also from a statistical point of view, through specific non-parametric analyses.

Finally, chapters 6 discusses the results presented in the previous chapter, focusing the attention on the comparison between the obtained outcomes with those found in literature. Chapter 7 proposed possible directions for future applications.

2. Nonlinear methods for time series analysis

The goal of this chapter is to outline the technical and mathematical details underlying the nonlinear algorithms selected for the current analysis. The five methods identified in this work are the power-law exponent, the analysis of detrended fluctuations, the approximate entropy, the sample entropy and the multiscale entropy. These methods were chosen for their clinical and applicative relevance. In this section, each of them is described in detail.

2.1 Fractal analysis

The concept of a fractal is usually associated with irregular geometric objects whose dimension is repeated on temporal and spatial scales. This means that the object exhibits self-similarity in the sense that each individual part of it represents a specific subunit that resembles the structure [24].

In this regard, fractal analysis represents a step forward in the study of biological systems. The nervous system and its activity over time depend on the interaction between numerous interconnected neurons that generate highly variable electro-

physiological signals. Organizations of this type are characterized by structural and functional redundancies that recall the concept of self-similarity of fractal processes [25].

Measures based on fractal theory relate complexity with long-range correlation and scaling behavior of the system [26], which explains why the power-law exponent (β) and the detrended fluctuation analysis (α) algorithms are grouped under the section fractal analysis.

To confirm this, since the indexes estimated through the algorithms measure the same feature, specific relationships exist among them that allow to derive one from the other [27]. However, these relationships also include another non-linear measure, the Hurst exponent (HE), introduced in 1951 by Harold Edwin Hurst, and currently used to estimate the presence or absence of long-range dependence or long-term memory, which is commonly referred to the ability of a time series to predict future events based on historical data [28]. For sake of clarity, it is important to underline that the Hurst exponent was not examined in this thesis due to the limited possibility of comparing its performance in terms of clinical information caused by the lack of studies in the literature linked to Alzheimer's and Parkinson's diseases.

Thus, summarizing, the relations that link α, β and HE are [27][29]:

$$\beta_{fBm} = 1 + 2HE \qquad \alpha_{fBm} = 1 + HE \tag{2.1}$$

$$\beta_{fBm} = 2HE - 1 \qquad \alpha_{fBm} = HE \tag{2.2}$$

where fGn and fBm are respectively abbreviations of fractional gaussian noise and fractional brownian motion, mathematical models introduced by Mandelbrot and Van Ness in 1968 and used to characterize and classify long-range dependent processes [30]

Without going into too much theoretical aspects, what is important to emphasize is that the concepts of "fractional gaussian noise" and "fractional brownian motion" are generalization of two classical processes. Indeed, fGn constitutes a stationary series of Gaussian random variables with constant variance and mean, whereas fBm is not stationary and represents the ordinary Brownian motion obtained by adding a parameter that assumes values between 0 and 1 and is indicated by *h*, which must not be confused with the HE, estimator of the self-similarity property of signals [31].

What is reported, highlights that the proper identification of these two processes is a crucial step in the application of the fractal algorithms. This aspect is accurately investigated in section 3.1, since the preliminary classification of the time series as fGn or fBm largely influences the performances of the nonlinear measures. In other words, the categorization of the long memory-process signals as stationary (fGn) or non-stationary (fBm) avoids the incorrect evaluation of the results, since the measuring range of the DFA exponent (α) and power-law exponent (β) are respectively [0:1] and [-1:1] for fGn and [1:2] and [1:3] for fBm.

2.1.1 Power-law exponent

The power law exponent, indicated by β-exponent, is a nonlinear measure defined starting from the power spectrum and used to quantify the self-affinity characteristics of the signals under analysis. This parameter, in other words, instead of considering how the time series varies in the time domain, carries out the study in the frequency domain, starting from the so-called power spectrum density (PSD) [32].

The PSD is widely used as a linear approach to evaluate the frequency content of a time series, quantifying the total power stored in different frequency bands. On the other hand, the non-oscillatory content of the signal results in a PSD inversely proportional to its frequency.

From the mathematical point of view, this means that the signals exhibit a 1/f non-oscillatory [33] component usually neglected in the application of linear methods, since the linear approach evaluates the PSD in linear coordinated enouncing only the oscillatory components of the specific signal. By considering, on the contrary, the PSD in log-log coordinates, the decay related to the $1/f^{\beta}$ component can be extracted through an operation of linear regression, whose slope corresponds to the scaling exponent β, then used as possible neurological biomarkers of disease status [1].

The reduction of complexity in Alzheimer's disease or the balance excitation/inhibition in Parkinson's disease, in fact, seems to be highly related to changes in PSD-slope (section 1.3.1).

The spectrum of the signal is usually obtained applying the modified Welch's periodogram, a method that consists of dividing the time series into segments, calculating the periodogram on each of them and then averaging the spectral estimated obtained. Thus, by considering a time series X(i)={x(1), x(2), …, x(N)} of length N divided into k=1, 2, … K segments of length M, for each subsequence of M values, indicated as $x_k(1)$, $x_k(2)$, …, $x_k(M)$, the PSD is calculated as:

$$Psd_k(f) = \frac{1}{MU}\left|\sum_{m=0}^{M-1} x_k(m)\, w(m) e^{-j2\pi f m}\right|^2 \qquad (2.3)$$

where Psd_k represents the periodogram of the k-th interval of M elements, W(m) ={w(1), w(2), …, w(M-1)} is the data window directly applied to each segment and U is the normalization factor given by:

$$U = \frac{1}{M}\sum_{m=0}^{M-1}|w(m)|^2 \tag{2.4}$$

At the end, the modified Welch's periodogram, that provides the PSD of the original time series, X(i) is obtained by the average of the modified periodograms Psd_k over all K intervals:

$$PSD_K(f) = \frac{1}{K}\sum_{k=0}^{K-1} Psd_k\,(f) \tag{2.5}$$

[34]. Once estimated the PSD, by showing its changes with respect to frequencies in a log-log graphic representation, the β-exponent is defined as the slope of the straight line obtained by performing an operation of linear regression. Figure 2.1 (a) shows a schematic representation of the power-law exponent procedure.

Nevertheless, to be sure that the estimation of β-exponent is not compromised by the rhythmic oscillatory components, as suggested by the work of Colombo et al. [35], the regression line can be identified following a peak removal operation. In other words, the peaks are accurately suppressed before performing the linear regression operation. Figure 2.1 (b) shows the two estimation of the power law exponent respectively before and after oscillatory peaks removal.

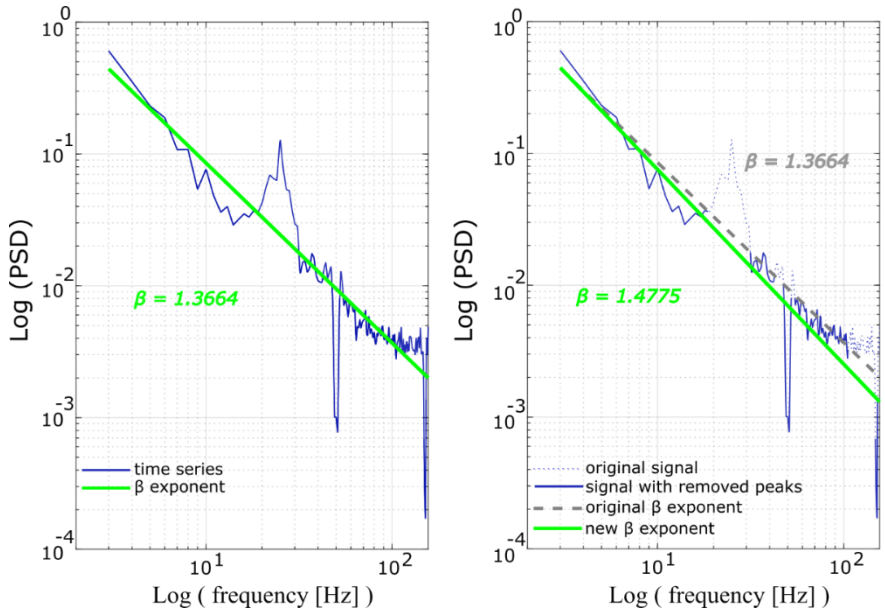


Figure 2.1 Example of application of the power-law exponent procedure applied to the same LFP signal from a PD patient before levodopa administration. The power spectral density (PSD) was obtained with the Welch's method setting windows of 2 seconds-length and overlap of 50%. The regression lines are the green ones respectively without (a) and after (b) the removal of oscillatory peaks, while the blue signal represent the PSD before (full line) and after (dotted line) the removal operation.

As regard the Matlab implementation, the spectra of the time series was obtained applying the modified Welch's periodogram with the *pwelch()* function defining, as input, the windows in terms of seconds-length and the percentage of overlapping segments. In this thesis, in relation to LFP and EEG time series, a Hamming window is used with an overlap of 50%.

Note that choice of the frequency resolution, given by sampling rate divided by the windows-length, may interfere with the estimation of the power-law exponent; thus, in this regard, the section 3.1.1 will describe in detail how these parameters affect the measurement.

2.1.2 Detrended fluctuation analysis

The detrended fluctuation analysis is the nonlinear parameter first proposed by Peng et al. [36], to evaluate the relationship between the variability of the time series and the length of the intervals over which the variability is measured. In this way, what is revealed is the level of self-similarity (or scale invariance) of the time series under analysis.

The index, indicated by α, is extracted by considering the root-mean-square fluctuations of integrated and detrended time series measured at different observation windows. Then, by representing in log-log coordinates how the fluctuations change against the size of the observation windows, the α-value is the slope obtained following an operation of linear regression on the embedded data.

In this regard, below the main mathematical steps involved in the implementation of the algorithm are outlined. Let X(i)={x(1), x(2), …, x(N)} of length N be the original time series under analysis, the average of the time series is calculated as:

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x(i) \tag{2.6}$$

with i=1, 2, ..., N. Then, the mean is subtracted from the time series and the integrated series is defined as:

$$Y(i) = \sum_{i=1}^{N} (x(i) - \bar{X}) \tag{2.7}$$

where Y(i) is the cumulative sum of the entire time series then divided into k=N/n non-overlapping intervals of length n. Within each k-th segment of length n, the *local trend* is calculated through the least-squares line fitted to the data and then removed.

In this way, the size of fluctuation (F_K) can be obtained as the standard deviation of the quantity (Y_k - $\bar{Y}_k$):

$$F_k(n) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(Y_k(j) - \bar{Y}_k)^2} \tag{2.8}$$

with k=1, 2, …, N/n and j=1, 2, … n. Y_k denotes the single integrated and detrended th-segment of the time series and $\bar{Y}_k$ the mean of Y_k defined as:

$$\bar{Y}_k = \frac{1}{n}\sum_{j=1}^{n} Y_k(j) \tag{2.9}$$

At the end, the size of fluctuation averaged over all the k intervals of size n is given by:

$$F(n) = \frac{1}{N/n}\sum_{k=1}^{N/n}(F_k(n)) \tag{2.10}$$

The steps are repeated for different values of n ∈ {N, N/2, N/3, …}.

Once a series of F(n) values is obtained with respect to a series of n window sizes, the operation of linear regression is performed in a double-logarithmic scale allowing the estimation of the fractal scaling α-exponent as the slope of the obtained straight line.

As mentioned before, the α-exponent provides the level of self-similarity in a time series giving information about the existing level of correlation among data. In detail, data result uncorrelated when α <0.5 and correlated when α >0.5. However, for sake of clarity it is important to underline that from an applicative point of

view, what seems to reflect neuronal activity in its dynamicity are the so-called "long range temporal correlations (LRTC)" present in the time series [37].

The analysis of the LRTC is linked to the DFA algorithm as indicated by Hohlefeld et al., that suggest to detect the α-exponent by applying the above-described steps on the amplitude envelopes of the oscillatory activity in specific frequency ranges. The envelopes are extracted by bandpass filtering the time series (finite impulse response filter, order 2000 and Hamming window) and by applying on the filtered signals the Hilbert transformation [20]. Figure 2.2 shows a schematic representation of the detrended fluctuation analysis procedure.

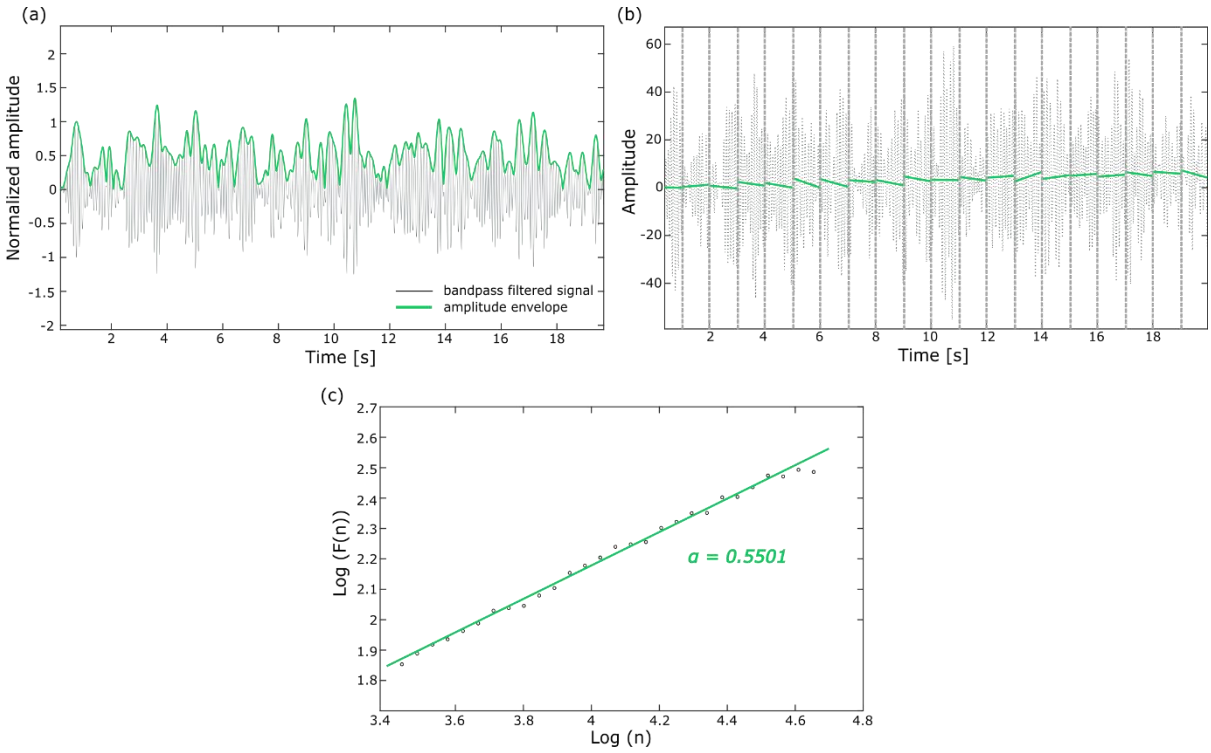


Figure 2.2 Example of application of the detrended fluctuation analysis in terms of LRTC approach. The procedure is applied to the same LFP signal from a PD patient before levodopa administration. In detail, (a) application of the bandpass filter and the Hilbert transformation to extract the envelope on which the DFA-method is then applied. (b) Division of the integrated series into non-overlapping segments of size n within which the data are fitted with the least-square method. The green lines represent the single trends. (c) Log-log representation of the size of the fluctuation F(n) in function of the window sizes (n). The green line represents the regression line fitting the data in the double logarithmic plot.

In this thesis, for the applications related to LFP and EEG time series, the LRTC approach, based on the amplitude envelopes of the neuronal oscillations, is implemented. From this perspective, as regard the Matlab implementation, the envelope was obtained by applying first the bandpass filter with the *filter()* function defining as inputs the numerator and denominator coefficients and the original time series and then the Hilbert transformation with the *hilbert()* function.

2.2 Approximate and sample entropy

The introduction of a family of measures called Approximate Entropy (*ApEn*) is ascribable to Pincus and Goldberger (1994). Such indices have been commonly used to quantify the amount of information or uncertainty content of physiological signals [38][39] because they estimate the level of randomness, the regularity and the predictability of a time series showing higher values for greater irregularity and lower values in opposite case.

Unlike the common Shannon entropy, defined as the information or the uncertainty associated with a set of possible events characterized by a certain probability *p* [40], Pincus' entropy does not estimate a probability distribution. The ApEn measures the logarithm of the relative frequencies with which two sequences in the series contain similar patterns [41].

From the mathematical point of view, ApEn is computed through the procedure described as follows. By considering a time series X(i)={x(1), x(2), …, x(N)} of N data points, the sequence of m (embedding dimension) points, defined *template*, is defined as:

$$x_m(i) = [x(i), \dots, x(i+m-1)] \tag{2.11}$$

with i=1, 2, ..., N-m+1. By considering the distance between each couple of sequences equal to:

$$d[x_m(i), x_m(j)] = \max_k [|x(i+k-1) - x(j+k-1)|] \tag{2.12}$$

where k=1, 2, …m, the number of sequences of m points similar to the template with a specific tolerance r is selected to compute $C_r^m(i)$ as:

$$C_r^m(i) = \frac{1}{N-m+1} \{\text{number of couples for which } d[x_m(i), x_m(j)] \leq r\} \tag{2.13}$$

that represents the regularity of the template in the time series for each i=1, 2, ..., N-m+1. Then, by taking the natural logarithm of each $C_r^m(i)$ and averaging it over a total of N-m+1, the quantity $\Phi^m(r)$ is given by:

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln\left(C_r^m(i)\right) \tag{2.14}$$

By repeating the previous steps after increasing the dimension to m+1, the entropic index is estimated as:

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \tag{2.15}$$

[42].

The sample entropy (*SampEn*) is another measure of complexity very similar to ApEn introduced by Richman and Moorman [43] to estimate the irregularity and the complexity of signals extracted from the heart rate variability.

The mathematical steps underlying the calculation of SampEn are reported following. Again, starting from a time series X(i)={x(1), x(2), …, x(N)} of N data points, the sequence of m points, defined *template*, is constructed as:

$$x_m(i) = [x(i), \dots, x(i+m-1)] \tag{2.16}$$

with i=1, 2, ..., N-m+1, and the distance between each couple of sequences is equal to:

$$d[x_m(i), x_m(j)] = \max_k [|x(i+k-1) - x(j+k-1)|] \tag{2.17}$$

where k=1, 2, …m and i ≠ j. The number of sequences of m points similar to the template with a specific tolerance r, is indicated as $B_r^m(i)$ and computed as:

$$B_r^m(i) = \frac{1}{N-m+1} \{\text{number of couples for which } d[x_m(i), x_m(j)] \leq r\} \tag{2.18}$$

In the same way, by considering the number of sequences similar for m+1 points, $A_r^m(i)$ is given by:

$$A_r^m(i) = \frac{1}{N-m+1} \{\text{number of couples for which } d[x_m(i), x_m(j)] \leq r\} \tag{2.19}$$

Then, $B_r^m(i)$ and $A_r^m(i)$ are averaged over i= N-m, providing:

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_r^m(i) \tag{2.20}$$

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_r^m(i) \tag{2.21}$$

and the entropic index is estimated as:

$$SampEn(m, r, N) = -\ln\left(\frac{A^m(r)}{B^m(r)}\right) \tag{2.22}$$

[42]. Figure 2.3 shows a schematic representation of the ApEn and SampEn computation. In practice, the main idea of both ApEn and SampEn is to calculate the conditional probability that two similar sequences of m points remain similar at the point m+1. Despite this, the two indices differ from each other based on how they manage the self-counting in the mathematical computation. Indeed, while for the first entropic index the procedure of self-counting is applied at each iteration in order to avoid the natural logarithm of zero, in the second index the natural logarithm is computed only once and the self-counting is excluded due to the necessary condition $i \neq j$ enounced in equation (2.15) [44].

Moreover, the performance of the estimator is dependent on the length N of the time series, on the sampling rate and on the two parameters that need to be specified before the application: the embedding dimension of the comparison pattern (m), representing also the detail level at which the signals is analyzed, and the margin of tolerance (r), that filters out the irregularities. Indeed, even if these factors are strongly correlated to each other, their impact plays a key role in ApEn and SampEn performances [42][43]. In this regard, section 3.2 will describe in detail how these parameters affect the measurement.

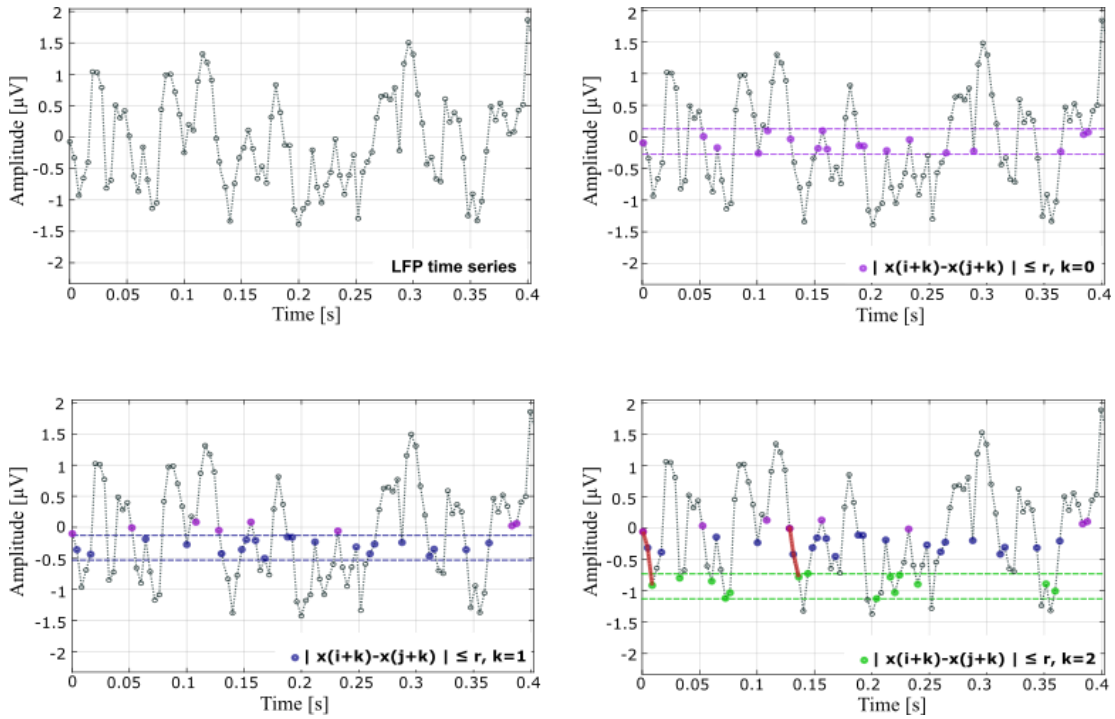


Figure 2.3 Example of the procedure implemented for the construction of the template on a LFP signal recorded from a PD patient. The template represents a fundamental process for calculating both ApEn and SampEn. Purple points verify that $|x(i+k)-x(j+k)| \leq r$ for k=0; blue points verify that $|x(i+k)-x(j+k)| \leq r$ for k=1; green points verify that $|x(i+k)-x(j+k)| \leq r$ for k=2. At the end, the template with m=2 is given by the red line in the bottom right panel of the figure.

2.3 Multiscale entropy

SampEn and ApEn represents traditional algorithms used to quantify the complexity; however, in some cases they seem to fail and not to quantify correctly the degree of randomness intrinsic in the time series. This happens when the series under analysis presents structures on multiple scales, as the case of LFP and EEG. For this reason, the need of introducing a new measure of complexity to capture the information at multiple time scales gave rise to the so-defined multiscale entropy (MSE) [45].

Two are the main steps used to estimate the MSE: the first regards the construction of a new time series generated from the original one, while the second involves the estimation of a measure of entropy, e.g. SampEn or ApEn, at each time scale. More specifically, by considering the time series X(i)={x(1), x(2), …, x(N)} of N data points, the coarse-grained sequence is originated by averaging consecutive points within specific non-overlapping windows of increasing length τ, named scale factor. The procedure implemented implies:

$$SampEn(m, r, N) = -\ln\left(\frac{A^m(r)}{B^m(r)}\right) \tag{2.23}$$

with $1 \leq j \leq N/\tau$. By repeating the previous step on different values of τ, multiple coarse-grained time series are defined and for each of them the SampEn (or ApEn) is calculated. At the end, all the entropic values are plotted against the scale factors [46]. In this thesis the MSE is implemented computing the SampEn for different scale factors.

Note that for τ=1 the coarse-grained time series is simply the original one. Figure 2.4 shows (a) the schematic representation of the construction of the coarse-graining

procedure for τ=2 and τ=3, while figure 2.4(b) depicts the graph of MSE complexity values versus scale factors.

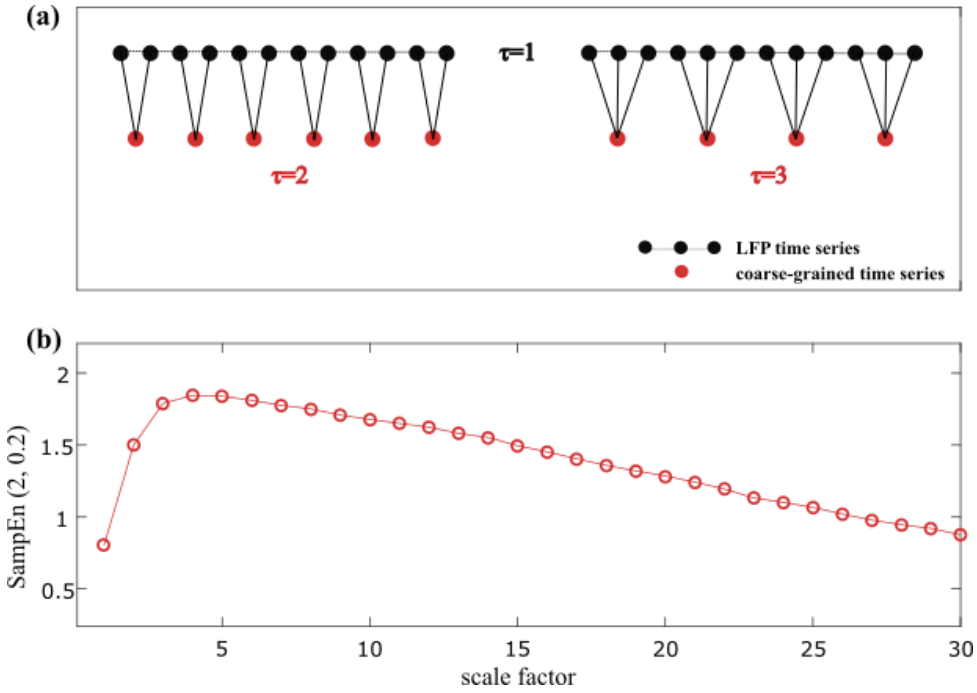


Figure 2.4 (a) Example of the coarse-graining procedure implemented on a LFP signal recorded from a PD patient. Starting from the original time series (dark circle), the average of consecutive samples within each window of length τ provides the coarse-grained time series (sequence of red circles). (b) Values of SampEn (2, 0.2) obtained at each scale factor τ. In this case the scale factor ranges from 1 to 30.

The described index can be used to compare the complexity among different time series. Indeed, by exploiting the MSE curves, a specific time series can be considered more complex than another if for the majority of the scales the entropy values (e.g. SampEn or ApEn) are higher than others; moreover, the monotonic decrease of the entropy values with scales suggests that the signal only contains information in the smallest scales [47].

3. Applicability criteria

The intention to identify and verify the conditions of applicability of the different nonlinear methods, which were accurately described in the previous chapter, represents the main objective of this thesis. It arises from the need to provide guidelines on the correct implementation of the measures and to the appropriate selection of the parameters which highly influence the application performance.

The analyses described in this chapter are carried out on a large number of artificial time series in order to investigate both parameters common to all the measures, such as data length and sampling frequency, and specific parameters linked to the implementation of each single measure. For sake of clarity, methods have been grouped into two different sections according to which aspects they analyze, i.e. whether the fractal behavior or the irregularity of the neural time series.

In this regard, section 3.1 outlines the applicability criteria relating to the fractal analysis that involves the power-law exponent (PLE) and the detrended fluctuation analysis (DFA); in this case, the discussion is centered on the discrimination between fractional gaussian noise (fGn) and fractional Brownian motion (fBm). On the other hand, section 3.2 investigates the conditions of application referred to the entropic analysis, including Approximate entropy (ApEn), Sample entropy (SampEn) and Multiscale entropy (MSE), whose parameters, such as embedding dimension of

patterns or tolerance, seem play an essential role in terms of performance of the algorithm.

3.1 Fractal analysis

As described in the previous chapter, the decision to group the PLE (β) and the DFA (α) under the section fractal finds confirmation in relationships which exists among them. These relations allow to theoretically derive one measure from the other and are different according to whether the time series under analysis is fGn or fBm [27]. The latter aspect suggests the possibility to consider them as useful tool to verify the robustness of the fractal algorithms, identifying also the more suitable parameter to classify the time series as fBm or fGn.

Hence, the idea is to carry out an analysis aimed at comparing the direct estimation of both β and α exponents and the values which are indirectly obtained by applying the theoretical relations (2.1) and (2.2).

At the beginning, the HE has been considered as discriminating parameter between the two processes. Thus, two sets of 1000 fGn and 1000 fBm were simulated for different values of HE as shown in figure 3.1, by implementing a function that receives as inputs the signal length and the HE.

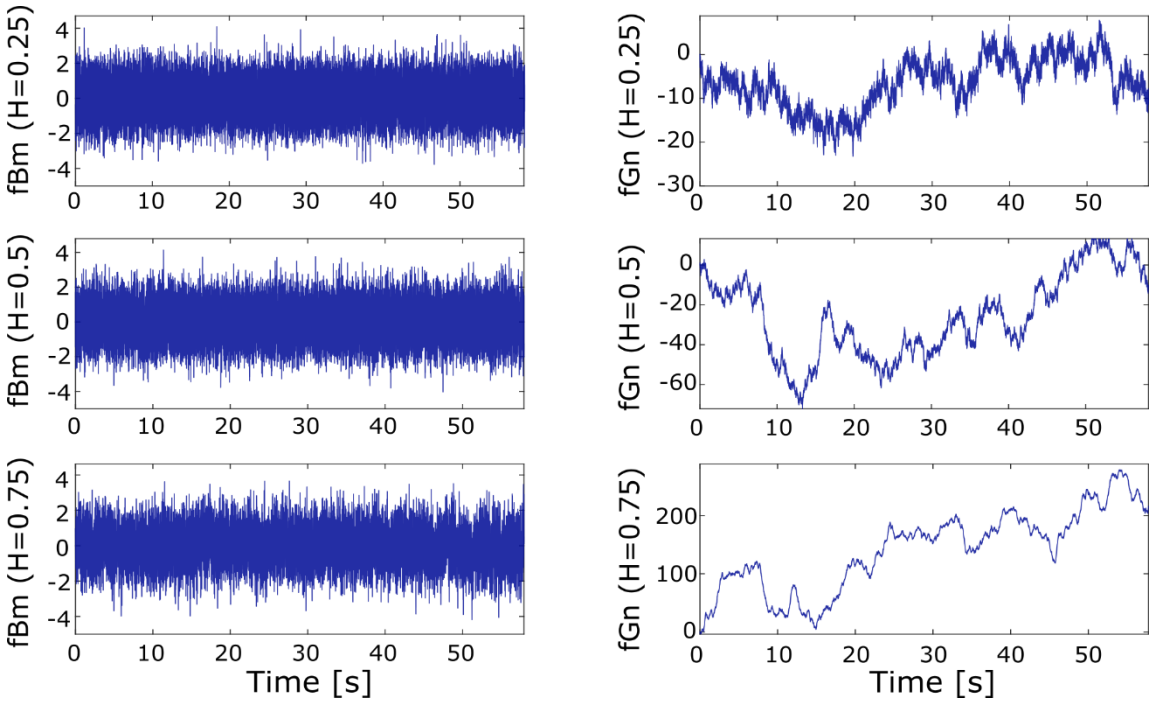


Figure 3.1 Simulated fGn and fBm paths for Hurst values equal to 0.25, 0.5, 0.75 respectively. The length of the signals is 60 seconds with a sampling rate of 512 Hz.

The signals are generated for HE values equal to 0.25, 0.5, 0.75, as suggested by the work of Eke et al. (2000), taken as reference for the current analysis [29]. The sampling rate was fixed equal to 512 Hz and the length of the signals equal to 60 seconds.

Table 3.1 summarizes the results obtained from the current analysis, reporting both the values of β and α estimated directly through the implementation of the nonlinear algorithms and β and α indirectly obtained by applying the theoretical relations (2.1) and (2.2).

	Results theoretically obtained		Results directly estimated	
	β - exponent	**α - exponent**	**β - exponent**	**α - exponent**
fGn (HE=0.25)	-0.5	0.25	-0.57 +- 0.02	0.25 +- 0.03
fGn (HE=0.5)	0	0.5	0.00 +- 0.02	0.49 +- 0.06
fGn (HE=0.75)	0.5	0.75	0.53 +- 0.02	0.74 +- 0.07
fBm (HE=0.25)	1.5	1.25	1.46 +- 0.02	1.25 +- 0.1
fBm (HE=0.5)	2	1.5	1.97 +- 0.02	1.48 +- 0.1
fBm (HE=0.75)	2.5	1.75	2.39 +- 0.03	1.72 +- 0.09

Table 3.1 Results indirectly obtained by the relations that link the two fractal measures selected in the current thesis and results directly estimated implementing each algorithm separately on the simulated fGn and fBm, originated from values of Hurst exponents (HE).

From table 3.1, the estimated results appear stable across the 1000 realizations and very close to the theoretically predicted values, thus demonstrating the validity of the fractal relations and the robustness of the implemented algorithms. However, it is worth noting that fBm and fGn are intrinsically characterized by the same HE, which falls within the range [0,1] in both cases. This makes the index less suitable for discerning the two classes. For this reason, a further simulation test was performed by selecting the PLE (β) as the discriminating parameter. The same analysis procedure was applied modulating the β-values, instead of HE, in order to obtain fGn (β equal to 0, 0.5 and 0.98) and fBm (β equal to 1.02, 1.5, 2).

In this regard, the signals have been generated through a specific Matlab function *dsp.ColoredNoise()* which requires, as input, the length in samples and a value between -2 and 2, which indicates, as the β-exponent, the slope of the spectrum in the frequency domain. The sampling rate was fixed equal to 512 Hz and the length of the signals equal to 60 seconds. The results both estimated and obtained are then compared and summarized in table 3.2.

	Results theoretically obtained		Results directly estimated	
	β - exponent	**α - exponent**	**β - exponent**	**α - exponent**
fGn (β =0)	0	0.5	0.00 +- 0.02	0.50 +- 0.06
fGn (β =0.5)	0.5	0.75	0.50 +- 0.02	0.74 +- 0.04
fGn (β=0.98)	0.98	0.99	0.98 +- 0.02	0.99 +- 0.04
fBm (β=1.02)	1.02	1.01	1.02 +- 0.02	1.01 +- 0.04
fBm (β=1.5)	1.5	1.25	1.48 +- 0.02	1.25 +- 0.04
fBm (β=2)	2	1.5	2.01 +- 0.02	1.48 +- 0.04

Table 3.2 Results indirectly obtained by the relations that link the two fractal measures selected in the current thesis and results directly estimated implementing each algorithm separately on the simulated fGn and fBm, originated from values of power-law exponents (β).

Table 3.2 displays results essentially similar to those found in table 3.1, confirming both the stability of the linear relationships existing between these fractal measures and the use of β-exponent as the most suitable parameter for the purposes of the classification between fBm and fGn.

However, for sake of clarity and especially in view of applications on real neurophysiological signals, it is important to underline that not always β values ranges are perfectly delineated as [-1:1] for fGn and [1:3] for fBm, with β=1 as critical boundary. This is due to the performance dependence of the PLE algorithm on transformations, like detrending or filtering, used to estimate the power spectrum. For this reason, several works in literature, starting from that of Eke et al. (2000) [29], have proposed to exclude high frequencies, i.e. 1/8 <f <1/2 Hz, from the calculation of the PSD. In this way, the ranges of β-values are redefined respectively as [-1: 0.38] for fGn and [1.04: 3] for fBm, introducing a new critical zone between 0.38 and 1.04 where further investigations are needed to distinguish the two processes correctly [27], [48].

Once the correct functioning of the selected fractal measures has been assessed on ad hoc simulations, the analysis proceeds with the evaluation of the parameters that play a key role in each single nonlinear algorithm.

3.1.1 Applicability criteria: power-law exponent

As demonstrated in section 3.1, the PLE is the method widely used to determine the signal class of time series. In the frequency domain, the fractal behavior is expressed through the power law:

$$PSD(f) \propto \frac{1}{f^{-\beta}} \tag{3.1}$$

where PSD(f) is the power spectrum density for the corresponding frequency f obtained by applying the modified Welch's periodogram. However, the choice of the frequency resolution and the parameters defined in the estimation of the power spectrum, such as the windows-length, need to be accurately selected to avoid misleading results. From this perspective, the effects that different frequency resolutions or different windows-length, as well as the sampling frequency and the data length, have on the β-values are evaluated below.

The analysis is carried out on two sets of 1000 time series generated through the Matlab function *dsp.ColoredNoise()* for β-values equal to 0.5 for fGn and 1.5 for fBm, respectively. Figure 3.2 shows simulated fGn and fBm paths for the selected β-values.

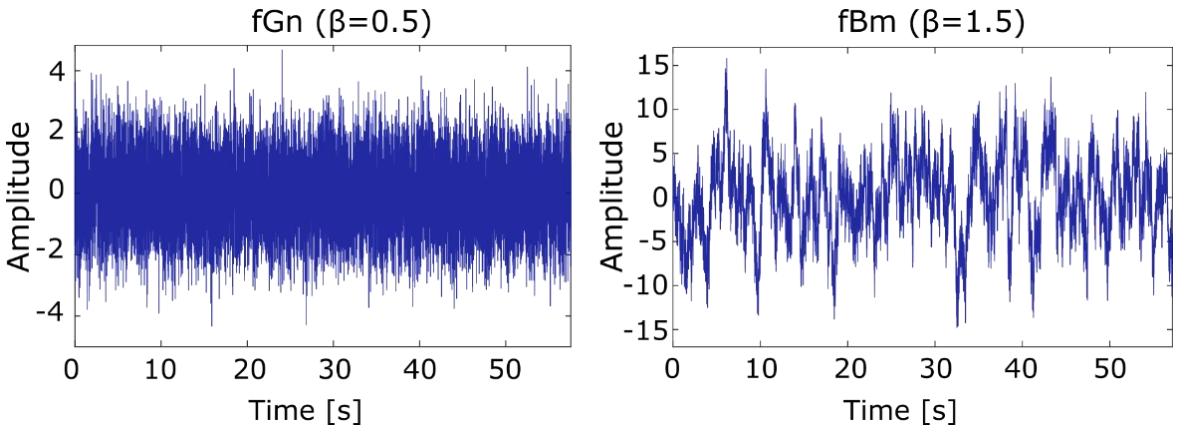


Figure 3.2 Example of simulated fGn and fBm paths for selected β-values equal to 0.5 and 1.5, respectively. The sampling frequency is 512 Hz, while the duration is about 10 seconds.

In order to test whether β-exponents change as function of the frequency resolution, the effect of using different window-lengths with a sampling frequency fixed as 512 Hz is examined. It is known that the frequency resolution is defined as:

$$\Delta f = \frac{f_s}{M} \tag{3.2}$$

where fs is the sampling frequency and M is the number of samples constituting the length of the windows into which the original signal is split. Thus, by considering different values of M, figure 3.3 shows the boxplot of β-exponents calculated for different frequency resolutions.

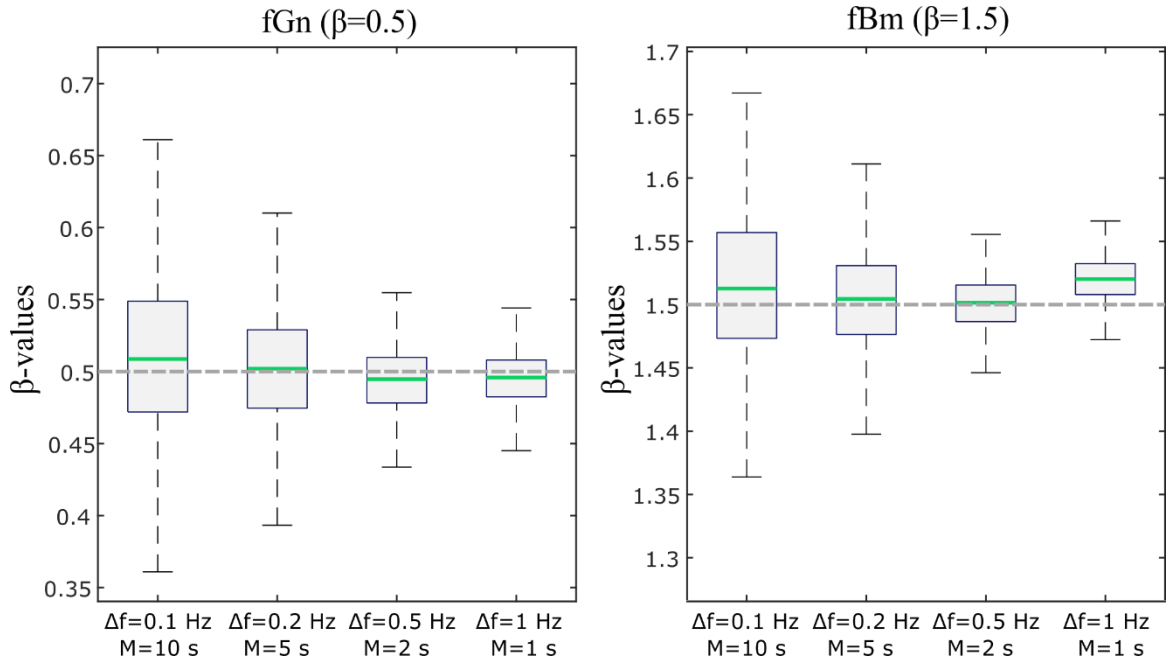


Figure 3.3 Boxplot comparison of β-exponents obtained by considering different values of frequency resolution (Δf). The dashed grey lines represent the expected β-exponents respectively for fBm and fGn, and the line in the middle of each box is the median value. The x-axes represent frequency resolution Δf, given by fs/M, with fs=512 Hz and M equals to 10, 5, 2, 1 seconds.

These results show that for time series characterized by a sampling frequency of 512 Hz and a length of 60 seconds, a frequency resolution of at least 0.5 provides an accurate evaluation of the exponent. This draws attention to the importance of correctly choosing the parameters that influence the PLE as fractal method used in biomedical signal processing.

In addition to this, it is also necessary to monitor other parameters that could influence the correct application of the algorithm, such as the length of the data and the sampling rate. However, in this case, these seem to depend on the frequency resolution. In fact, by repeating the same procedure aimed to test the influence of Δf on β-exponents for different lengths of the time series, for example 60, 120 or 180 seconds, the obtained results are similar to the one reported in figure 3.3.

3.1.2 Applicability criteria: detrended fluctuation analysis

The DFA-based method acts in the time domain to evaluate the self-similarity (or scale invariance) of the time series. The α-index provided by the algorithm expresses the fractal behavior as:

$$F(n) \propto n^{\alpha} \tag{3.3}$$

where $F(n)$ represent the fluctuations of the signal and n is the window size in which the fluctuations are evaluated. Despite being a powerful tool for determining if the data exhibit fractal characteristics, it is also highly dependent on input parameters, such as the data length (N) and the minimum (n_{min}) and maximum (n_{max}) window, or 'box' sizes, in which the fluctuations are estimated [49], [50]. Thus, the analysis outlined below investigates the main effects produced by such parameters on the performance of the DFA method.

The first aspect that must be taken into account is related to the way in which the window sizes are plotted against $F(n)$ forming the so-called "diffusion plot" on which the α-index is obtained as the slope of the linear regression (section 2.2). Indeed, the double-logarithmic nature of the plot appears to cause greater uncertainty, since the density of points increases for higher window sizes with respect to shorter ones. For this reason, a possible solution for overcoming this limitation is to consider the window sizes, in the diffusion plot, evenly spaced on the logarithmic scale [49], [50].

Once established the correct procedure to place the window sizes along the abscissa axis, the analysis aimed to evaluate the effects of changing N, n_{min}, n_{max} on the estimated α-values can be carried out on the same two sets of 1000 fGn and fBm, previously described, with sampling frequency equals to 512 Hz

- Influence of data length (N)

To verify the sensitivity of the algorithm to N, the following values were chosen: 500, 1000, 2500, 5000, 7500, 10000 samples. By selecting n_{min}, n_{max} equals to the smallest and largest possible values on the basis of existing literature [49], [51], i.e. 4 and N/10 respectively, figure 3.4 displays changes in α-values for the different data lengths.

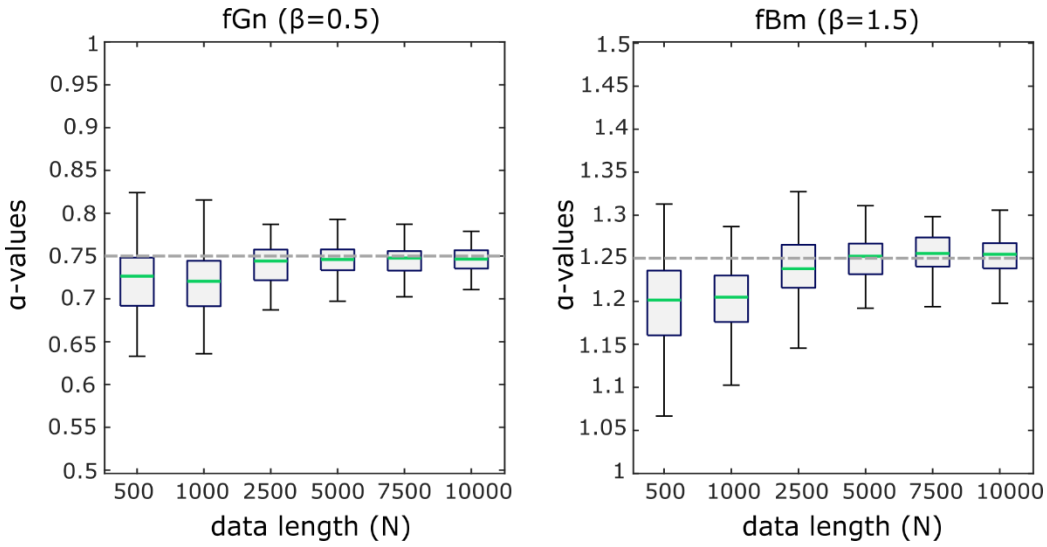


Figure 3.4 Boxplot comparison of α-values calculated for different data lengths. The sampling frequency is 512 Hz, while the minimum and the maximum window sizes are 4 and N/10. The grey dashed lines represent the expected α-exponents on the basis of relations (3.1) and (3.2) starting from β-values used to originate the simulated fGn and fBm time series, and the line in the middle of each box is the median value.

The findings revealed by the boxplots are consistent with ranges of α-values defined for fBm and fGn, i.e. [0:1] for the first class of time series and [1:3] for the second one. Size larger than 2500 samples are desirable to perform an appropriate fractal investigation through the DFA-method, while smaller data lengths, like 500 or 1000, seem to provide less accurate estimates of α-exponents.

- Influence of maximum window size (n_{max}).

This parameter proved to be the most influencing factor when evaluating the accuracy of the DFA method on simulated data. Larger values of window sizes may provide unreliable estimations of fluctuations, since the number of intervals involved in the calculation of the α-slope is smaller for increasing maximum window size. By considering for example n_{max} equal to N/2, the time series is divided into two intervals which contribute to the calculation of the fluctuation F(n). This contribution is less than the one provided by the ten intervals obtained by setting n_{max} = N/10. Hence, the n_{max} parameters chosen for the investigation of DFA are given by N/2, N/4, N/6, N/8, N/10. Figure 3.5 summarizes the obtained results, displaying boxplots of α-distributions for each selected window sizes. The y-axis indicates the values of DFA-exponents, while the x-axis the different values of n_{max} examined.

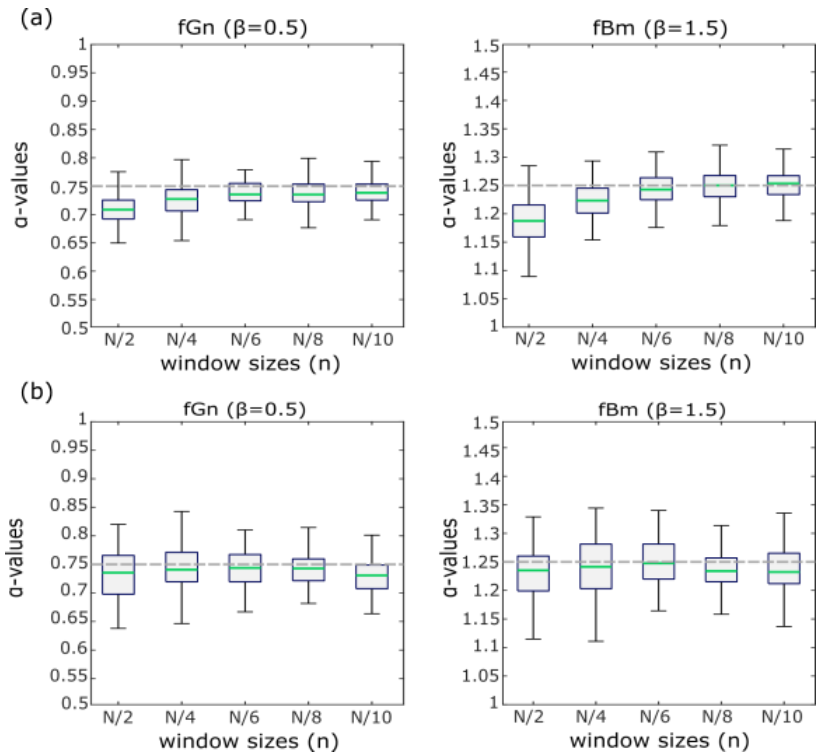


Figure 3.5 Boxplot comparison of α-values calculated for window sizes. The sampling frequency of the signals is 512 Hz, while the minimum window sizes is 4 samples and the data length are (a) 2500 samples and (b) 7500 samples. The grey dashed lines represent the expected α-exponents on the basis of relations (2.1) and (2.2) starting from β-values used to originate the simulated fGn and fBm time series, and the line in the middle of each box is the median value.

Again, the findings are consistent with ranges of α-values defined for fBm and fGn. However, the most interesting aspect carried out from the analysis is related to the high dependency between N and n_{max}. From this perspective, figure 3.5 clearly shows how N/2 and N/4 values of maximum window sizes provide generally unreliable estimates of DFA-exponent for time series with data length of 2500 samples (figure 3.5 (a)). Interestingly, an opposite trend is indeed shown when values of n_{max} equals to N/10 are applied on signals characterized by a data length of 7500 (figure 3.5 (b)). In this latter case, in fact, lower values of maximum window sizes, such as N/2 or N/4, offer more accurate estimation of α-exponents.

- Influence of minimum window size (n_{min})

This factor, unlike n_{max}, is not dependent on the data length. Indeed, by considering an analysis similar to the one carried out for n_{max}, thus evaluating changes in α-values, no differences were found. This shows that selection of n_{min} is not the most impactful parameter selection choice in terms of accuracy in estimating DFA α-exponents.

3.2 Entropy analysis

The entropy analysis includes indices that measure the level of complexity of time series. As largely outlined in previous chapters, the concept of complexity is related to the randomness of the signal and to the predictability of its future samples based on the previous ones [41]. In this study, it was decided to investigate in detail the Approximate Entropy (ApEn), the Sample Entropy (SampEn) and the Multiscale

Entropy (MSE), since such indices perform the quantification of the complexity through a similar procedure, as accurately described in chapter 2.

It is no coincidence, in fact, neither that SampEn was introduced to overcome limitations of ApEn, such as the presence of self-matches in the measure or the instability related to increasing length of the time series, nor that MSE represents an extension of the previous indices, since by calculating one of them at different scales, it provides information about the complexity of signals characterized by a multi-scale structure [43], [47].

However, as for fractal analysis, also the correct estimation of the entropic measures may be guaranteed by appropriately selecting the parameters involved in the algorithm implementation. Thus, in this regard, sections 3.2.1 and 3.2.2 discuss the main factors that influence the selected entropy indices, through an analysis carried out on 1000 random white noise sequences generated through the Matlab function *dsp.ColoredNoise()* with β=0, and 1000 pure periodic signals given by $\sin(10\pi t)$. An example of such simulated signals characterized by 10 seconds-length and sampling frequency equals to 512 Hz, is shown in figure 3.6.

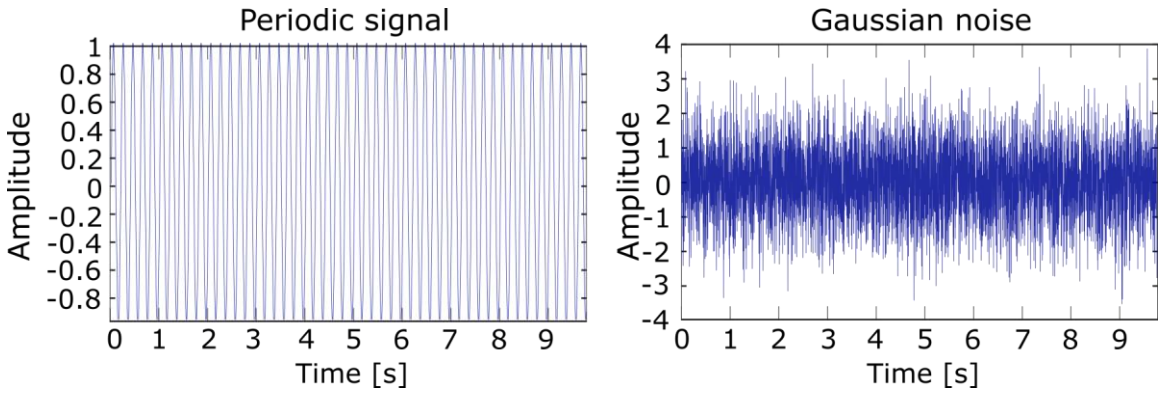


Figure 3.6 Example of simulated gaussian and periodic paths with sampling frequency of 512 Hz and duration of 10 seconds.

3.2.1 Applicability criteria: approximate and sample entropy

Approximate entropy and Sample entropy are methods that quantify the dynamical complexity of time series by computing the conditional probability that a given sequence close to another one over m consecutive points, maintains this property when one or more samples are added.

Generally, the factors that influences ApEn and SampEn indices are the length of the time series (N), the pattern dimension (m) and the tolerance (r). These, even if strongly correlated to each other, are investigated separately below. The works taken as references in the current analysis are those of Chen et al. (2005) and Richman e Moorman (2000) [42], [43].

- Influence of data length (N)

The length of the time series is linked to the embedding dimension (m), since the number of sample points constituting the time series is recommended about 10^m or at least 30^m for ApEn estimation [41] and about 10^m - 20^m for SampEn [43]. For this reason, entropy values change as function of N and in order to examine to what extent this happens, a set of four different data lengths consisting of 100, 1000, 5000 and 20000 is evaluated considering m=2. The simulated signal considered in the analysis is the gaussian noise shown in figure 3.6 with sampling frequency of 512 Hz.

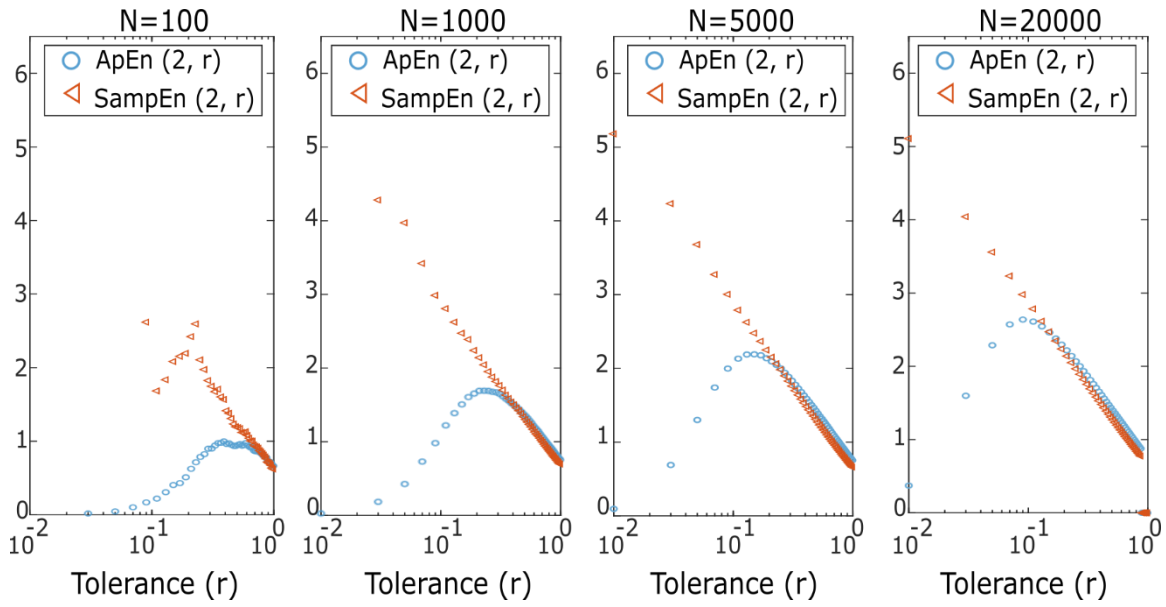


Figure 3.7 Effect of changing the N values on ApEn (left) and SampEn (right) for a Gaussian noise characterized by sampling frequency of 512 Hz and data length of 5000 samples. The parameter r ranges from 0.01 to 1with step 0.02, while the embedding dimension is fixed equal to 2.

The results related to Gaussian noise reveal two different trends for ApEn and SampEn, due to the respectively presence or absence of self-matches in the measure. Indeed, figure 3.7 shows that the ApEn values, characterized by self-matches, increase steadily up to a certain value of r and then decrease with increased r. On the other hand, SampEn values decrease with increased r, as expected, since the measure is designed not to count self-matches [42].

However, in this regard, it is important to underline that in both cases, the expected shape of the entropic distribution should linearly increase as r decreases [43], but if this seems true for SampEn, it appears not be the same for ApEn, which assumes the expected shape only for values of tolerance higher than a specific peak in the distribution. The value of such peak value is provided at lower r with increasing the number of data points.

All these findings suggest that the smaller the data length, the higher is the value of recommended r to avoid misleading entropic estimations.

- Influence of tolerance (r)

The parameter r balances the quality of logarithmic likelihood estimates with loss of time series information. For this reason, a good selection of r is needed, since lower values appear to cause poor conditional probability estimates while larger values, in contrast, seem related to the loss of too much information [52].

Existing rules in literature suggest to use values of r between 0.1 and 0.25 times the standard deviation of the original time series [41]. Thus, in this regard, figure 3.8 displays the results obtained implementing the analysis aimed to illustrate the sensitivity of SampEn and ApEn with respect to the choice of tolerance r.

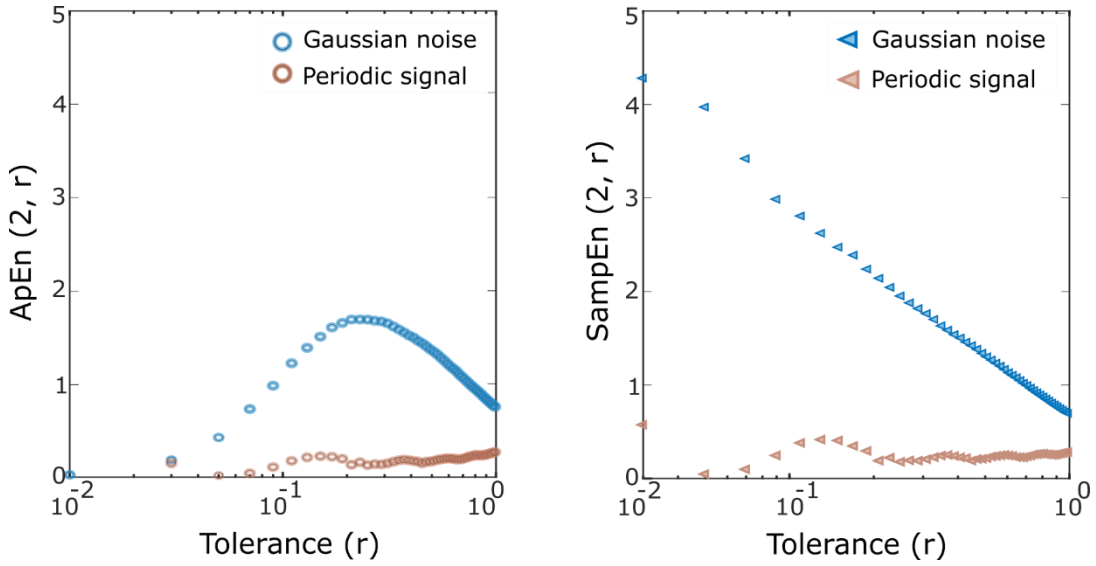


Figure 3.8 Effect of changing the r values on ApEn (left) and SampEn (right) for a Gaussian noise and periodic signal characterized by sampling frequency of 512 Hz and data length of 5000 samples. The parameter ranges from 0.01 to 1with step 0.02. The embedding dimension in both cases is m=2.

Figure 3.8 shows that the separation between the gaussian and periodic signals is guaranteed for almost all values of tolerance, even if it becomes smaller as the factor r increases suggesting a necessary caution when apply ApEn and SampEn with too large values of r.

Moreover, the peak value for ApEn is seen in correspondence of r=0.1 that means that the tolerance should be greater than 0.2 for the specific combination of parameters N=5000 samples and m=2. In the same way, even if more stable with respect to ApEn, also for SampEn the values of r should range between 0.1 and 0.2 since the distribution of entropic values doesn't assume the expected linear shape for $r \leq 0.1$. All these findings are in line with suggestion of Pincus et al in the perspective of obtaining more accurate estimations of entropic indices [41].

- Influence of embedding dimension (m)

The embedding dimension represents the length of sequences compared in the entropic algorithms and is highly dependent on the time series length (N), as mentioned before. In clinical applications, the most popular choices proposed in literature are m=1 and m=2. Thus, in this regard, figure 3.9 investigates how values of m ranging from 1 to 5 impact on ApEn and SampEn measures, requiring also further adjustment of the value of tolerance (r).

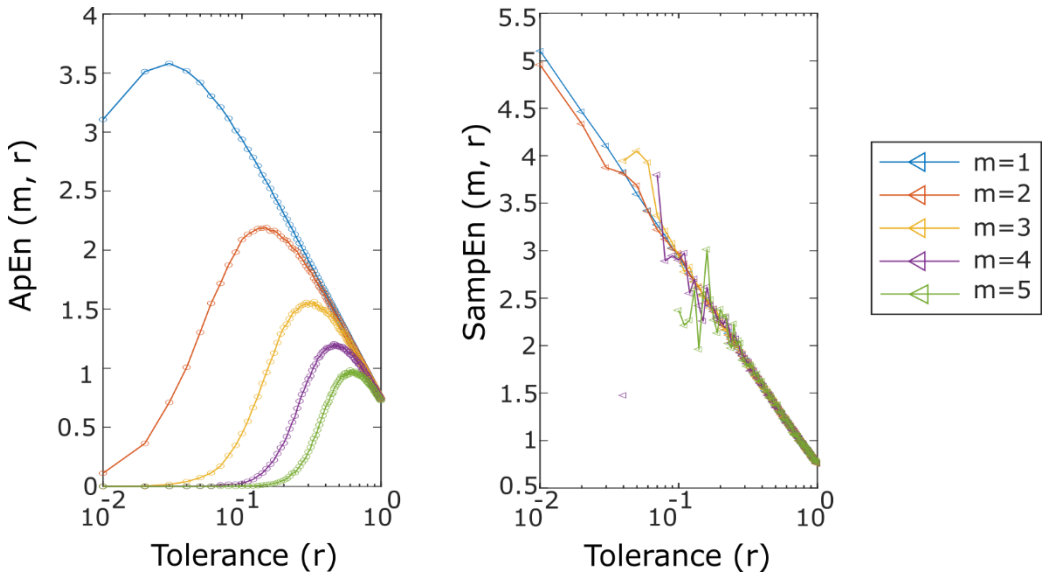


Figure 3.9 Effect of changing the m values on ApEn (left) and SampEn (right) for a Gaussian noise characterized by sampling frequency of 512 Hz and data length of 5000 samples. The x-axis represents values of tolerance (r) ranging between 0.01 and 1with step 0.02.

The results shown in figure 3.9 confirm the suggestions of Pincus et al., since m=1 and m=2 appear to provide more reliable indices estimations for both ApEn and SampEn. The proof of this is clearly pointed out by the shape of the distributions that deviates more and more from the expected one, with increased embedding dimensions ($m \geq 3$). This is true both for ApEn algorithm, since reducing the values of embedding dimension the range of possible values of tolerance appears also reduced, and for SampEn algorithm which seems not provide stable results with values of $m \geq 3$.

3.2.2 Applicability criteria: multiscale entropy

The multiscale entropy is a method designed for the analysis of complexity at different temporal scales which are generated using the so-called coarse-grained procedure, a combination of moving average filter and down-sampling process [53]. The factor which needs to be accurately discussed is the scale factor (τ), that determines the length of the coarse-grained time series on which the SampEn or ApEn is then computed.

The mathematical procedure, described in section 2.3, clearly shows that the number of samples in the time series provided by the coarse-graining procedure decrease considerably as the scale factor increases. This may cause an unstable estimation of entropy [53]. Thus, to evaluate the sensitivity of MSE to the signal length, 1000 realizations of white noises and 1/f noises are considered as function of sample points (N). By considering the parameters of SampEn equal respectively to m=2 and r=0.15 times the standard deviation of the signal, the MSE values are computed for data lengths equal to 800, 1000, 5000 and 20000. Figure 3.10 shows an example of the two types of time series for N=5000, while figure 3.11 depicts the MSE profiles for each value of N.

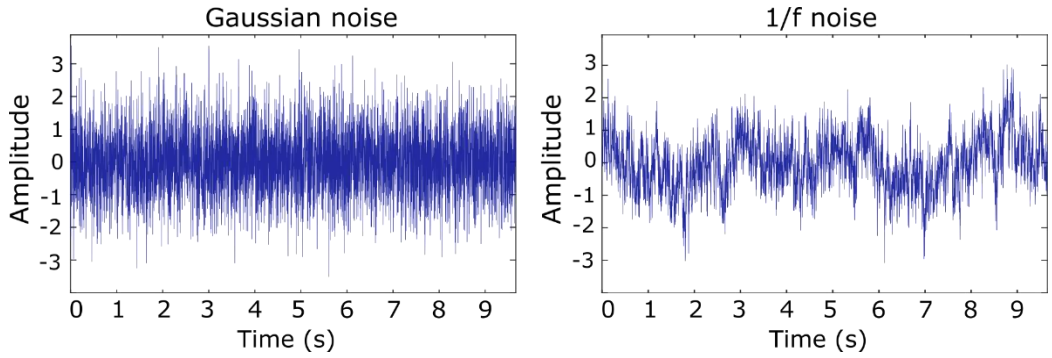


Figure 3.10 Example of simulated gaussian and periodic paths with sampling frequency of 512 Hz and duration of 10 seconds.

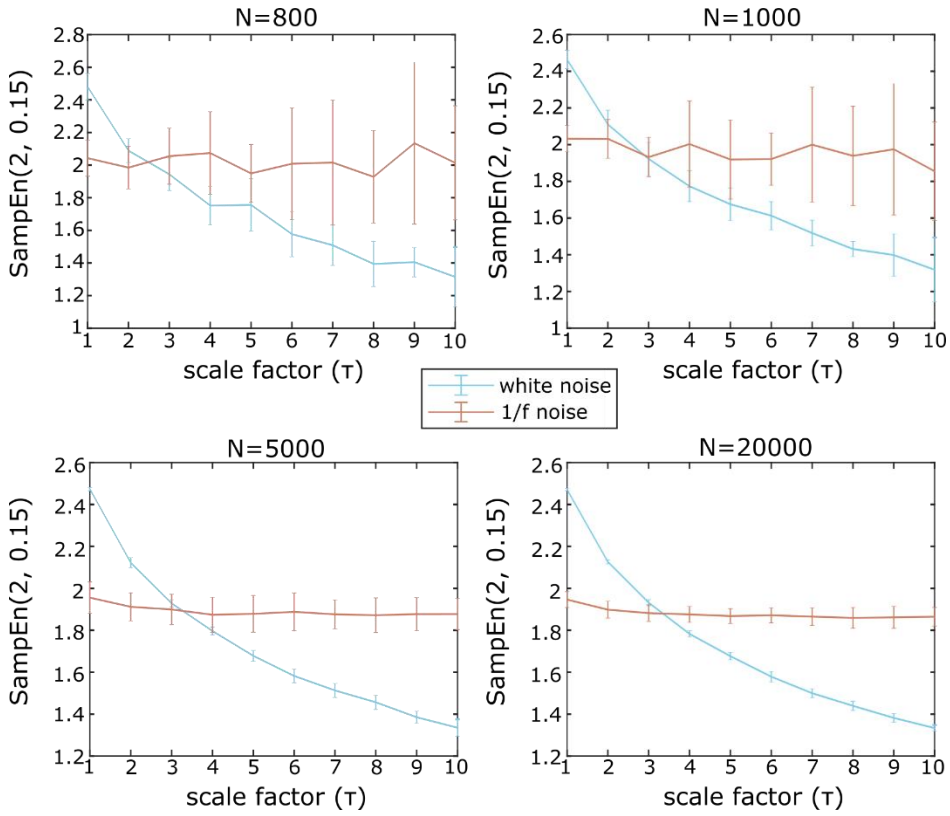


Figure 3.11 MSE profiles as function of data length (N) from 1000 white noises and 1000 1/f noises. The entropy values are expressed in terms of mean and standard deviation. Blue profile represents the white noise, while peach profile the 1/f noise.

MSE analysis, depicted in figure 3.11, points out that, at lower scale ($\tau \leq 3$) white noise has higher values of entropy than the correlated 1/f scaling noise. With increasing scales, the entropy of the white noise decreases monotonically and falls below the entropy value of the 1/f noise, which instead remains constant with scale. This finding suggests that correlated signals are more complex than uncorrelated ones, in the sense that the former contain complex structures across multiple time scales whereas the latter does not. In addition to this, the MSE curves suggest also that the higher the value of N, the more robust the multiscale sample entropy estimations.

Besides the influence of the data length, also the value of the maximum scale factor may interfere with the entropy measure. Thus, by considering two different values of scale factor, like 10 and 30, changes of MSE profiles are evaluated and shown in figure 3.12 for signal length of 5000 and 20000 samples. The embedding dimension and the tolerance are equal respectively to 2 and 0.15 times the standard deviation of the signal.

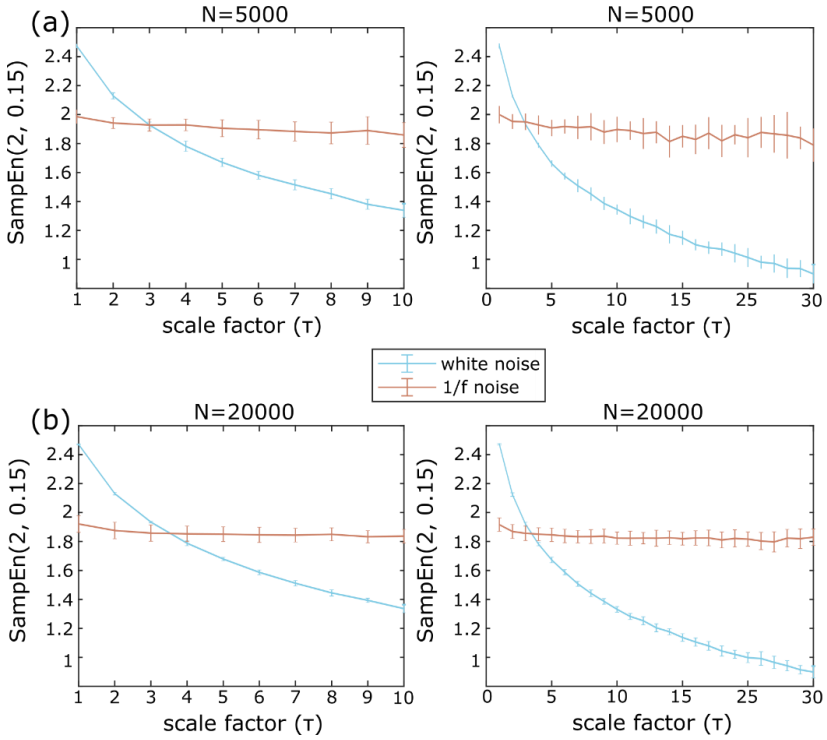


Figure 3.12 MSE profiles as function of scale factor (τ) from 1000 white noises and 1000 1/f noises. (a) indicates the evaluation for N=5000, (b) for N=20000 samples. The entropy values are expressed in terms of mean and standard deviation. Blue profile represents the white noise, while peach profile the 1/f noise.

The influence of the scale factor on MSE curves may be considered as an operation of down-sampling. In fact, as described in section 2.3, the coarse graining procedure provides new sequences starting from τ samples of the original time series. This suggests to select the most suitable τ_{max}, since when it is higher it causes a lower number of samples in the coarse-grained sequence.

4. Application of nonlinear measures to LFP and EEG time series

The usefulness of the applicability criteria investigated for the nonlinear measures will be demonstrated by an application to local field potentials recorded from Parkinson's Disease (PD) patients and an application to electroencephalograms acquired from both Alzheimer's Disease (AD) patients and from subjects affected by Lewy body dementia (LBD), frontotemporal dementia (FTD) and mild cognitive impairment (MCI).

In both cases, the two available datasets are used as a sort of test bench for the selected nonlinear measures described in chapter 2. The purpose is to evaluate the implementation of the nonlinear approaches on real neurophysiological time series, examining also the results in order to verify if they are in agreement with those present in literature.

In this regard, the current chapter provides an introduction of the two conducted studies, in terms of database description, data acquisition and pre-processing requirements. Such aspects are discussed in separate sections. In detail, section 4.1 outlines the analysis carried out on LFP of PD patients, while section 4.2 focuses on EEG of AD patients.

4.1 Study case: LFP analysis in Parkinson's disease

Levodopa represents the most common treatment strategy used to reduce parkinsonian motor symptoms. In this regard, the objective of the analysis carried out was to assess the nonlinear parameters before and after the levodopa administration, in order to evaluate which is the measure that better enounce the possible effects of the drug on patients.

4.1.1 Study population

The data used for the analysis have been provided by twenty-four subjects affected by Parkinson disease (12 females and 12 males). All patients recruited for the study gave their written informed consent to undergo data acquisition and the experiment was approved by the local ethics committees, in accordance with the Declaration of Helsinki.

The subjects, aged between 38 and 70 years, were submitted to bilateral implant for STN stimulation with macroelectrodes (model 3389 Medtronic Minneapolis, USA) and were considered in two different dopaminergic conditions, i.e. before and after levodopa medication [54]. All patients had predominantly rigid-akinetic phenotype with severe motor fluctuations and motor scores were assessed with the Unified Parkinson's Disease Rating Scale - UPDRS III after surgery (off medication, for patients recorded before levodopa and on medication for patients recorded after levodopa).

4.1.2 LFP recordings

During DBS electrode implantation, raw LFP recordings were acquired bilaterally from the subthalamic nucleus (STN) at rest before and after medication. The two different conditions were labeled “pre medication” and “post medication”, respectively and, in details, the first condition was referred to 12 h after withdrawal of levodopa treatment, while the second one to at least 30 minutes after drug administration.

In total, 56 LFP time series were recorded from right and left STN nuclei, 35 in condition of pre medication and 21 in condition of post medication. Table 4.1 summarizes further acquisition details.

All data, on which the work of thesis is conducted through an off-line analysis implemented in Matlab, were acquired with a sampling rate of 2500 Hz, while the number of bit and the voltage needed for quantization process were 12 and 5V range, respectively [54].

Patient	Gender	Recorded side	Recording condition	Time recording
P1	F	DX	Pre levodopa / post levodopa	356 sec /164 sec
P2	F	DX	Pre levodopa / post levodopa	60 sec /60 sec
		SX	Pre levodopa / post levodopa	81 sec /83 sec
P3	F	DX	Pre levodopa / post levodopa	56 sec /59 sec
		SX	Pre levodopa / post levodopa	57 sec /63 sec
P4	F	DX	post levodopa	94 sec
		SX	post levodopa	78 sec
P5	F	DX	pre levodopa	195 sec
		SX	pre levodopa	171 sec
P6	F	SX	Pre levodopa / post levodopa	375 sec /59 sec
P7	F	DX	pre levodopa	115 sec
		SX	pre levodopa	112 sec
P8	F	DX	pre levodopa	295 sec
		SX	pre levodopa	289 sec
P9	F	DX	pre levodopa	178 sec
		SX	pre levodopa	172 sec
P10	F	DX	Pre levodopa / post levodopa	54 sec /57 sec
		SX	Pre levodopa / post levodopa	60 sec /57 sec
P11	F	DX	pre levodopa	112 sec
P12	F	DX	Pre levodopa / post levodopa	81 sec /60 sec
		SX	Pre levodopa / post levodopa	289 sec /66 sec
P13	M	DX	Pre levodopa / post levodopa	26 sec /86 sec
		SX	Pre levodopa / post levodopa	66 sec /66 sec
P14	M	DX	Pre levodopa / post levodopa	63 sec /83 sec
		SX	Pre levodopa / post levodopa	63 sec /69 sec
P15	M	SX	Pre levodopa / post levodopa	108 sec /89 sec
P16	M	SX	Pre levodopa / post levodopa	336 sec /43 sec
P17	M	DX	pre levodopa	173 sec
		SX	pre levodopa	80 sec
P18	M	DX	pre levodopa	179 sec
		SX	pre levodopa	184 sec
P19	M	DX	pre levodopa	163 sec
		SX	pre levodopa	143 sec
P20	M	SX	Pre levodopa / post levodopa	106 sec /94 sec
P21	M	DX	pre levodopa	302 sec
		SX	pre levodopa	147 sec
P22	M	DX	pre levodopa	120 sec
		SX	pre levodopa	123 sec
P23	M	DX	post levodopa	66 sec
		SX	post levodopa	53 sec
P24	M	DX	post levodopa	69 sec
		SX	post levodopa	57 sec

Table 4.1 Details of LFP recordings analyzed.

4.1.3 LFP preprocessing

Data pre-processing was necessary to remove movements and power-line artifacts, prior to the application of the nonlinear measures. Four are the main steps included in this phase.

- Application of high pass filter through a sixth-order Butterworth filter with a cut-off frequency of 2 Hz with the purpose to highlight the activities of interest, reducing the noise components which affect the signals at lower frequencies (below 2Hz).
- Application of notch filter, commonly implemented to remove the power line noise (50 Hz) and its harmonics (multiples of 50 Hz), using a fourth-order Butterworth filter. The choice of this type of filter was related to the drop-off at the desired cut-off frequency as steep as possible and to the absence of "ripples" in the stopband. In order to avoid the introduction of new artifact during the notch filter application, the bandwidth of the filter needed to be accurately defined. In this regard, the bandwidth was set up equal to 4 Hz except for the patient P3 whose signals required a wider bandwidth (10 Hz). To avoid phase distortion, the filter was implemented as 'zero-phase' using the 'filtfilt' MATLAB function.
- Application of the normalization procedure aimed to reduce signal variability and ensure matching background noise in all recordings [55], [56]. LFPs were normalized by subtracting the mean and dividing the result by the standard deviation of the 600–1000 Hz band-pass filtered signals [54]. The operation is summarized in the formula 4.1:

$$\text{signal} = \frac{\text{signal} - \text{mean}(\text{signal}_{\text{filtered}})}{\text{std}(\text{signal}_{\text{filtered}})} \tag{4.1}$$

- Application of the artifact removal procedure that requires the division of the signal into non-overlapping segments of 100 ms; in each segment, the calculation of the peak-to-peak distance is performed in order to eliminate those higher than two times the standard deviation of the segmented signal.

All the described steps of the pre-processing phase were conducted on the original time-lengths of recorded time series (ranging from 24 seconds up to 6 minutes, as accurately indicated in Table 4.1). However, to enable a quantitative and qualitative comparison amongst nonlinear measures, the definition of a unique length equal to 43 seconds was necessary. Such temporal length allowed not only to treat separately just the signal presenting inferior duration (24 seconds) but also to work on a number of points contained in each time series sufficiently high (107500 given by 43 seconds multiplied by the sampling rate).

4.2 Study case: EEG analysis in Alzheimer's disease

Among the causes of cognitive impairment on neurodegenerative bases, Alzheimer's disease (AD) is the most frequent. To this group of pathologies are also involved other brain disorders, such as mild cognitive impairment (MCI), frontotemporal dementia (FTD) and Lewy body dementia (LBD). Briefly, MCI is a clinical state between elderly normal cognition and dementia, featuring memory complaints and cognitive impairment. It is a brain disorder with a high risk to

progress to AD, in fact it was estimated that annually between an 8 and 15% of people with MCI progress to AD [57]. FTD is a common dementia encompassing progressive dysfunction in executive functioning, behavior, language and movement. It is a disease still poorly recognized due to the overlapping clinical symptoms with respect to the AD [58], [59]. LBD is a degenerative disease which exhibits equivalent deficits in many cognitive abilities affected by AD, but assessing better scores than AD patients on most of tests focused on verbal memory.

Starting from these descriptions, it is easy to understand why the diagnosis and the effective treatments of MCI, FTD and LBD are still critically debated. Indeed, the lack of a clear characterization of these pathologies is related to both misleading definition of the clinical symptoms and the brain region interested by the disease, usually confounded with those that affect the common AD. In this regard, the purpose of the analysis carried out is to assesses the application of the nonlinear parameters in order to verify if they may be used to differentiate the four groups of patients.

4.2.1 Study population

The data used for the analysis have been provided by 8 subjects affected by Alzheimer's disease (5 females and three males), 4 subjects affected by Frontotemporal dementia (3 males and 1female), 2 patients affected by Lewy body dementia (males) and 3 patients affected by Mild cognitive impairment (two females and one male). All patients, aged between 63 and 85 years, recruited for the study were followed up at the "Centro per I disturbi cognitive e demenze" (CDCD) of the "Ospedale Maggiore Policlinico di Milano" and gave their written informed consent

to undergo data acquisition. Global cognitive impairment was assessed with Mini-Mental State Examination (MMSE), a screening test for mental deterioration [60].

4.2.2 EEG recordings

The EEG recordings, performed from 19 electrodes positioned according to the International 10-20 System, were acquired from the scalp of each subject at rest in both eyes-open and eyes-closed conditions. The two different conditions were labeled "eyes-closed" and "eyes-open", respectively, and the time recording was about 10 minutes with 5 minutes acquired when patient had their eyes open.

In total, 608 EEG time series were recorded from 19 channels, 304 in eyes-open condition and 304 in eyes-closed condition. Table 4.2 summarizes further acquisition details. The data, on which the work of this thesis was conducted through an off-line analysis implemented in Matlab, were acquired with sampling rates of 512 Hz in some cases and 2048 Hz in others. The proposed parameters are calculated for all electrodes and for nine different subcategories of electrodes defined as follow:

- (LF) left frontal: Fp1, F7, F3 positions;
- (RF) right frontal: Fp2, F4, F8 positions;
- (LC) left central: T3, C3 positions;
- (RC) right central: T4, C4 positions;
- (LT) left temporal: T5, P3, O1 positions;
- (RT) right temporal: T6, P4, O2 positions;
- (LFT) left fronto-temporal: Fp1, F7, F3, T3 positions;

- (RFT) right fronto-temporal: Fp2, F4, F8, T4 positions;
- (Z) z-axis: Fz, Cz, Pz positions.

The ground electrode was placed at FCz and connected mastoid electrodes will be used as references. Contact impedances will be kept below 10 kΩ each.

Patient	Gender	Disease	Recording condition	Sampling frequency
P1	F	AD	Eyes-closed	512 Hz
P2	M	AD	Eyes-closed/eyes-open	512 Hz
P3	F	AD	Eyes-closed/eyes-open	512 Hz
P4	M	AD	Eyes-closed/eyes-open	512 Hz
P5	F	AD	Eyes-closed/eyes-open	512 Hz
P6	F	AD	Eyes-open	2048 Hz
P7	M	AD	Eyes-closed/eyes-open	2048 Hz
P8	F	AD	Eyes-closed/eyes-open	2048 Hz
P9	F	FTD	Eyes-closed/eyes-open	512 Hz
P10	M	FTD	Eyes-closed/eyes-open	512 Hz
P11	M	FTD	Eyes-closed/eyes-open	2048 Hz
P12	M	FTD	Eyes-closed/eyes-open	512 Hz
P13	M	LBD	Eyes-closed/eyes-open	2048 Hz
P14	M	LBD	Eyes-closed/eyes-open	2048 Hz
P15	M	MCI	Eyes-closed/eyes-open	2048 Hz
P16	F	MCI	Eyes-closed/eyes-open	2048 Hz
P17	F	MCI	Eyes-closed/eyes-open	2048 Hz

Table 4.2 Details of EEG recordings analyzed

4.2.3 EEG preprocessing

Data pre-processing, prior to the application of the nonlinear measures, were necessary to remove movements and power-line artifacts. Five are the main steps included in this phase.

- Application of the artifact removal procedure based on the implementation of the independent component analysis (ICA). The ICA technique is commonly applied to EEG time series to reduce the artifacts without altering the brain activity. It is based on the research of independent components after the separation of the signal into its subcomponents and on correlation coefficients calculated between the initial and the "cleaned" signal in order to quantify the changes induced in each recording channels by removing the desired component[61]. The aim is to find a more significant representation of the data through a linear transformation of the same.
- Extraction of continuous 60-seconds artifact-free epochs, selected by visual inspection.
- Application of band-pass filter through a fourth-order Butterworth filter with cut-off frequencies of 2 Hz and 120 Hz with the purpose to highlight the portion of the signal characterized by information.
- Application of notch filter, commonly implemented to remove the power line noise (50 Hz) using a fourth-order Butterworth filter. The choice of this type of filter was related to the drop-off at the desired cut-off frequency as steep as possible and to the absence of "ripples" in the stopband. In order to avoid the introduction of new artifact during the notch filter application, the bandwidth of the filter needed to be accurately defined. In this regard, the bandwidth was set up equal to 2 Hz for all the patients. To avoid phase distortion, the filter was implemented as 'zero-phase' using the 'filtfilt' MATLAB function.
- Application of the normalization procedure based on the simple subtraction of the mean to reduce signal variability and to ensure matching background noise in all recordings.

All the described steps of the pre-processing phase were conducted on the original recorded time series sampled with sampling frequencies ranging from 512 Hz up to 2048 Hz (as accurately indicated in Table 4.2). However, to enable a quantitative and qualitative comparison amongst nonlinear measures, the definition of a unique sampling frequency equal to the minimum 512 Hz was necessary. This means that a further preprocessing step referred to the operation of down-sampling was implemented for the signal recorded at 2048 Hz.

5. Results

The purpose of this chapter is to present the main results obtained from the application of nonlinear measures on real data. The ability of the proposed methods is inspected by the use of real LFP and EEG time series, outlined in the previous chapter. In detail, the current chapter is divided into two parts. Section 5.1 related to the results of the study conducted on PD patients and section 5.2 focused on the outcomes carried out from the detection of dynamical alterations in AD subjects and patients affected by other neurological dementias.

In relation to the statistical analysis, it is worth stressing that for the first case study the differences were statistically evaluated in: i) a subset of patients whose LFP was recorded in conditions of both pre and post levodopa administration; ii) a set which includes all nuclei of PD patients considered individually. In this regard, the nonparametric Wilcoxon signed rank test *signrank()* was used to compare groups in condition i and the nonparametric Mann-Whitney rank sum test *ranksum()* was used to compare groups in condition ii. As regard, instead, the second case study, the small sample size of the dataset under analysis makes it difficult to evaluate differences from a statistical point of view. This is an obvious limitation of this analysis that may be object of future investigations.

5.1 Results: LFP analysis in Parkinson's disease

5.1.1 LFP classification

The first analysis conducted on pre-processed LFP signals concerned the classification of the signals as fractional Brownian motion (fBm) or fractional Gaussian noise (fGn), in order to evaluate the applicability of nonlinear measures.

The power-law exponent algorithm was examined by means of the slope of the regression line defined on the power spectral density (PSD) of the signal between 2 and $\frac{f_{max}}{8}$ Hz. The idea of estimating the parameter (β) within the aforementioned range has been clarified in the chapter 3 concerning the investigation of the applicability criteria on fractal measures (in detail, section 3.1).

In this regard, figure 5.1 displays the distributions of β-values estimated in condition of pre and post levodopa medication. The two vertical lines indicate the critical boundary, necessary for the signal classification, and are positioned at β = 0.38 and β = 1.04 for fGn and fBm respectively. More in details, β-values positioned on the left of the first critical boundary can be classified as fGn, while β-values positioned on the right of the second critical boundary can be classified as fBm.

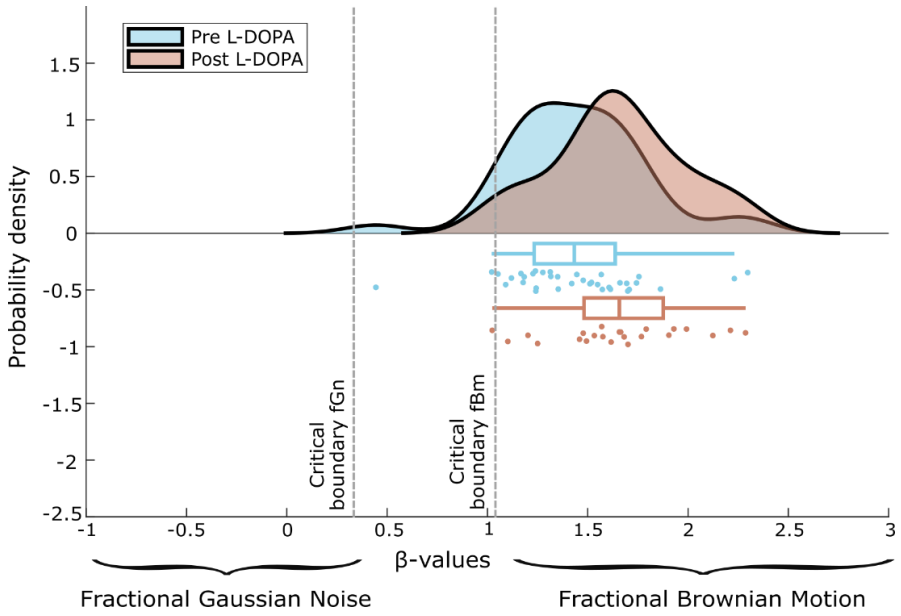


Figure 5.1 Result of LFP time series classification as fBm or fGn. The figure shows the distributions of β-values estimated in condition of pre (light-orange) and post (light-blue) levodopa medication. The two vertical lines mark the boundary beyond which the signal can be correctly classified as fBm or fGn. The range between the two boundaries presents the 40% uncertainty in signal class discrimination.

The analysis, made on all LFP time series, reveals that most of the signals appeared to be correctly considered as fBm while very few cases (three) have fallen into the range $0.38 < \beta < 1.04$ where there is about 40% uncertainty in defining the signal class [29]. These few cases were included in the nonlinear analysis, but marked in order to be distinguished from the fBm LFP signals.

5.1.2 LFP linear spectral analysis

In accordance with literature [55], [56], the power spectrum is used to characterize the STN oscillatory pattern in five main frequency bands, i.e. δ (2-7), α (8-12), low-β (13-20), high-β (20-30), low-γ (60-80), high-γ (250-350) Hz. The PSD values within each band are calculated through the steps described in section 2.1.1; then, the spectral power is expressed in normalized units by considering:

$$P_{(f_1-f_2)} = \frac{1}{f_2 - f_1}\int_{f_1}^{f_2} PSD(f)\, df \tag{5.1}$$

where f_1 and f_2 are the boundary frequencies of the considered band (f_1-f_2). The values of normalized spectrum appear as negative values because they are provided in logarithmic scale. Table 5.1 summarizes the p-values of the two different non-parametric tests performed to compare logarithmic values of normalized PSD in conditions of pre and post levodopa administration. The differences considered significant were those that present p-value<0.05.

Analyzed frequency bands	Log (PSD) values pre-post levodopa pairs	Log (PSD) values pre-post levodopa distributions
δ (2-7) Hz	0.7564	0.4647
α (8-12) Hz	0.1477	0.3166
low-β (13-20) Hz	0.0980	**0.0345**
high-β (20-30) Hz	0.5350	0.4843
low-γ (60-80) Hz	**0.0200**	**< 0.001**
high-γ (250-350) Hz	**0.0437**	**0.0388**

Table 5.1 P-values of the two different non-parametric tests performed to compare logarithmic normalized PSD concerning pre and post Levodopa conditions. The second column shows p-values obtained from the test conducted only on patients that present LFPs both in pre and post Levodopa conditions. The third column shows p-values obtained from the test conducted on all patient's LFP. The first column defines the frequency bands in which the two tests are carried out.

Significant differences are found in low-γ (60 - 80 Hz) and high-γ (250 - 350 Hz) frequency bands. The findings reveal how the various LFP rhythms are differently affected by levodopa, since in the first band the power spectra decrease in condition of post levodopa administration while oppositely in the second band increase. In addition to this, also the low-β frequency band provides significant outcomes, even if only in the condition in which all nuclei of PD patients are considered individually. In detail, higher values of logarithmic PSD characterize the condition of post

levodopa medication. All Significant changes in normalized power spectra are displayed in figure 5.2.

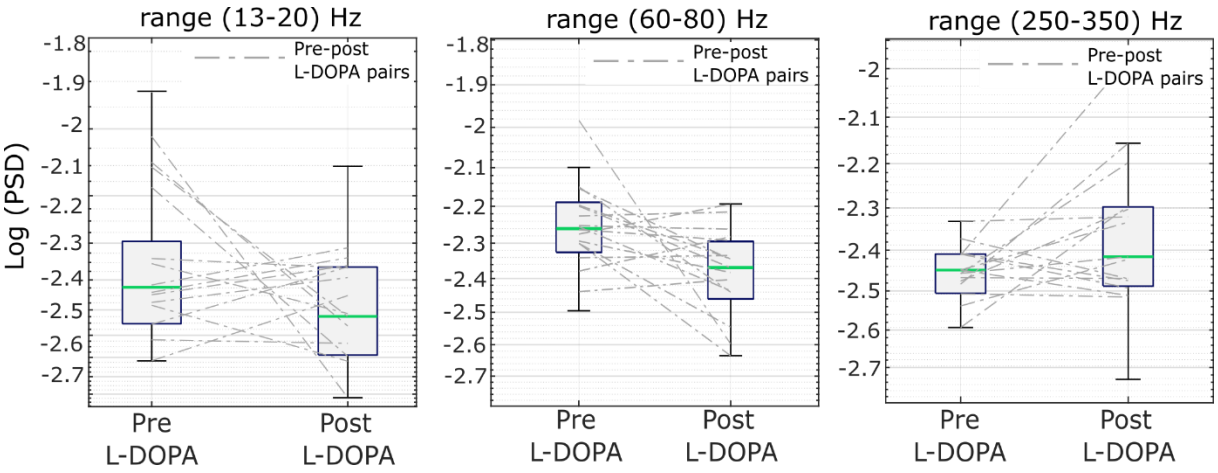


Figure 5.2 Boxplot comparison of logarithmic power spectra values for low-β, high-β and high-γ frequency bands. Tops and bottoms of each box are the 25th and 75th percentiles, the green line is the median value of each distribution while the dashed grey lines link the values referred to pre and post conditions respectively. Each boxplot provides the distribution of values estimated in condition of pre and post levodopa administration.

5.1.3 LFP power law exponent

The power-law exponent algorithm estimates β-exponents both on the whole spectrum (2 - $\frac{f_{max}}{8}$) and on the spectra computed on signals filtered in the ranges (2-40) and (30-40) Hz. The parameters set for the algorithm application are those described in section 3.1.1, while the choice of the bands, in which the analysis is computed, is suggested by works present in literature. Indeed, the range (2-40) Hz is supposed to be referred to the presence of activity at lower frequencies (f <50 Hz) in PD, while the range (30-40) is related to the excitation/inhibition ratio of neurons [17]. Table 5.2 again displays the p-values of the two different non-parametric tests performed to compare β-values in conditions of pre and post levodopa

administration. The differences considered significant were those that present p-value<0.05.

Analyzed frequency bands	β-values pre-post levodopa pairs	β-values pre-post levodopa distributions
(2 - $\frac{f_{max}}{8}$) Hz	**0.0199**	**0.0171**
(2-40) Hz	**0.0151**	**0.0469**
(30-40) Hz	0.5014	0.8115

Table 5.2 P-values of the two different non-parametric tests performed to compare β-exponents of pre and post Levodopa conditions. The second column shows p-values obtained from the test conducted only on patients that present LFPs both in pre and post levodopa conditions. The third column shows p-values obtained from the test conducted on all patient's LFP. The first column defines the frequency bands in which the two tests were carried out.

Results reveals no significant differences in the frequency band (30-40) Hz, while the analysis conducted on the other two ranges of frequency shows that, after levodopa medication, the β-exponents are significantly higher with respect to the condition before the drug administration. Such changes in the (2 - $\frac{f_{max}}{8}$) and (2 - 40) Hz frequency bands are displayed in figure 5.3.

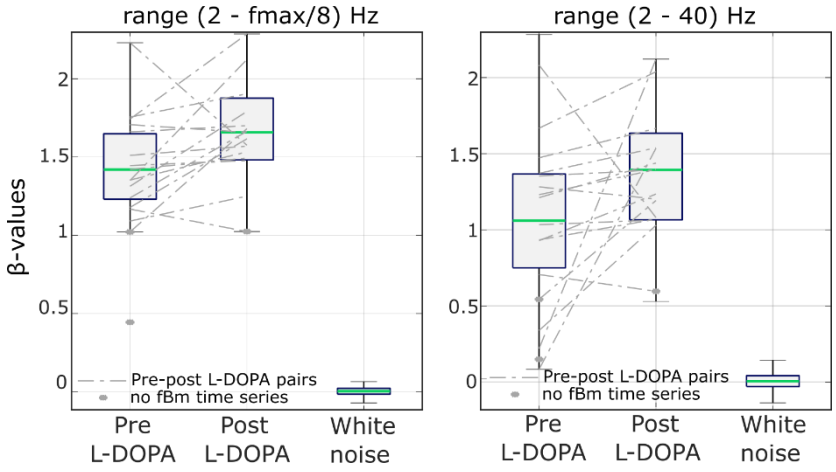


Figure 5.3 Boxplot comparison β scaling exponent quantifying power-law exponent on both total spectrum (left) and low-frequency range (right). Tops and bottoms of each box are the 25th and 75th percentiles, the green line is the median value of each distribution while the dashed grey lines link the values referred to pre and post conditions respectively. Each boxplot provides the distribution of β-values estimated in condition of pre and post levodopa administration and β-values estimated on 1000 simulated white noise.

5.1.4 LFP: detrended fluctuation analysis

The long-range temporal correlations[1], characterized by the DFA algorithm, are evaluated on the time series filtered in the common frequency bands adopted from the EEG literature: α (8-12), low-β (13-20), high-β (20-30), low-γ (60-80), high-γ (250-350) Hz. The parameters set for the algorithm application are slightly different with respect to those descripted in section 3.1.2, since the procedure implemented for the detection of LRTC requires a filtering operation and Hilbert transformation before the application of the classical DFA-method. For this reason, the analysis was conducted on 30 windows sizes logarithmically-spaced from 1 seconds to 20 seconds. The choice of considering 1 second as minimum window sizes allows to avoid temporal correlation introduced by signal filtering, while the largest window of 20 seconds allows to have at least 2 segments of that size.

Results are summarized in table 5.3 that shows the p-values of the two different non-parametric tests performed to compare α-values estimated in the pre and post levodopa conditions. The differences considered significant were those that present p-value<0.05.

[1] The long-range temporal correlation, through DFA, was also examined centering the analysis in proximity of a pronounced peak. However, no statistical results were found.

Analyzed frequency bands	α-values pre-post levodopa pairs	α-values pre-post levodopa distributions
α (8-12) Hz	0.3519	0.7265
low-β (13-20) Hz	0.3258	0.4175
high-β (20-30) Hz	0.7959	0.5043
low-γ (60-80) Hz	0.2343	0.1979
high-γ (250-350) Hz	**0.0299**	**0.0100**

Table 5.3 P-values of the two different non-parametric tests performed to compare α-exponents of pre and post levodopa conditions. The second column shows p-values obtained from the test conducted only on patients that present LFPs both in pre and post Levodopa conditions (paired test). The third column shows p-values obtained from the test conducted on all patient's LFP. The first column defines the frequency bands in which the two tests were carried out (test for independent samples).

The α-values discloses no significant differences in all the bands of interest except for that referred to the high-frequencies (250-350) Hz, in which the scaling exponent was higher for the post than for the pre levodopa condition, as depicted in figure 5.4.

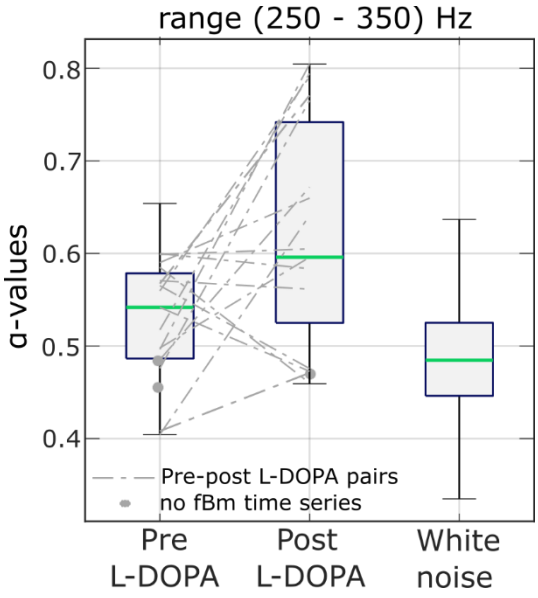


Figure 5.4 Boxplot of scaling exponent α quantifying long-range temporal correlation. Tops and bottoms of the box are the 25th and 75th percentiles, the green line is the median value of each distribution while the dashed grey lines link the values referred to pre and post conditions respectively. The boxes provide the distribution of α -values estimated in condition of pre and post levodopa administration compared to α -values estimated on 1000 simulated white noise.

5.1.5 LFP: multiscale entropy analysis

The multiscale sample entropy estimates the sample entropy on several time scales. The procedure is applied both on pre-processed LFP time series without any filtering operation and on the signals filtered in the frequency band (2-45) Hz. The number of scale factors considered is 30, while the parameters set for the SampEn application are m=2 and r=0.2 times the standard deviation of the given signals. In this regard, figure 5.5 displays the profiles of MSE curves obtained through the application of MSE algorithm.

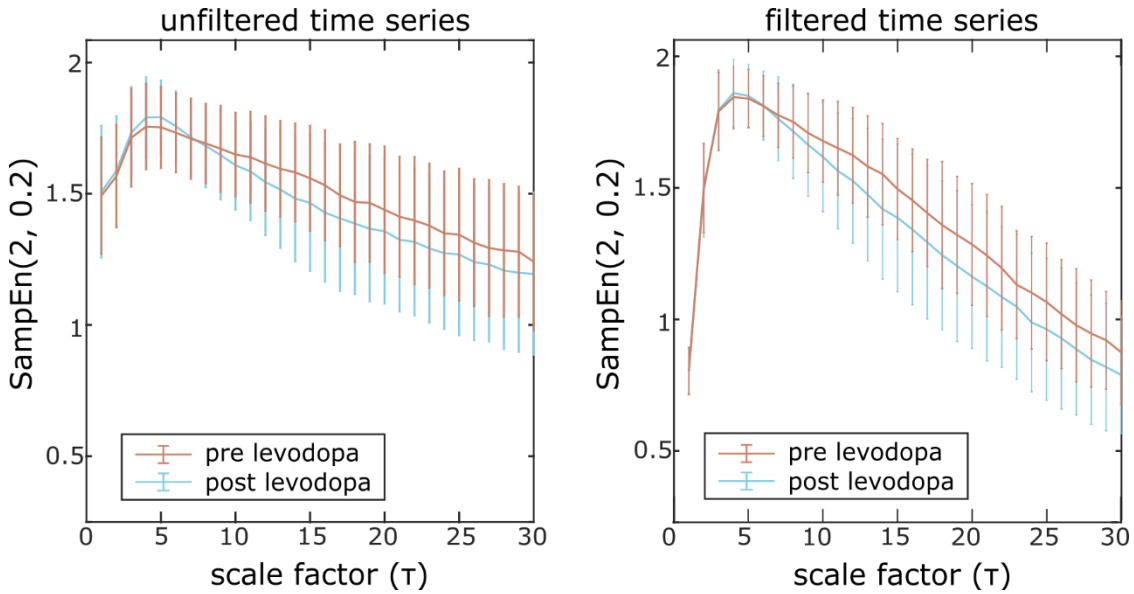


Figure 5.5 MSE curves of the SampEn values obtained in conditions of pre and post levodopa administration. The curves are shown in terms of mean value and standard deviation of the results. The embedding dimension is m=2, while the tolerance is r=0.2 times the standard deviation of the data sequence.

Figure 5.5 shows a less steep decreasing curve from scale 4 to 30 for condition of pre levodopa with respect to post-levodopa. However, no significant differences are found from a statistical point of view, as reported in table 5.4. The measure extracted to compare changes of the entropic index between pre and post levodopa conditions are the slopes computed on MSE curves. In detail, such slopes are evaluated both on the smaller time scales, i.e. $1 \leq \tau \leq 3$, and on the larger ones, i.e. $4 \leq \tau \leq 30$.

Analyzed range of scale factors	MSE-slopes pre-post levodopa pairs	MSE-slopes pre-post levodopa distributions
$1 \leq \tau \leq 3$, unfiltered time series	0.5014	0.9746
$4 \leq \tau \leq 30$, unfiltered time series	0.2775	0.2332
$1 \leq \tau \leq 3$, filtered time series	0.7960	0.5353
$4 \leq \tau \leq 30$, filtered time series	0.7960	0.1119

Table 5.4 P-values of the two different non-parametric tests performed to compare MSE slopes concerning pre and post Levodopa conditions. The second column shows p-values obtained from the test conducted only on patients that present LFPs both in pre and post Levodopa conditions. The third column shows p-values obtained from the test conducted on all patient's LFP. The first column defines the frequency bands in which the two tests were carried out.

The results are obtained through the evaluation of the two different non-parametric tests performed and reveal no significant correlations between changes in complexity and levodopa medication. This absence of differences can be clearly seen in the following figure 5.6, in which values of MSE slopes are displayed for the unfiltered LFP time series. Note that similar outcomes have been provided also considering the filtered time series.

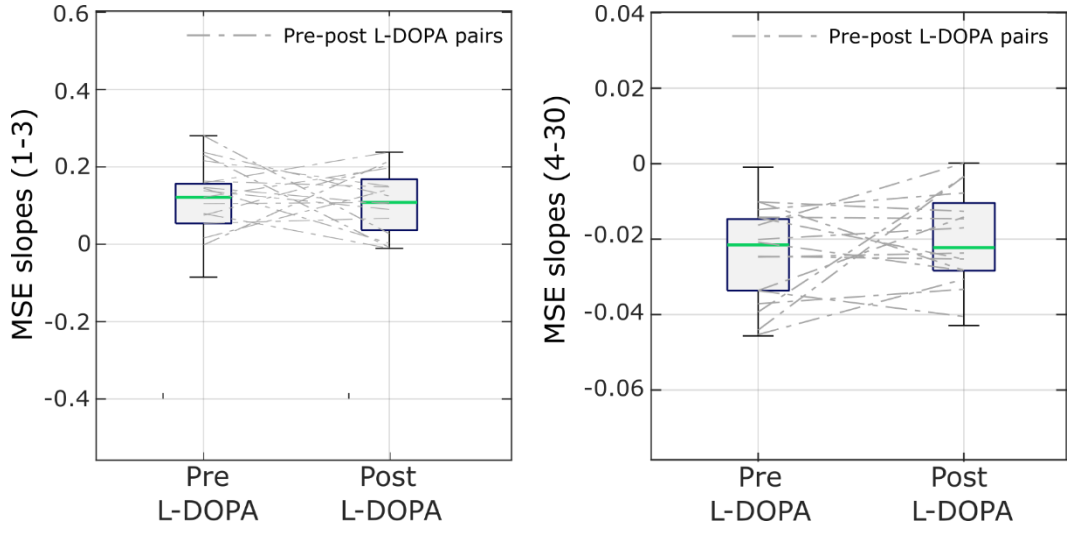


Figure 5.6 Boxplot of MSE slopes quantifying the grade of complexity in PD patients before and after levodopa medication. Tops and bottoms of the box are the 25th and 75th percentiles, the green line is the median value of each distribution while the dashed grey lines link the values referred to pre and post conditions respectively. The boxes provide the distribution of MSE slope -values estimated in condition of pre and post levodopa administration.

5.2 Results: EEG analysis in Alzheimer's disease

5.2.1 EEG classification

The first analysis conducted on pre-processed EEG signals concerned the classification of the signals as fractional Brownian motion (fBm) or fractional Gaussian noise (fGn), in order to evaluate the applicability of nonlinear measures.

The power-law exponent algorithm was examined by means of the slope of the regression line defined on the power spectral density (PSD) of the signal between 2 and $\frac{f_{max}}{8}$ Hz. The idea of estimating the parameter (β) within the aforementioned range has been clarified in the chapter 3 concerning the investigation of the applicability criteria on fractal measures (section 3.1).

In this regard, figure 5.7 displays the distributions of β-values estimated in eyes-open and eyes-closed conditions for the different groups of diseases. The two vertical lines indicate the critical boundary, necessary for the signal classification, and are positioned at β = 0.38 and β = 1.04 for fGn and fBm, respectively.

More in details, β-values positioned on the left of the first critical boundary can be classified as fGn, while β-values positioned on the right of the second critical boundary can be classified as fBm.

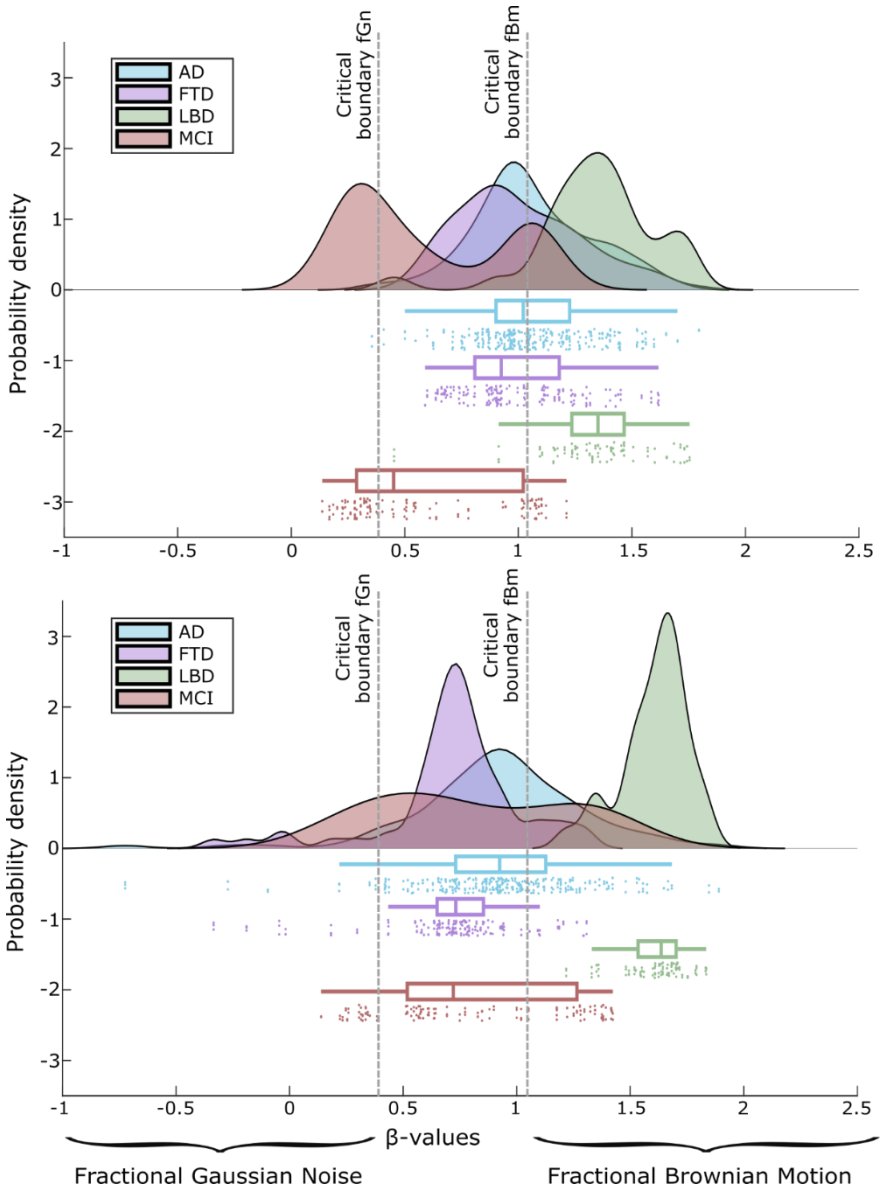


Figure 5.7 Result of EEG time series classification as fBm or fGn. The figure shows the distributions of β-values estimated in condition of eyes-closed (up) and eyes-open (down) for the four different categories of disease. The two vertical lines mark the boundary beyond which the signal can be correctly classified as fBm or fGn. The range between the two boundaries presents the 40% uncertainty in signal class discrimination.

The analyses, made grouping together all the EEG time series recorded from the different electrodes of acquisition, reveal that most of the signals appeared to fall into the range $0.38 < \beta < 1.04$ where there is about 40% uncertainty in defining the signal class [29]. Interestingly, this is not true for signals recorded from patients affected by LBD, for which the signals may be correctly considered as fBm.

5.2.2 EEG: linear spectral analysis

The analysis in the frequency domain is used to evaluate the power spectra in the following spectral bands: θ (2-7), α (8-12), β (13-35) Hz. The PSD within each band is calculated through the steps described in section 2.1.1; then, the spectral power is expressed in normalized units by considering

$$P_{(f_1-f_2)} = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} PSD(f)\, df \tag{5.2}$$

where f_1 and f_2 are the boundary frequencies of the considered band (f_1-f_2). The values of normalized spectrum appear as negative values because they are provided in logarithmic scale. Figure 5.8 displays changes in normalized PSD in relation to the frequency band 2-7 Hz.

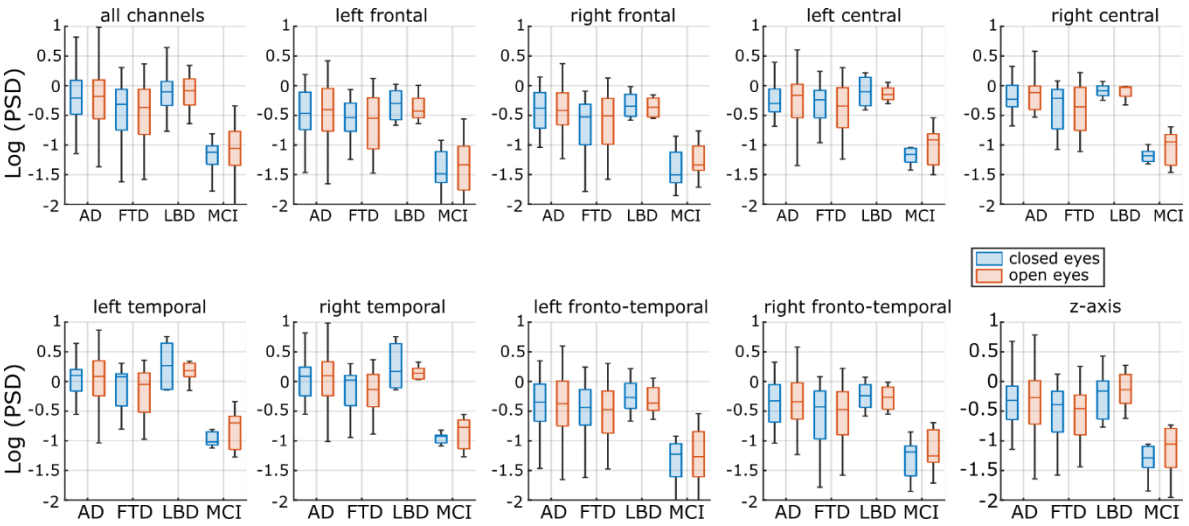


Figure 5.8 Boxchart comparison of normalized power spectra of EEG signals for θ frequency band. The blue boxes represent the closed-eyes condition, while the peach ones the open-eyes condition. Tops and bottoms of the boxes are the quartiles, while the horizontal lines defined within each distribution correspond to the median.

Results reveal an overall decrease of normalized PSD values for MCI patients in both eyes-open and eyes-closed conditions. In comparison to the other groups, this is valid for all the different subcategories and also in the α (8-12 Hz) and β (13-35 Hz) frequency bands not shown below since they exhibit variations similar to those presented in figure 5.8.

5.2.3 EEG: power law exponent

The power-law exponent algorithm was used to estimate β-values of the power spectra computed on EEG signals filtered in the ranges (2-45) Hz. The parameters set for the algorithm application are those described in section 3.1.1, while the choice of the band is related to two main aspects. The first one is application of the high-pass filtering with the frequency cut off at 2 Hz to which the EEG time series used in this thesis have been subjected. This makes it necessary to apply the power-law exponent algorithm in a band whose lower frequency is $f_L \geq 2$ Hz. The other aspect is

related to the application of the notch filter, with the purpose to avoid its influence on the β-values estimations. Figure 5.9 displays changes in β-values estimated for the different subcategories of electrodes in closed-eyes and open-eyes conditions. The x-axis indicates the class of diseases to which the boxplots belong, while the y-axis represents the value of β-exponents.

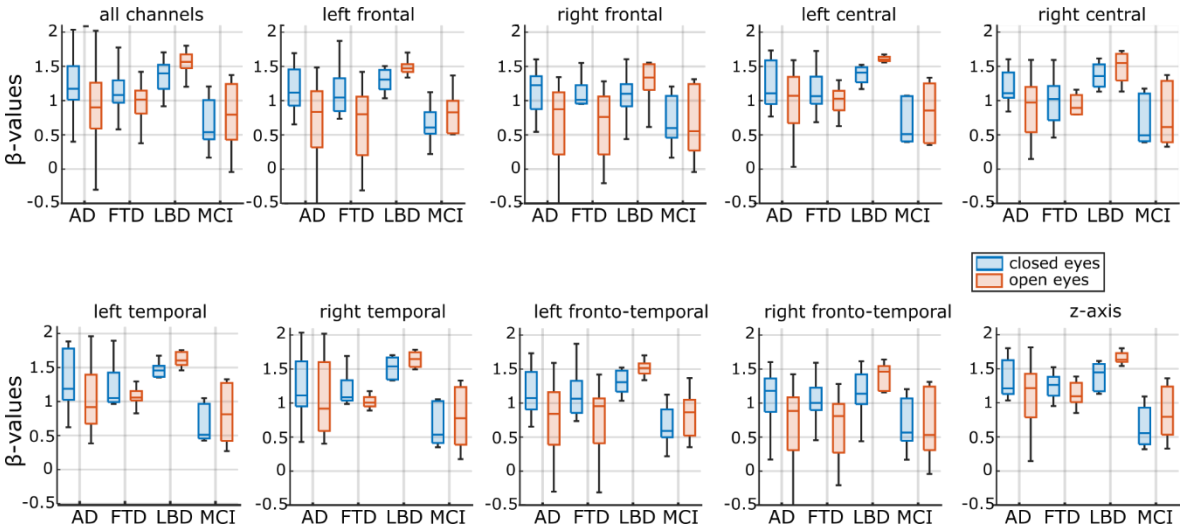


Figure 5.9 Boxchart comparison of power-law exponents (β) of EEG signals filtered between 2 and 45 Hz. The blue boxes represent the closed-eyes condition, while the peach ones the open-eyes condition. Tops and bottoms of the boxes are the quartiles, while the horizontal lines defined within each distribution correspond to the median.

The findings reveal lower β-values of MCI patients compared to the other groups in the eyes-closed conditions for all the different subcategories, and that LBD patients exhibited higher β-values than the other groups in the eyes-opened conditions for all the different subcategories.

The small number of patients involved in the study lead to consider the changes shown in figure 5.9 as a trend rather than something statistically defined.

5.2.4 EEG: detrended fluctuation analysis

The detrended fluctuation analysis, in terms of long-range temporal correlation approach, estimates α-exponents in the common frequency bands adopted from the EEG literature: α (8-12), β (13-30), low-γ (30-40), high-γ (60-80) Hz. The parameters set for the algorithm application are the same involved in the LFP analysis (section 5.1.4). Figures 5.10 and 5.11 display changes in α-values for all the subcategories of electrodes in closed-eyes and open-eyes conditions, evaluated respectively in the frequency bands α (8-12) and low-γ (30-40) Hz. Again, the x-axis indicates the class of diseases, while the y-axis represents the values of α-exponents.

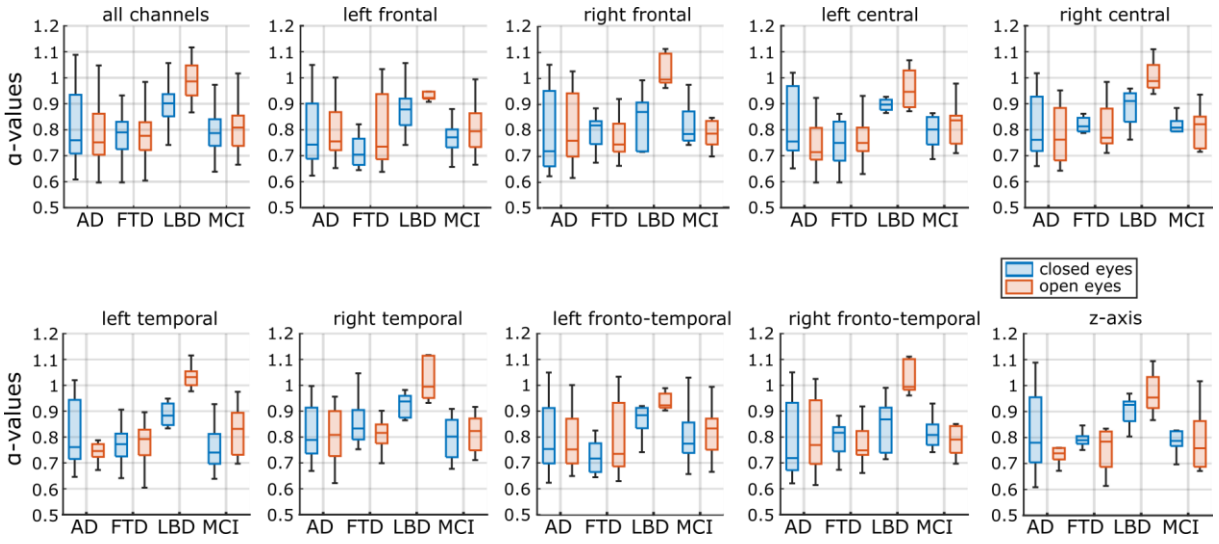


Figure 5.10 Boxchart comparison of detrended fluctuation analysis of EEG signals filtered between 8 and 12 Hz. The blue boxes represent the closed-eyes condition, while the peach ones the open-eyes condition. Tops and bottoms of the boxes are the quartiles, while the horizontal lines defined within each distribution correspond to the median.

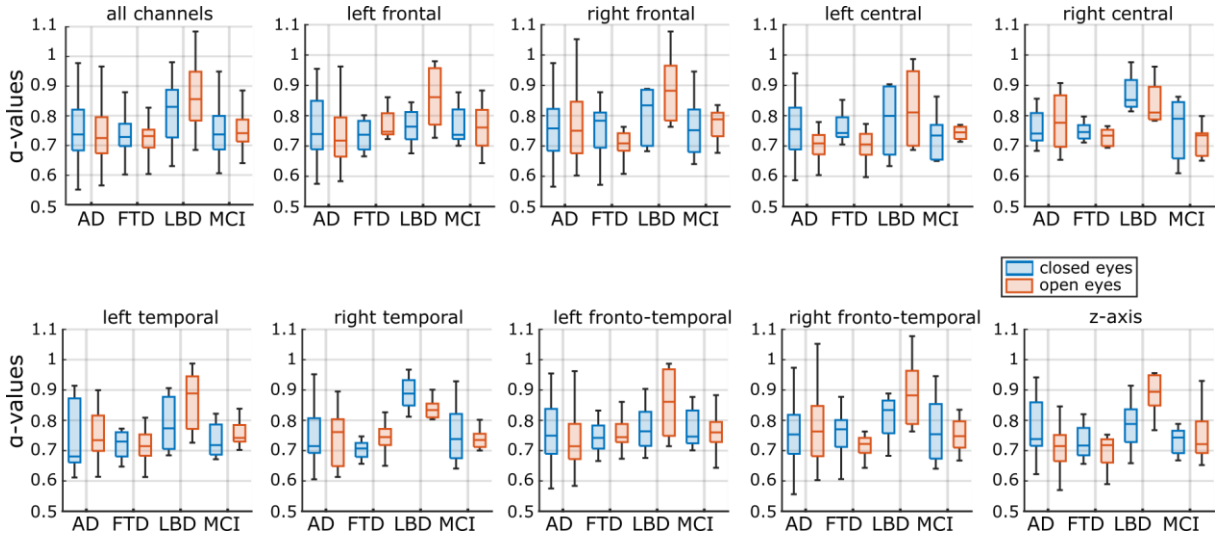


Figure 5.11 Boxchart comparison of detrended fluctuation analysis of EEG signals filtered between 30 and 40 Hz. The blue boxes represent the closed-eyes condition, while the peach ones the open-eyes condition. Tops and bottoms of the boxes are the quartiles, while the horizontal lines defined within each distribution correspond to the median.

The considerations that may be extracted from the outcomes shown in figure 5.10 are referred to an increasing trend for the LBD group compared to the other patients in the eyes-opened conditions for all the different subcategories of electrodes. The same trend is also found in the frequency band (30-40) Hz. In this latter case, higher α-values are visible in figure 5.11 not only in the eyes-opened condition for all the different subcategories, but also in the eyes-closed condition limitedly to the right central and right temporal subcategories of electrodes.

The choice of showing only figures related to the frequency band (8-12) and (30-40) Hz is simply due to the presence of interesting differences, even if not statistically evaluated among the different categories.

5.2.5 EEG: multiscale entropy analysis

The multiscale sample entropy estimates the sample entropy on several time scales. The procedure is applied both on pre-processed EEG time series both in open-eyes and closed-eyes conditions, filtered in the frequency band (2-45) Hz. The idea to apply the MSE algorithm on the filtered signals have been suggested from works present in literature [21], [22] and the choice to define the cut-off frequencies of the band pass filter equal to 2 and 45 Hz is related to the aspects described in section 5.2.3. The number of scale factors considered is 30, while the parameters set for the SampEn application are m=2 and r=0.2 times the standard deviation of the given signals. In this regard, figures 5.12 and 5.13 display the profiles of MSE curves obtained for the different subcategories of electrodes respectively in open-eyes and closed-eyes conditions.

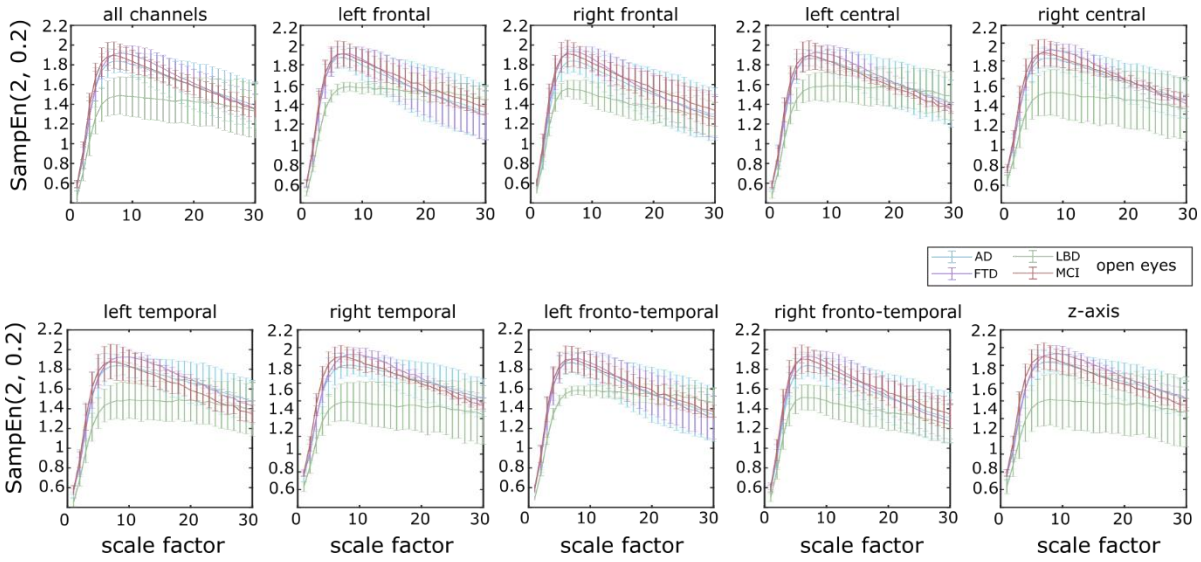


Figure 5.12 MSE curves computed from different subcategories of electrodes in open-eyes conditions. Blue curves represent AD, green LBD, purple FTD and red MCI. The y-axis indicates the SampEn values, that are shown in terms of mean and standard deviation in the curves, while the x-axis indicates the scale factors.

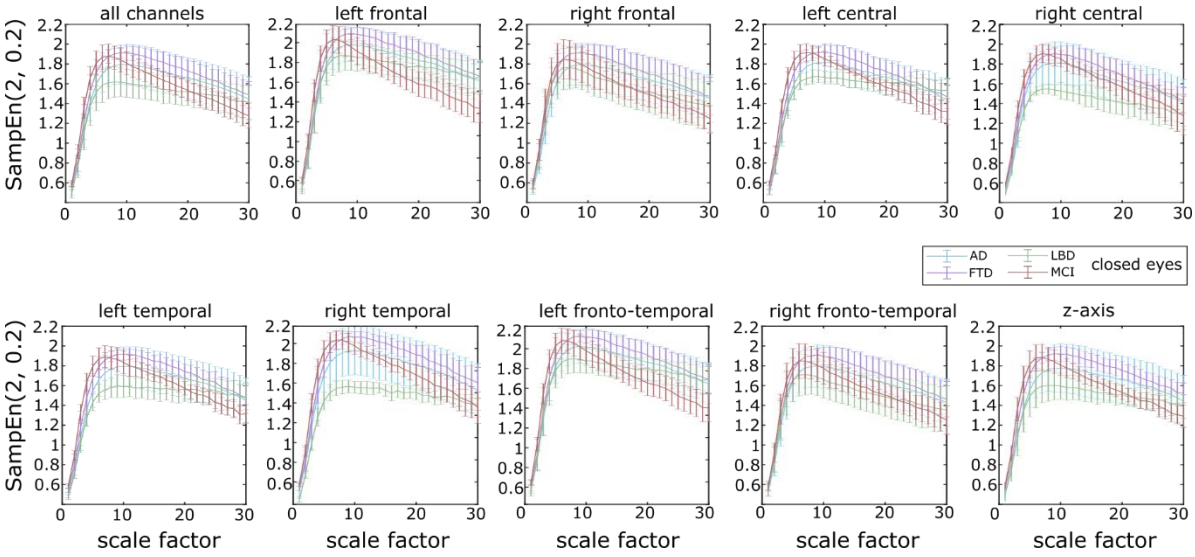


Figure 5.13 MSE curves computed from different subcategories of electrodes in closed-eyes conditions. Blue curves represent AD, green LBD, purple FTD and red MCI. The y-axis indicates the SampEn values, that are shown in terms of mean and standard deviation in the curves, while the x-axis indicates the scale factors.

From a visual examination of figure 5.12, at larger-scale factors, i.e. $\tau \geq 8$, the group of LBD patients appears characterized by entropy values approximately constant, while the other categories of diseases present values that decrease in function of the scale factor τ. This is true for almost all subcategories in the open-eyes conditions.

On the other hand, as regard the closed-eyes condition, all categories of diseases present again a decreasing trend at larger scale factors ($\tau \geq 8$) less pronounced for LBD and steeper for MCI category of patients (figure 5.13). The MSE curves are used to compare qualitatively the signal complexity of the time series [22], but to carry out the analysis from a quantitative point of view, changes of the entropic index can be evaluated extracting the slopes computed on MSE profiles. In this case, the slopes are evaluated both on the smaller time scales, i.e. $1 \leq \tau \leq 7$, and on the larger ones, i.e. $8 \leq \tau \leq 30$.

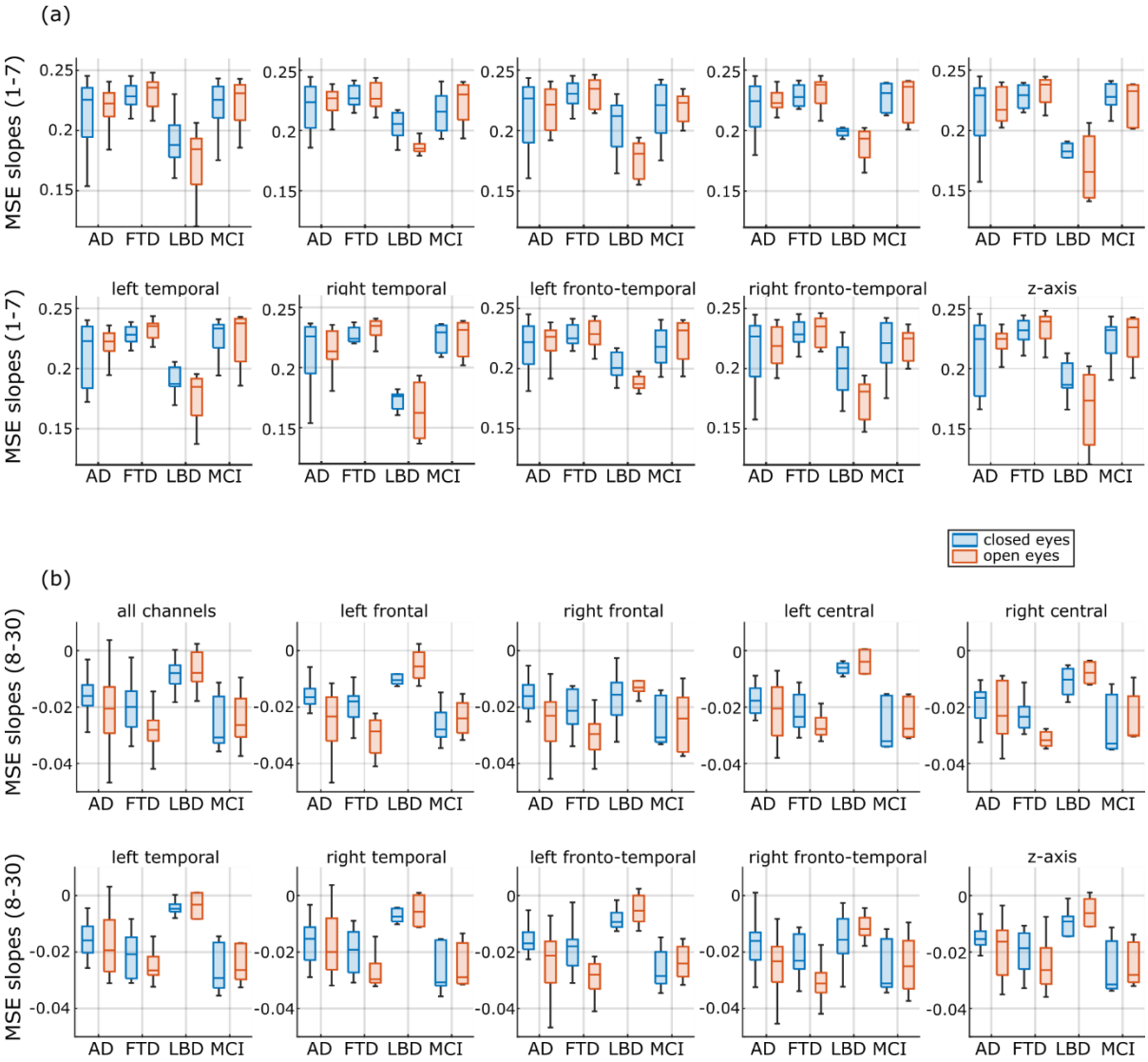


Figure 5.14 Boxchart of MSE slopes quantifying the grade of complexity in patients affected by four different neurological diseases. The blue boxes represent the closed-eyes condition, while the peach ones the open-eyes condition. Each box provides the distribution of MSE slope-values. Tops and bottoms of the boxes are the quartiles, while the horizontal lines defined within each distribution correspond to the median.

Figure 5.14 reveals higher values of MSE slopes (flatter slopes) evaluated at scale factors (8-30) for LBD group then the other patients in the eyes-closed condition (figure 5.14 b). This is consistent for all subcategories except for right frontal, right central, right fronto-temporal electrodes. Moreover, MSE slope evaluated at scale factors (8-30) shows lower values (steeper slopes) for MCI patients in the eyes-closed condition for all subcategories.

As regard the eyes-opened condition, at scale factors (1-7) (figure 5.14 a) lower MSE slopes (less steep slopes) are found for LBD group compared to the others for all the subcategories, while at scale factors (8-30) (figure 5.14 b) higher values for MSE slopes (flatter slopes) are shown for LBD compared to the other groups for all subcategories.

All considerations made on the outcomes interpreted from the figure 5.14 should be evaluated as trend due to the small number of patients involved in the study

6. Discussions

In this chapter, a discussion is provided about the results obtained by the investigation of the applicability criteria and the outcomes achieved in the reported case study on real datasets. Section 6.1 offers a detailed methodological discussion about the usefulness and importance of the simulation results, provided in chapter 3, about the proper setting of the parameters involved in the nonlinear algorithms proposed in this thesis. Section 6.2, instead, summarizes the results outlined in chapter 5 about the application of the nonlinear parameters to the dataset of real neurophysiological time series, i.e. LFP and EEG, in the context of Parkinson's and Alzheimer's diseases.

6.1 Methodological discussions

The aim of this thesis was to investigate the conditions of applicability of a set of selected nonlinear measures, providing guidelines for the correct setting of the parameters influencing their performance when applied to neurophysiological signals. There is a clear need to analyze the effects and interactions of changing input parameters, since both fractal and entropic methods appear very sensitive to them [42][49][50].

The first simulation study carried out in this thesis, tested the relationships (2.1) and (2.2) existing among the measures grouped under the section fractal, i.e. the power law exponent (β) and the detrended fluctuation analysis (α). Giving the fact that one measure can be theoretically derived by the others suggests the opportunity to verify the robustness of the fractal algorithms by comparing the direct estimations of β and α exponents with the values indirectly obtained (estimated) by applying the theoretical relations.

For sake of completeness, two comparisons were carried out: the first performed on 1000 fractional Brownian motions (fBm) and 1000 fractional Gaussian noises (fGn) time series generated by implementing a function that receives as input the hurst exponent (HE), while the second focused on similar artificial data but generated starting from known values of power law exponents (i.e. β). In both cases, as expected, tables 3.1 and 3.2 clearly showed that the measures estimated and the theoretically predicted ones were coherent. The congruence in the results finds confirmation in literature [27], [29] and allows to identify the more suitable parameter able to classify the time series as fBm or fGn, i.e. the β-exponent. The other simulation studies, instead, aimed at investigating the influence of the main factors that appear to play a key role in each nonlinear algorithm.

- Power law exponent

The way in which the frequency resolution affects the power spectrum density (PSD) needs to be evaluated before the application of the power-law exponent. The algorithm, in fact, exploits the inversely proportionality between the PSD and the frequency to measure changes in the scale-free dynamic properties of the time series. Thus, an incorrect estimation of PSD may invalidate the calculation of the β-slope of the regression line defined in log-log coordinates. An increase of the frequency

resolution led to a more stable evaluation of the β-exponent characterized by a reduced dispersion of the values measured across the 1000 surrogate time series. This is clearly shown in figure 3.3 and it may be explained considering the inverse proportionality between the frequency resolution and the window size M, exploited in the equation 3.2. M represents the length of segments in which the entire time series is divided into when applying the modified Welch's periodogram. This means that the larger M, the smaller the number of available segments, the smoother the PSD shape.

- Detrended fluctuation analysis

The minimum and the maximum window sizes *n* in which fluctuations F(n) are estimated to provide the measure of α-exponent supposed to be the influencing factors that characterize the DFA-method. The algorithm, in fact, defines a relationship between the variability of the signal and the length of the intervals over which the variability is computed to estimate the so-called long-range temporal correlation present in the signals. Thus, the appropriate choice of n_{min} and n_{max} and the way in which the different values of $n \in [n_{min}:n_{max}]$ are plotted against F(n) may affect the calculation of the α-slope of the regression line in log-log coordinates.

As regard the procedure to place the window size along the x axis, the most suitable solution appears to be the evenly spaced distribution [49], [50]; on the other hand, the selection of n_{min} and n_{max} was tested on the same two sets of 1000 realizations defined for the evaluation of the applicability criteria of the power law exponent. Interestingly, the accuracy of DFA in estimating the α-exponent appeared not influenced by the minimum window size but by a careful choice of maximum window size and data length combination. The results of the analysis carried out, in fact, revealed that for a fixed value of data length (N), lower values of n_{max} provide

unreliable estimates of the α-value with respect to higher values; on the other hand, instead, by considering an increase in data length, an opposite trend was found since more accurate estimations of DFA-exponent are guaranteed when lower values of n_{max} were selected. These findings are clearly depicted in figure 3.5 and may be explained by the fact that the number of intervals involved in the calculation of the α-slope depends on the length of the window sizes chosen. In other words, with an increase in data length, a larger number of intervals is required to guarantee the stability of the measure, and this corresponds to the selection of a lower value of n_{max}. If the time series are shorter, instead, a smaller number of intervals, that means higher value of n_{max}, is necessary to properly estimate the fluctuations.

- Entropy

Pincus and Goldberger (1991) and Richman and Moorman (2000) shows that when the embedding dimension (m) and the tolerance (r) are properly set, reliable approximate and sample entropy estimations may be provided [41], [43]. The algorithms, in fact, associates the conditional probability that two similar sequences of m points remain similar at the point m+1 to the grade of predictability in a time series. The expected trend of both entropic indices on Gaussian time series is a linear decrease with an increase of tolerance r. However, figure 3.8 clearly depicts how this is true for SampEn but not for ApEn, which assumes a distribution of values that increases up to a certain peak and then decreases. In addition to this, a closer inspection of the results showed how specific values of r allow to provide more accurate estimation of entropic indices. Such values are in the range between 0.1 and 0.25 times the standard deviation of the data and are valid both for ApEn and for SampEn for values of $N < 5000$. These findings did not change with different values of m (figure 3.9), where more unstable and unreliable entropic measures were found for embedding dimension $m \geq 3$.

As regard the multiscale entropy, the use of different temporal scales τ has no considerable effects on the results (figure 3.12). The tests on both white noise (β=0) and 1/f noise (β=1) time series showed that with an increase of the scale factor the shape of MSE profiles maintains the same trend. However, it is worth to consider that higher values of scale factor τ causes a lower number of samples in the coarse-grained sequence, thus the selection of the most suitable τ_{max} is necessary to avoid unreliable multiscale sample entropy estimations.

- Data length

Simulation studies were carried out to test the influence of the data length, a parameter that commonly affect all the selected nonlinear algorithms. The results, shown in figures 3.4, 3.7 and 3.11, revealed that with an increase in data length led to a better estimation of both fractal and entropic measures.

However, for sake of clarity, it is important to underline that the input parameters, previously described for each nonlinear method, are dependent on the data length. This means that changes of N values required further adjustment of n_{max} for DFA, of m and r for SampEn and ApEn and of τ_{max} for MSE.

6.2 Discussion of practical application

The knowledge about the behavior of the entropic and fractal measures, enriched with the simulation studies carried out in chapter 3, is crucial for correctly applying the algorithms to real datasets of neurophysiological signals. In this thesis, nonlinear measures were investigated on two types of neural recordings, i.e local field potential and electroencephalography.

- Local field potential (LFP) analysis in Parkinson's disease (PD)

The first case study included in the thesis was designed to assess the activity of the brain in the subthalamic region of Parkinson's disease patients in conditions of pre and post levodopa administration. The aim was to identify the influence of the antiparkinsonian medication in 24 PD subjects, analyzing spectral and dynamic nonlinear behavior of LFP time series.

From a linear point of view, the oscillatory patterns recorded from the subthalamic nucleus (STN) of patients at rest were quantified by the standard power spectral analysis. The results showed that the various LFP rhythms were differently affected by levodopa, since the normalized power in low-β (13-20 Hz) and low-γ (60-80 Hz) frequency bands was lower after levodopa medication, while the power in high-γ (250-350 Hz) increased after levodopa administration (figure 5.2). These differences were statistically evaluated in: i) a subset of patients whose LFP was recorded in conditions of both pre and post levodopa administration; ii) a set of nuclei of all PD patients, considered individually. In this regard, the decrease in low-β rhythm after drug medication was significant only in condition (i), while the decrease in low-γ after levodopa administration and the increase in high-γ rhythms were significant in both conditions (i) and (ii). These findings are summarized in table 5.1 and are in accordance with previous studies conducted on the same dataset under analysis [54].

From the fractal point of view, the analysis of LFP time series showed significant differences in both the proposed measures between the conditions of pre and post levodopa administration (see tables 5.2 and 5.3). In particular, as shown in figure 5.3, β-exponents appeared to be lower in condition of pre levodopa in both the frequency bands (2-45) and (2-156) Hz evaluated. Lower β-values reflected into less

steep PSD-slope were also found in the fronto-temporal sites of EEG by the recent study of Mostile et al. [62]. However, the different nature of the electrophysiological recordings may not provide a reliable comparison. Works based on LFP time series, instead, are mostly concentrated on the demonstration of the link between the excitation-inhibition balance and the nonlinear β-parameter. In this regard, flatter PSD slopes were related to an increasing excitation/inhibition ratio, that means an increase in inhibition and/or a decrease in excitation. (Gao).

On the other hand, figure 5.4 exhibited higher α-exponents, in terms of long-range temporal correlations, after the levodopa administration at high frequency range (250-350 Hz). This result finds confirmation in the work of Hohlefeld et al, 2012 that showed the presence of prominent LRTC in the high frequency (<200 Hz) when PD patients are treated with levodopa. As suggested by the authors, decreased LRTC are supposed to be associated with less efficient information processing and excessive neuronal hyperactivity in STN, while increased LRTC may be related to the optimal balance between stability and excitability. All these aspects highlight how LRTC were modulated by levodopa and lead to propose LRTC as a possible biomarker for PD [20].

From the entropic point of view, findings suggest that levodopa-related changes in the complexity properties of LFP are not so evident. In fact, the MSE proposed parameter showed no significant changes among the condition of pre and post levodopa administration (see table 5.4), but an interesting trend in correspondence of high scale factors ($\tau \geq 7$) seemed to reflect increases in SampEn(2, 0.2) values in condition of pre levodopa. This behavior was depicted in figure 5.5 and may be considered in line with works found in literature. Chung et al (2013) and Alam et al (2016), in fact, found a higher entropy in brain activity of PD patients with respect to control [63][64]. Thus, since higher entropy of the basal ganglia activity is supposed

to be associated to a more error-prone and less efficient motor information transfer [65], lower values of entropy were expected in conditions of post levodopa medication, due to the fact that antiparkinsonian medication improves the motor symptoms in PD, leading to consider the brain activity of patients to be closer to healthy control.

- Electroencephalography (EEG) analysis in Alzheimer's disease (AD)

The second case study of this thesis was designed to assess the activity of the brain in patients affected by four different neurological disorders, i.e. Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), Lewy Body dementia (LBD) and Frontotemporal dementia (FTD), in closed-eyes and open-eyes resting conditions. The aim was to verify if the use of parameters, evaluated on spectral and dynamic behavior of EEG time series, might point out possible distinctions among the groups of patients. Even if not statistically relevant, results revealed interesting differences particularly with regard to LBD and MCI against the other diseases. Despite this, the lack of studies aimed at discriminating the different disorders and the small sample size of the dataset under analysis, make it difficult to both understand the physiological outcomes and also to compare them with the results present in literature.

From a linear point of view, the oscillatory patterns recorded from the scalp of patients at rest were quantified by the standard power spectral analysis. The results showed that the various EEG rhythms were different for MCI patients with respect to the other diseases. Indeed, the normalized power in θ (2-7 Hz), α (8-12 Hz) and β (13-35 Hz) frequency bands revealed lower values for MCI in both eyes-closed and eyes-open conditions. These findings, exhibited in figure 5.8, are valid for all subcategories

of electrodes evaluated and are in accordance with previous studies conducted on EEG between MCI and AD[66]

From the fractal point of view, the analysis of EEG time series showed differences among AD, FTD, MCI and LBD in eyes-open and eyes-closed conditions in both the proposed measures (i.e. LRTC and power-law exponent). In particular, as regard the power-law exponent, experimental results revealed lower β-values for MCI in eyes-closed condition and higher β-values for LBD patients in eyes-open conditions, in comparison to the other groups. This is clearly visible in figure 5.9 and is valid for all the different subcategories of electrodes. In general, it has been demonstrated that decreased EEG complexity in AD patients with respect to healthy control was reflected by lower values of β-exponents (flatter PSD-slopes), which is associated to the presence of brain atrophy that affects the disease [15], but the lack of studies carried out on the other categories of disease make it difficult to infer the proper physiological interpretation of the measure in case of MCI, LBD and FTD.

On the other hand, as regard the DFA algorithm in terms of LRTC, results summarized in section 5.2.4 revealed higher α-exponents for LBD group of patients in eyes-open condition. This is clearly visible in figures 5.10 and 5.11 and is valid for all subcategories in case of both α (8-12 Hz) and low-γ (30-40 Hz) frequency bands. Moreover, a similar increasing trend was also found in eyes-closed condition, even if it was not confirmed in all the subcategories of electrodes under analysis. Indeed, the right central and right frontal are the only groups of electrodes in which the phenomenon is evident.

From the entropic point of view, all MSE profiles were characterized by trends that increased up to the maximum SampEn(2, 0.2) value. In line with literature, the maximum value is usually reached on smaller time scales [21], [22], and in the

present study such value was found both in eyes-close and eyes-open conditions for all the subcategories of electrodes. Regarding the behavior of the signals at larger scale factors ($\tau \geq 8$), the results suggested that in eyes-closed condition, LBD patients revealed values of SampEn(2, 0.2) approximately constant with increasing τ with respect to the other categories. MCI, AD and FTD, in fact, presented decreasing values of entropy with increasing τ. In eyes-open condition, all the diseases showed decreasing profiles at higher scale factors, even if less pronounced for LBD and also steeper values for MCI against the other two categories.

All the described findings may be visual examined in figures 5.12 and 5.13 in which MSE curves are displayed for each category of patients, and find confirmation in figure 5.14 in which slopes computed on profiles are evaluated in order to quantitatively compare the entropic measures.

7. Conclusions and future researches

Power-law exponent, detrended fluctuation analysis and multiscale entropy are the nonlinear techniques widely applied to quantify the information process capacity of the brain in Alzheimer's and Parkinson's disorders. However, the lack of application guidelines for the applicability of measures on neurophysiological signals and to the appropriate selection of the parameters, which highly influence the performance, impedes a considerable comparison between the different works in literature.

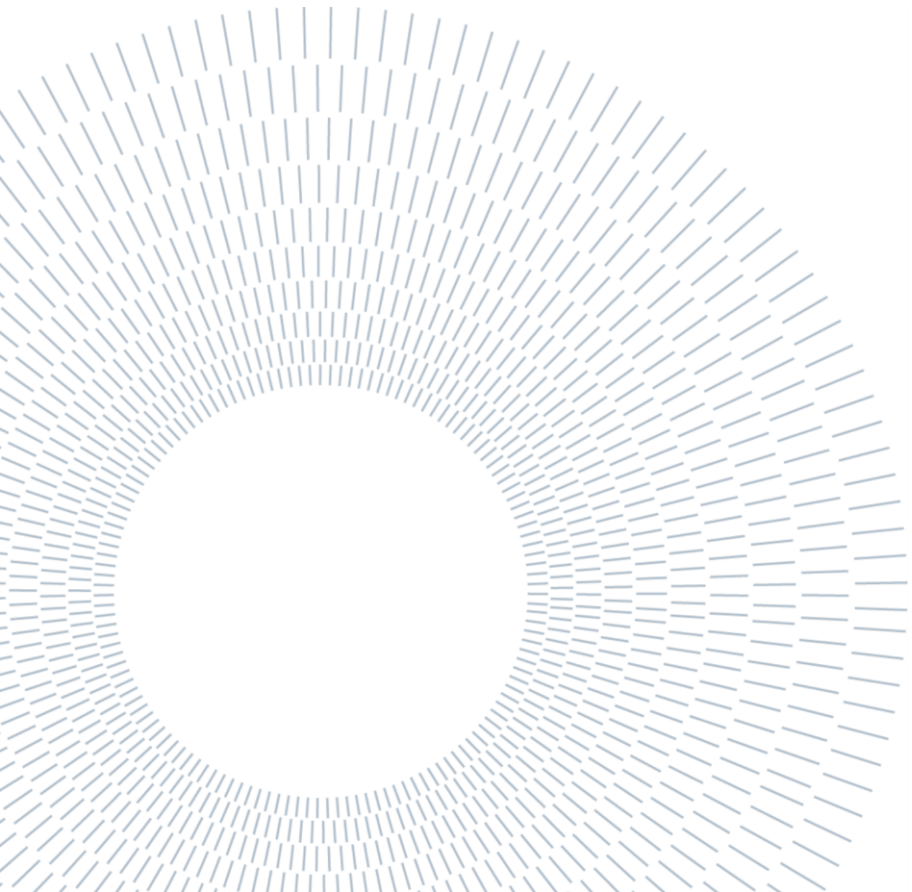
Even if interesting results were obtained in the literature on the application of such nonlinear measures on LFP and EEG, there is the suspect that the selection of input parameters for the algorithms was made empirically without a deep investigation of the applicability criteria in each specific case. Thus, the need of defining specific criteria for the appropriate parameter selection may increase the repeatability of the results along different signals and provide estimations in line with the rationale behind their definition.

Despite this study aims to highlight how selected nonlinear algorithms perform under different combinations of input parameters, there are limitations whose should be considered to improve the reliability of these methods for the future development of data-driven nonlinear approaches. The first obstacle is related to the ranges in which the input parameters were analyzed. These were selected based on preexisting

works on literature and were mainly applied on fractional Brownian motion and fractional Gaussian noise. Thus, future studies should investigate wider ranges of parametric inputs or either evaluate other time series, such as pink noise for fractal analysis or chirp signal for entropy analysis. In this way, additional insights for nonlinear analysis may be provided in order to improve the reliability of these methods and for defining reference rules for different kind of investigators.

The second limitation is a clear consequence of the previous one and regards the difficult interpretation of the results for clinical applications. The case studies presented in this thesis, in fact, must be taken in consideration as a test bench for the selected nonlinear measures. They have been introduced with the sole purpose to help understanding the basic implementation of the measures on real biological signals. Indeed, the reduced number of subjects involved and the limited in-depth studies in literature make difficult to highlight the great potential of the nonlinear techniques for application in neuroscience. Further investigations, thus, may lead to verify the grade of correlation between the nonlinear parameters and the motor impairment scores as regard the PD as well as the cognitive impairment scores in case of AD.

In conclusions, the present thesis investigated some of methodological issues on the basis of theoretical simulations. The obtained results provide concrete advices on the correct choice of nonlinear measures and parameters helping to minimize errors made by the utilizer. This set of nonlinear tools for studying the neurophysiological signals and applying, if possible, the measures alongside the classical linear ones, could also facilitate clinical use. The idea of developing a graphical interface, which makes the tool easy to use for clinicians, may represent, in fact, a great challenge for research, helping to better answer the still open questions on the clinical interpretation of neural data.



Bibliography

[1] M. B. Tayel and E. I. Alsaba, "Review: Nonlinear Techniques for Analysis of Heart Rate Variability," 2016. [Online]. Available: www.ijres.org

[2] H. Hallez *et al.*, "Review on solving the forward problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 4. 2007. doi: 10.1186/1743-0003-4-46.

[3] B. Pesaran *et al.*, "Investigating large-scale brain dynamics using field potential recordings: Analysis and interpretation," *Nature Neuroscience*, vol. 21, no. 7, pp. 903–919, Jun. 2018, doi: 10.1038/s41593-018-0171-8.

[4] S. P. Burns, D. Xing, and R. M. Shapley, "Comparisons of the dynamics of local field potential and multiunit activity signals in macaque visual cortex," *Journal of Neuroscience*, vol. 30, no. 41, pp. 13739–13749, Oct. 2010, doi: 10.1523/JNEUROSCI.0743-10.2010.

[5] A. Moran and I. Bar-Gad, "Revealing neuronal functional organization through the relation between multi-scale oscillatory extracellular signals," *Journal of Neuroscience Methods*, vol. 186, no. 1, pp. 116–129, Jan. 2010, doi: 10.1016/j.jneumeth.2009.10.024.

[6] S. Katzner, I. Nauhaus, A. Benucci, V. Bonin, D. L. Ringach, and M. Carandini, "Local Origin of Field Potentials in Visual Cortex," *Neuron*, vol. 61, no. 1, pp. 35–41, Jan. 2009, doi: 10.1016/j.neuron.2008.11.016.

[7] U. Herwig, P. Satrapi, and C. Schönfeldt-Lecuona, "Using the International 10-20 EEG System for Positioning of Transcranial Magnetic Stimulation," 2003. [Online]. Available: www.easycap.de

[8] W. Poewe *et al.*, "Parkinson disease," *Nature Reviews Disease Primers*, vol. 3, no. 1, p. 17013, Dec. 2017, doi: 10.1038/nrdp.2017.13.

[9] H. Bergman and G. Deuschl, "Pathophysiology of Parkinson's disease: From clinical neurology to basic neuroscience and back," *Movement Disorders*, vol. 17, no. S3, pp. S28–S40, Mar. 2002, doi: 10.1002/mds.10140.

[10] J. Jankovic, "Current approaches to the treatment of Parkinson’s disease," *Neuropsychiatric Disease and Treatment*, p. 743, Sep. 2008, doi: 10.2147/NDT.S2006.

[11] M. Rosa, G. Giannicola, S. Marceglia, M. Fumagalli, S. Barbieri, and A. Priori, "Neurophysiology of Deep Brain Stimulation," 2012, pp. 23–55. doi: 10.1016/B978-0-12-404706-8.00004-8.

[12] M. Revi, "Alzheimer's Disease Therapeutic Approaches," 2020, pp. 105–116. doi: 10.1007/978-3-030-32633-3_15.

[13] D. Abásolo, R. Hornero, J. Escudero, and P. Espino, "A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in Alzheimer's disease," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2171–2179, Sep. 2008, doi: 10.1109/TBME.2008.923145.

[14] L. Chu, "Alzheimer's disease: early diagnosis and treatment," 2012. [Online]. Available: www.hkmj.org

[15] O. Vyšata *et al.*, "Change in the characteristics of EEG color noise in alzheimer's disease," *Clinical EEG and Neuroscience*, vol. 45, no. 3, pp. 147–151, 2014, doi: 10.1177/1550059413491558.

[16] E. M. Belova, U. Semenova, A. A. Gamaleya, A. A. Tomskiy, and A. Sedov, "Voluntary movements cause beta oscillations increase and broadband slope decrease in the subthalamic nucleus of parkinsonian patients," *European Journal of Neuroscience*, vol. 53, no. 7, pp. 2205–2213, Apr. 2021, doi: 10.1111/ejn.14715.

[17] R. Gao, E. J. Peterson, and B. Voytek, "Inferring synaptic excitation/inhibition balance from field potentials," *NeuroImage*, vol. 158, pp. 70–78, Sep. 2017, doi: 10.1016/j.neuroimage.2017.06.078.

[18] Y. Huang *et al.*, "Dynamic changes in rhythmic and arrhythmic neural signatures in the subthalamic nucleus induced by anaesthesia and tracheal intubation," *British Journal of Anaesthesia*, vol. 125, no. 1, pp. 67–76, Jul. 2020, doi: 10.1016/j.bja.2020.03.014.

[19] C. J. Stam *et al.*, "Disturbed fluctuations of resting state EEG synchronization in Alzheimer's disease," *Clinical Neurophysiology*, vol. 116, no. 3, pp. 708–715, 2005, doi: 10.1016/j.clinph.2004.09.022.

[20] F. U. Hohlefeld *et al.*, "Long-range temporal correlations in the subthalamic nucleus of patients with Parkinson's disease," *European Journal of Neuroscience*, vol. 36, no. 6, pp. 2812–2821, Sep. 2012, doi: 10.1111/j.1460-9568.2012.08198.x.

[21] T. Mizuno *et al.*, "Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy," *Clinical Neurophysiology*, vol. 121, no. 9, pp. 1438–1446, Sep. 2010, doi: 10.1016/j.clinph.2010.03.025.

[22] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, Nov. 2006, doi: 10.1088/0967-3334/27/11/004.

[23] J. Syrkin-Nikolau *et al.*, "Subthalamic neural entropy is a feature of freezing of gait in freely moving people with Parkinson's disease," *Neurobiology of Disease*, vol. 108, pp. 288–297, Dec. 2017, doi: 10.1016/j.nbd.2017.09.002.

[24] A. L. Goldberger, "Complex systems," in *Proceedings of the American Thoracic Society*, 2006, vol. 3, no. 6, pp. 467–471. doi: 10.1513/pats.200603-028MS.

[25] P. Herman, L. Kocsis, and A. Eke, "Fractal characterization of complexity in dynamic signals: Application to cerebral hemodynamics," *Methods in Molecular Biology*, vol. 489. Humana Press, pp. 23–40, 2009. doi: 10.1007/978-1-59745-543-5_2.

[26] A. Eke, P. Herman, L. Kocsis, and L. R. Kozak, "Fractal characterization of complexity in temporal physiological signals," 2002.

[27] D. Delignieres, S. Ramdani, L. Lemoine, K. Torre, M. Fortes, and G. Ninot, "Fractal analyses for 'short' time series: A re-assessment of classical methods," *Journal of Mathematical Psychology*, vol. 50, no. 6, pp. 525–544, Dec. 2006, doi: 10.1016/j.jmp.2006.07.004.

[28] H. Hurst, "Long Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.

[29] A. Eke *et al.*, "Physiological time series: distinguishing fractal noises from motions," *Pflügers Archiv - European Journal of Physiology*, vol. 439, no. 4, pp. 403–415, Feb. 2000, doi: 10.1007/s004249900135.

[30] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, Oct. 1968, doi: 10.1137/1010093.

[31] F. J. Molz, H. H. Liu, and J. Szulga, "Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions," *Water Resources Research*, vol. 33, no. 10, pp. 2273–2286, 1997, doi: 10.1029/97WR01982.

[32] T. Donoghue *et al.*, "Parameterizing neural power spectra into periodic and aperiodic components," *Nature Neuroscience*, vol. 23, no. 12, pp. 1655–1665, Dec. 2020, doi: 10.1038/s41593-020-00744-x.

[33] G. Perez, B. Rappazzo, and C. Gomes, "Extending the capacity of 1/f noise generation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11008 LNCS, pp. 601–610. doi: 10.1007/978-3-319-98334-9_39.

[34] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Aver. aging Over Short, Modified Periodograms Ik(fn) =-I A h (%) [a k," 1967.

[35] M. A. Colombo *et al.*, "The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine," *NeuroImage*, vol. 189, pp. 631–644, Apr. 2019, doi: 10.1016/j.neuroimage.2019.01.024.

[36] C.-K. Peng, S. v Buldyrev, S. Havtin, M. Simons, H. E. Stanley, and A. L. Goldbergerz, "Mosaic organization of DNA nucleotides," 1994.

[37] Y. H. Shao, G. F. Gu, Z. Q. Jiang, W. X. Zhou, and D. Sornette, "Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series," *Scientific Reports*, vol. 2, 2012, doi: 10.1038/srep00835.

[38] S. Keshmiri, "Entropy and the Brain: An Overview," *Entropy*, vol. 22, no. 9, p. 917, Aug. 2020, doi: 10.3390/e22090917.

[39] T. Takahashi, "Complexity of spontaneous brain activity in mental disorders," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45. pp. 258–266, Aug. 01, 2013. doi: 10.1016/j.pnpbp.2012.05.001.

[40] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.

[41] S. M. Pincus, A. L. Goldberger, and A. L. Goldberger Physiologi-, "Physiological time-series analysis: what does regularity quantify?"

[42] X. Chen, I. C. Solomon, and K. H. Chon, "Comparison of the use of approximate entropy and sample entropy: Applications to neural respiratory signal," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2005, vol. 7 VOLS, pp. 4212–4215. doi: 10.1109/iembs.2005.1615393.

[43] J. S. Richman, J. Randall Moorman, J. Randall, and M. Physi, "Downloaded from journals.physiology.org/journal/ajpheart (131.175.028.198) on," 2000.

[44] A. Delgado-Bonal and A. Marshak, "Approximate entropy and sample entropy: A comprehensive tutorial," *Entropy*, vol. 21, no. 6. MDPI AG, Jun. 01, 2019. doi: 10.3390/e21060541.

[45] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Physical Review Letters*, vol. 89, no. 6, 2002, doi: 10.1103/PhysRevLett.89.068102.

[46] M. Costa, ¡£¢, and J. A. Healey, "Multiscale Entropy Analysis of Complex Heart Rate Dynamics: Discrimination of Age and Heart Failure Effects," 2003. [Online]. Available: http://physionet.org

[47] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 71, no. 2, Feb. 2005, doi: 10.1103/PhysRevE.71.021906.

[48] T. Stadnitski, "Some Critical Aspects of Fractality Research," 2012. [Online]. Available: https://www.researchgate.net/publication/221977930

[49] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao, "A tutorial introduction to adaptive fractal analysis," *Frontiers in Physiology*, vol. 3 SEP, 2012, doi: 10.3389/fphys.2012.00371.

[50] Z. M. H. Almurad and D. Delignières, "Evenly spacing in Detrended Fluctuation Analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 451, pp. 63–69, Jun. 2016, doi: 10.1016/j.physa.2015.12.155.

[51] Z. Chen, P. C. Ivanov, K. Hu, and H. E. Stanley, "Effect of nonstationarities on detrended fluctuation analysis," *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 65, no. 4, p. 15, 2002, doi: 10.1103/PhysRevE.65.041107.

[52] L. Zhao *et al.*, "Determination of Sample Entropy and Fuzzy Measure Entropy Parameters for Distinguishing Congestive Heart Failure from Normal Sinus Rhythm Subjects," *Entropy*, vol. 17, no. 12, pp. 6270–6288, Sep. 2015, doi: 10.3390/e17096270.

[53] H. Azami and J. Escudero, "Coarse-graining approaches in univariate multiscale sample and dispersion entropy," *Entropy*, vol. 20, no. 2, Feb. 2018, doi: 10.3390/e20020138.

[54] S. Marceglia *et al.*, "Gender-related differences in the human subthalamic area: A local field potential study," *European Journal of Neuroscience*, vol. 24, no. 11, pp. 3213–3222, Dec. 2006, doi: 10.1111/j.1460-9568.2006.05208.x.

[55] G. Foffani, "300-Hz subthalamic oscillations in Parkinson's disease," *Brain*, vol. 126, no. 10, pp. 2153–2163, Jun. 2003, doi: 10.1093/brain/awg229.

[56] A. PRIORI *et al.*, "Rhythm-specific pharmacological modulation of subthalamic activity in Parkinson's disease," *Experimental Neurology*, vol. 189, no. 2, pp. 369–379, Oct. 2004, doi: 10.1016/j.expneurol.2004.06.001.

[57] J. P. Amezquita-Sanchez, N. Mammone, F. C. Morabito, S. Marino, and H. Adeli, "A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals," *Journal of Neuroscience Methods*, vol. 322, pp. 88–95, Jul. 2019, doi: 10.1016/j.jneumeth.2019.04.013.

[58] R. Nardone *et al.*, "Usefulness of EEG Techniques in Distinguishing Frontotemporal Dementia from Alzheimer's Disease and Other Dementias," *Disease Markers*, vol. 2018. Hindawi Limited, 2018. doi: 10.1155/2018/6581490.

[59] J. J. Young, M. Lavakumar, D. Tampi, S. Balachandran, and R. R. Tampi, "Frontotemporal dementia: latest evidence and clinical implications," *Therapeutic Advances in Psychopharmacology*, vol. 8, no. 1, pp. 33–48, Jan. 2018, doi: 10.1177/2045125317739818.

[60] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state,'" *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, Nov. 1975, doi: 10.1016/0022-3956(75)90026-6.

[61] J. Iriarte *et al.*, "Independent Component Analysis as a Tool to Eliminate Artifacts in EEG: A Quantitative Study," 2003.

[62] G. Mostile *et al.*, "Complexity of electrocortical activity as potential biomarker in untreated Parkinson's disease," *Journal of Neural Transmission*, vol. 126, no. 2, pp. 167–172, Feb. 2019, doi: 10.1007/s00702-018-1961-6.

[63] C.-C. Chung *et al.*, "Multiscale Entropy Analysis of Electroencephalography During Sleep in Patients With Parkinson Disease," *Clinical EEG and Neuroscience*, vol. 44, no. 3, pp. 221–226, Jul. 2013, doi: 10.1177/1550059412475066.

[64] M. Alam *et al.*, "Globus pallidus internus neuronal activity: a comparative study of linear and non-linear features in patients with dystonia or Parkinson's disease," *Journal of Neural Transmission*, vol. 123, no. 3, pp. 231–240, Mar. 2016, doi: 10.1007/s00702-015-1484-3.

[65] A. D. Dorval and W. M. Grill, "Deep brain stimulation of the subthalamic nucleus reestablishes neuronal information transmission in the 6-OHDA rat model of parkinsonism," *J Neurophysiol*, vol. 111, pp. 1949–1959, 2014, doi: 10.1152/jn.00713.2013.-Pathophysiologi.

[66] J. C. McBride *et al.*, "Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 2, pp. 153–163, 2014, doi: 10.1016/j.cmpb.2014.01.019.

List of Figures

List of Tables

Acronyms

N	data length
fs	sampling frequency
LFP	local field potential
EEG	electroencephalography
PLE	power-law exponent
β	power-law exponent
PSD	power spectrum density
Δf	frequency resolution
DFA	detrended fluctuation analysis
α	detrended fluctuation analysis exponent
n_{min}	minimum window size
n_{max}	maximum window size
ApEn	approximate entropy
SampEn	sample entropy
MSE	multiscale entropy
m	embedding dimension
r	tolerance
τ	scale factor
PD	Parkinson's disease

STN	subthalamic nuclei
L-DOPA	levodopa
DBS	Deep brain stimulation
UPDRS	Unified Parkinson's Disease Rating Scale
AD	Alzheimer’s disease
MMSE	Mini Mental State Examination
MCI	Mild Cognitive Impairment
FTD	Frontotemporal dementia
LBD	Lewy body dementia

Acknowledgements

First of all, I would like to express my sincerest gratitude to the Professor Anna Maria Bianchi for her guidance and patience. I feel grateful to have worked under her supervision.

I want also to thank biomedical engineers Stefania Coelli and Alberto Averna for their collaboration and their willingness to meet, discuss and advise me for every step of my work in this thesis. It would be impossible to carry out it without their help and constant support.

I wish also to express my acknowledgment to the Professor Alberto Priori for granting me the possibility to attend a curricular stage in the Dipartimento di Scienze della Salute (DISS), Università degli Studi di Milano.

Last but not least, my special thanks to all my colleagues for giving me happiness and joy during the last three years and especially to my family for all their endless care, support and encouragement throughout my exciting graduate career.