Scaling properties and heterogeneous dynamics of epileptic activity

Lucas Gabriel Souza França

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Clinical and Experimental Epilepsy

UCL Queen Square Institute of Neurology

University College London

06 May 2020

I, Lucas Gabriel Souza França, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Epilepsy is a brain disorder characterised by recurrent seizures with a wide range of cognitive, psychological, and social consequences. Seizures are the transient occurrence of signs and symptoms due to abnormal excessive synchronous neural activity. Epilepsy is currently treated with medication or by a resective surgical procedure; however, about 30% of patients still do not achieve seizure control. The unpredictability of when seizures occur means that even infrequent seizures can have a devastating effect on someone's life and has hampered progress in response mode treatments.

There has, therefore, been a major research drive for the development of seizure prediction/detection techniques. These have suffered from poor sensitivity and/or specificity. Moreover, many of the advanced mathematical methods for seizure detection/prediction do no more than reproduce the results of linear and simpler approaches.

This thesis focuses on the scaling properties of intracranial electrophysiological measures from both humans and rats. Firstly, a novel approach to evaluate multifractal properties of brain signals is presented and its properties are demonstrated. Secondly, the developed approach is applied in a series of clinical challenges: dis-

tinguishing sleep phases; a study on characteristic time for seizure detection; a putative classification scheme for focal epileptic seizures; and an evaluation of data from an animal model of epileptogenesis.

The proposed technique was capable of extracting distinct information from the EEG signals that has clinical utility. Moreover, the multifractal nature of the brain signals hints at how neural structures work and lends support to results from previous research.

Impact Statement

Epilepsy is a disease that affects over 70 million people around the globe with around 30% of these patients not being able to achieve freedom from seizures — the main symptom of the syndrome — despite optimum treatment. One of the most disabling aspects of seizures is their unpredictability.

This thesis investigates properties of the electrical activity of epileptic seizures, aiming to determine methods of detecting and predicting seizures that could help the development of new 'response-mode' treatments. Moreover, a better understanding of seizure dynamics could provide insights into novel treatment approaches. Thus, this work could have an impact on the treatment and care of people with epilepsy.

Discoveries arising from the studies presented in this thesis could impact academia by bringing new information on mechanisms and dynamics of the brain — in both health and disease. This way it could change concepts in neurology and foster the development of new research ideas, projects and questions.

These results would have a global impact in both academia and healthcare, changing the lives of millions of people around the world. One article resulting from the working developed during the studies presented in this thesis has already been published in an international peer-reviewed journal.

Acknowledgements

This section is aimed to acknowledge those whose direct or indirect actions made this work possible. These are randomly listed below (excepting the supervisors of this work — who head the list).

Prof. Matthew Walker (supervisor), Dr Yujiang Wang (supervisor), Maria Centeno, Jacopo Fantini, Bruna Montuori, André Van Graan, Fernanda Queirós, Jane de Tisi, Vejay Vakharia, João Cardoso, Vanissa Wanick, Michael Everson, José Garcia Vivas Miranda, Sjoerd Vos, Marian Galovic, Igor Dutra, Simon Bennett, Peter Taylor, Niraj Sharma, Frank Mulhern, Tim Wehner, Charles Santana, Roderick Ritchie, Lorenzo Caciagli, Prof. John Duncan, Helena Martins, Caroline Osbourn, Armando Castro, Tabada Dutra, Prof. Malcolm von Schantz, Ann-Marie Douglas, Mayra Cotta, Alessandra Marcon Gasperini, Tawfeeq Shekh-Ahmad, Andrea Hill, Isis Costa, Matthew Osbourn, Rafael Magalhães, Jane Burdett, Juliet Solomon, Denise Neves Santos, David Warner, Luke Allen, Paul Bray, Anderson Brito, Marco Leite, David Osbourn.

This work was supported by a grant from the Brazilian National Council for Scientific and Technological Development (CNPq) (206907/2014-1).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 On epilepsy

Epilepsy is a brain disorder characterised by recurrent seizures with a wide range of cognitive, psychological and social consequences. Seizures are defined as the paroxysmal (transient) occurrence of signs and symptoms due to abnormal, excessive synchronous neural activity (Fisher et al., 2005). The existence of an enduring alteration, in the brain, that increases the likelihood of seizures (epileptogenicity), is the key concept in the definition of epilepsy (Fisher et al., 2005).

Epilepsy is believed to affect over 70 million people all over the world (Thijs et al., 2019), with a higher impact in under-resourced countries (Meinardi et al., 2008; Ngugi et al., 2010). This disparity can be explained by the lack of effective healthcare systems and eventual "higher risk of brain damage due to parasitic (e.g., cysticercosis) and bacterial (e.g., tuberculosis, meningococci) infections, and due to substandard perinatal care" (Meinardi et al., 2008).

In wealthy countries the incidence of epilepsy is approximately 50/100,000

cases per year, whereas in resource-poor countries the incidence is two to threefold higher (Sander, 2003). The prevalence is approximately 0.6-1 % (Sander, 2003).

The epilepsies are classified by the International League Against Epilepsy (ILAE) as either generalised, which "are conceptualised as originating at some point within, and rapidly engaging, bilaterally distributed networks", or focal, "originating within networks limited to one hemisphere" (Berg et al., 2010).

1.1.1 Classifying epilepsies and epileptic seizures

Classifying neurological syndromes as diverse as the epilepsies, and their clinical manifestations epileptic seizures, can bring some challenges. Proposals of international classifications were made in late 1960s and published in 1970 (Gastaut, 1970; Merlis, 1970; Berg and Cross, 2010). The initial classification scheme included clinical features of epileptic seizures (Berg and Cross, 2010) and mainly divided epileptic seizures in two groups: of those "that are generalised from the beginning", and "those that are partial or focal at onset and become generalised secondarily" (Penry, 1981).

A proposal of revision of such classification was published in 1981; it encompasses knowledge obtained from — at the time — novel clinical technologies for documenting epileptic seizures, e.g. prolonged EEG and video recordings. Named International Classification of Epileptic Seizures (ICES), the revision also removed historical information or data considered speculative (such as aetiology), age factors, and anatomic substrate (Penry, 1981). It is important to emphasise that this was still a seizure classification and kept some of the elements from the 1970 scheme.

The ILAE proposed another classification in 1989, called International Classification of Epilepsies and Epileptic Syndromes (ICE). In this new scheme the syndromes rather than the seizures - expressions of those syndromes - are classified. The new proposal still maintained the main concept of previous work with the two major classes being defined as epilepsies with generalised or partial/focal seizures, respectively. Aetiological criteria were also adopted by the 1989 proposal with epilepsies being classed as symptomatic or "secondary" (of known aetiology), idiopathic (primary), and cryptogenic (Commission on Classification and Terminology of the International League Against Epilepsy, 1989). In 2010, the ILAE Commission on Classification and Terminology published a document reviewing the "concepts, terminology, and approaches for classifying seizures and forms of epilepsy" (Berg et al., 2010). This new classification departs from both 1981's and 1989's ICES and ICE.

The revised terminology separates epileptic seizures into two groups: generalised and focal. Also, the aetiologies of the syndromes were redefined as genetic, structural-metabolic, and unknown, to replace idiopathic, symptomatic, and cryptogenic, respectively (Berg et al., 2010). It is important to emphasise that some seizures, and epilepsies, may fall into subtypes and/or exhibit special features.

1.1.2 The epileptic seizure

Overall, epileptic seizures are the symptoms caused by "abnormal excessive or synchronous neuronal activity in the brain" (Fisher et al., 2005). An epileptic seizure is a transient event with a clear start and finish; however, the termination is often less

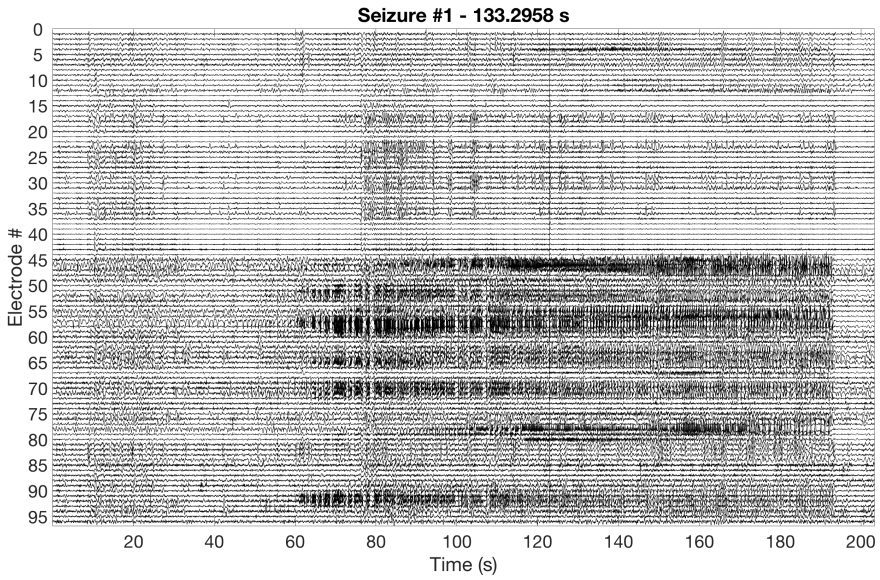

Figure 1.1: Example of seizure electroencephalography obtained with intracranial electrodes in a patient undergoing presurgical evaluation. The seizure starts at 60 seconds and lasts for about 133 seconds.

evident than its onset — this is due to the fact that the postictal state can at times blur the termination of a seizure (Fisher et al., 2005).

There is a wide range of clinical manifestations that can arise from epileptic seizures; these signs will depend on location of the onset in the brain, the propagation, medications and various other factors (Fisher et al., 2005). "Sensory, motor, and autonomic function; consciousness; emotional state; memory; cognition; or behaviour" can be affected by epileptic seizures (Fisher et al., 2005). Nevertheless, not all seizures compromise all these functions and capacities but at least one of these will be affected (Fisher et al., 2005).

This section follows the terminology proposed in 2010 by the ILAE commission (Berg et al., 2010) and describes the main seizure types classed as either generalised or focal. These seizures may also exhibit further subtypes and special features not discussed in this text.

1.1.2.1 Generalised seizures

Generalised seizures are defined "as originating at some point within, and rapidly engaging, bilaterally distributed networks" that can involve cortical and subcortical structures (Berg et al., 2010). The generalised seizures can be classified in 6 main subgroups: tonic-clonic, absence, myoclonic, clonic, tonic, and atonic seizures (Berg et al., 2010).

Tonic-clonic seizures are characterised by a sudden attack and loss of consciousness without warning or auras (Greenberg et al., 2012). This class of seizure is defined by three phases: tonic, clonic, and recovery stages. In rare cases, they may evolve into convulsive status epilepticus (prolonged seizures, defined below in section 1.1.2.3) (Greenberg et al., 2012).

The initial tonic stage, with unconscious tonic muscular contractions, lasts between 10 and 30 seconds. In this type of seizure, the patient falls to the ground and can be injured. Contraction of respiratory muscles may cause vocalisation, resulting in a "cry" or "moan", and cyanosis. Some patients may experience tongue trauma due to contraction of masticatory muscles (Greenberg et al., 2012).

The second phase is characterised by "a clonic (alternating muscle contraction and relaxation) phase of symmetric limb jerking" and persists for 30-60 seconds or more (Greenberg et al., 2012). This phase is also marked by cessation of cyanosis, mouth froth with saliva, and sometimes urinary incontinence (Greenberg et al., 2012).

The third phase is marked by postictal confusion and often headache. Full orientation should take 10-30 minutes but can take longer in cases with status epilep-

ticus (Greenberg et al., 2012).

Absence seizures are characterised by a brief loss of consciousness that lasts from 5 to 10 seconds. These seizures may also feature subtle motor manifestations, e.g. eye blink and head turning. Consciousness recovers immediately after the cessation of the seizure (Greenberg et al., 2012).

Myoclonic seizures exhibit sudden and brief contractions localised to some muscles, to one or more extremities, or distributed more generally (Greenberg et al., 2012).

Clonic seizures show similar features to the second phase of tonic-clonic seizures without the initial tonic phase, whereas the tonic seizures exhibit properties of the first phase of tonic-clonic seizures without the subsequent clonic phase (Greenberg et al., 2012). Atonic seizures feature loss of postural tone leading to a fall or drop attack and may be preceded by a myoclonic jerk (Greenberg et al., 2012).

1.1.2.2 Focal seizures

Focal seizures are defined "as originating within networks limited to one hemisphere". These networks can be more localised or widely spread and can also have their origin in subcortical structures. Focal seizures are not grouped into sets of natural classes, but are defined by loss, or not, of awareness, and the presence of clinical features depending upon origin and spread (Berg et al., 2010).

It is also vital to underline that, in focal epilepsies, seizures are often a clinical expression of a disease process; in other words, ongoing structural and functional

changes following a brain insult or lesion can be expressed clinically as seizures in susceptible individuals (Walker et al., 2002).

1.1.2.3 Status epilepticus

Status epilepticus is defined by the ILAE in the 1981 ICE as when "a seizure persists for a sufficient length of time or is repeated frequently enough that recovery between attacks does not occur". This official definition however is deemed imprecise as it lacks a specific duration of seizure activity (Lowenstein et al., 1999). The condition is thought to be caused by the failure of mechanisms that would normally abort epileptic seizures (Lowenstein and Alldredge, 1998).

Attempts have been made to quantify the maximum duration of an epileptic seizure; these efforts were based on the time it takes for the episode to cause neuronal damage in animal models (Lowenstein and Alldredge, 1998; Lowenstein et al., 1999). However, the studies suggest different values for a putative maximum duration (status epilepticus inclusion criterion), e.g., 30 (Celesia, 1976; Brodie, 1990), 20 (Bleck, 1991), and 10 minutes (Treiman et al., 1998). Another problem with such criteria is their incompatibility with clinical practice, in which assistance would be given much sooner (Lowenstein et al., 1999).

To try to solve this inconsistency Lowenstein et al. (1999), established an operational definition of status epilepticus "in adults and older children ($>$ 5 years old)" as "5 minutes of (a) continuous seizures or (b) two or more discrete seizures between which there is incomplete recovery of consciousness". A further mechanistic definition states that "convulsive status epilepticus refers to a condition in which

there is a failure of the 'normal' factors that serve to terminate a typical generalised tonic-clonic seizure" (Lowenstein et al., 1999).

In a recent effort, an ILAE commission merged the two definitions given by Lowenstein et al. (1999). The new definition says "status epilepticus is a condition resulting either from the failure of the mechanisms responsible for seizure termination or from the initiation of mechanisms which lead to abnormally prolonged seizures (after time point t_1). It is a condition that can have long-term consequences (after time point t_2), including neuronal death, neuronal injury, and alteration of neuronal networks, depending on the type and duration of seizures" (Trinka et al., 2015). The values of t_1 and t_2 were defined as 5 and 30 minutes for tonic-clonic status epilepticus, 10 and $>$ 60 minutes for focal status epilepticus with impaired consciousness, and 10-15 minutes, and unknown for absence status epilepticus (Trinka et al., 2015).

Status epilepticus is an extremely serious and life-threatening condition (Brodie, 1990) that brings serious risks to the health of the patient, such as permanent brain damage from hyperpyrexia, circulatory collapse, and/or excitotoxic neuronal damage (Greenberg et al., 2012).

1.1.3 Aetiologies

Epilepsy is a condition with multiple potential causes. The ILAE 2010 terminology revision defined three potential types of cause (aetiology) for epileptic syndromes: genetic, structural/metabolic, and unknown cause (Berg et al., 2010).

Genetic epilepsies are defined as "the direct result of a known or presumed

genetic defect(s) in which seizures are the core symptom of the disorder". The diagnosis of this subtype of epilepsy must come from a specific genetic test (Berg et al., 2010).

Structural/metabolic epilepsies are caused by "distinct other structural or metabolic condition or disease", such as lesions caused by stroke, trauma, and infections. These conditions may have a genetic cause, such as some malformations, but the separate disorder causes the epilepsy as opposed to a genetic aetiology (Berg et al., 2010).

Unknown epilepsies include all the other cases that are yet to have an aetiology found/confirmed. In the ILAE commission's words "the nature of the underlying cause is as yet unknown" (Berg et al., 2010).

Epileptogenesis and the epilepsies have numerous aetiologies and it seems unlikely that there exists a universal mechanism underlying its development. Even the epilepsies in which a causative gene is identified, environmental variables seem to be important for the development of epilepsy (Walker et al., 2002).

1.1.4 Epileptogenesis

Epileptogenesis refers to the process of the development of epilepsy that leads to the occurrence of spontaneous seizures. It consists of a phenomenon in which the brain undergoes molecular and structural changes after an insult that eventually lead to the occurrence of spontaneous seizures (Pitkänen et al., 2007; Goldberg and Coulter, 2013; Williams et al., 2007). The period between the brain insult and the onset of the first spontaneous seizure — chronic epilepsy — is termed the latent

period (Hellier et al., 1998; Williams et al., 2007).

The definition of latent period brings some challenges to clinical practice. For an accurate evaluation of such period, in humans, both the time of the brain injury and first seizure must be correctly defined. However, these two are not always clearly marked. This is even more challenging when it comes to the first seizure as it might not be recognised by the carer or occur during sleep (Williams et al., 2007). The average time for epileptogenesis in humans is 7.5 years (French et al., 1993). Common factors for epileptogenesis in humans include traumatic brain injury (TBI), stroke, and cerebral infections (Pitkänen et al., 2007).

In animal models, epileptogenesis shows latent periods of several days to months (Williams et al., 2007). This period has, however, shown dependence on the animal model and the severity of the brain injury (Williams et al., 2007).

There is evidence shown by some studies that epilepsy worsens in time; however, some do not show very clear signs of change in the frequency or duration of seizures (French et al., 1993; Bertram and Cornett, 1993, 1994; Hellier et al., 1998; Williams et al., 2007; Mazarati et al., 2002). In the latter case, the latent period is a phase which includes all the molecular and structural changes that facilitate epileptic seizures and the dynamics of the disease would be described by two discrete states: without and with seizures. The former scenario would still feature changes after the first epileptic seizures; the frequency of these would converge to a plateau at some point (Williams et al., 2007).

Knowledge of epileptogenesis and epileptic seizures dynamics can provide insights for clinical practice and enable the discovery and creation of novel treatments,

including anti-epileptic drugs and surgery.

1.1.5 Treatment

1.1.5.1 Current treatment approaches

The two main current treatment options for epilepsy are the use of medication, and in some cases, a resective surgical procedure. Seizure control is achieved in two-thirds of patients with focal epilepsy through the use of antiepileptic drugs (AEDs) (Mormann et al., 2007).

AEDs protect against seizures by acting on, and affecting, the function of a variety of cellular targets, suppressing the abnormal hypersynchronous activity in brain circuits (Rogawski and Cavazos, 2015). However, for a significant proportion of epilepsy patients ($\sim 30\%$), these drugs are not effective (Mormann et al., 2007).

In cases in which there is effective control and reduction of the ictal episodes, treatment may be limited by adverse effects. Even when successful, long-term treatment with AEDs can cause several unwanted effects, such as cognitive and other neurological dysfunction (Elger, 2001).

The epilepsies that cannot be controlled by medication are termed “drug resistant” and are defined, by the ILAE, as a condition that features “a failure of adequate trials of two tolerated and appropriately chosen and used AED schedules (whether as monotherapies or in combination) to achieve sustainable seizure freedom” (Kwan et al., 2010).

In societies with advanced health care systems, seizure control can be achieved in drug resistant cases by a resective surgical procedure (Mormann et al., 2007).

Surgical therapy is based on the principle of removing and/or disconnecting a critical mass of the epileptogenic network. However, the long-term seizure outcome tends to change over time and the likelihood of success of the surgical procedure seems to be inversely related to the complexity of the epileptogenic network (Wiebe and Jette, 2012).

In a postsurgical outcome study of 615 patients, 52% remained seizure free at 5 years after the surgery, and 47% at 10 years after the procedure. Surgery can also result in functional deficit and is not without risk, including death (de Tisi et al., 2011).

1.1.5.2 New therapeutic approaches

The ability to identify earlier seizure markers could lead to new therapeutic methodologies, and improve current techniques, e.g. targeted AED administration and/or surgical procedures through improved localisation based on the spread seizure patterns knowledge, thereby providing a more effective treatment option for those patients who do not attain sustained seizure freedom.

Also, the treatment concepts could move from a preventive strategy towards an on-demand therapy by the delivery of anticonvulsants substances or by electrical stimulation of the brain prior to the seizure onset, that would depend only on the correct characterisation, and detection or prediction of a possible ictal state (Mormann et al., 2007).

1.2 On fractals

1.2.1 Beyond Euclid: fractals

Geometry is the branch of mathematics dealing with the properties of figures in space. Euclid compiled knowledge on geometry available in 300 BC, which was subsequently called Euclidian geometry. The figures of such geometry are constructed from various elements such as points, lines, planes, curves and surfaces. Euclid also defined axioms and relationships between these elements (Wolfe, 2012).

Euclidean geometry, with its axioms, is unable to capture and describe some naturally occurring shapes (e.g. shapes of clouds, galaxies, mountains, leaves, or neurons). The French mathematician Benoit Mandelbrot describes such natural patterns as "fragmented and irregular", emphasising the discrepancies with Euclids figures (Mandelbrot, 1983). This led him to conceive the concept of fractal geometry, enabling the description of objects with much higher complexity that transcends traditional geometric shapes (Mandelbrot, 1983).

1.2.2 What is a fractal?

The term fractal was coined by Mandelbrot (1983), from the Latin adjective *fractus*. The verb *frangere* means, in Latin, to break, to create irregular fragments. So, a fractal would be a broken and irregular object. A fractal is a shape that is "self-similar" or "scale-invariant" and has a Hausdorff-Besicovitch dimension, or simply Hausdorff dimension, that exceeds the topological dimension (Feder, 1988). If the object has these characteristics, the Hausdorff dimension is also called fractal dimension (Mandelbrot, 1983).

The topological dimension corresponds to the amount of information needed to localise a coordinate in an object, e.g. in a line only one index is required to differentiate a point from any other possible points. It can also be defined based on the dimension of the object needed to split the initial figure in two pieces; this must have a dimension $n-1$, where n is the topological dimension of the figure, e.g. a point (conventionally treated as a zero-dimensional element) can split a line (topological dimension = 1) in two pieces (Mandelbrot, 1983; Feder, 1988).

Scale invariance is the property of an object to retain its features when the scale changes by a fixed factor. Such a property is possessed by several objects in nature, such as the $1/f$ noise (Bak et al., 1987). For fractal objects, scale invariance can be considered as the property of being made of parts similar to the whole object (Feder, 1988). Examples of fractal objects can be seen in figure 1.2.

1.2.3 Hausdorff dimension

The Hausdorff dimension quantifies the way in which an object occupies the space in which it is immersed and can have a fractional value (Feder, 1988). The length L of a straight line can be measured by line segments of size δ. The number of elements used, $N(\delta)$, will give the value of the line's size according to the equation:

$$L = N(\delta)\delta \tag{1.1}$$

The same concept can be extended to the measure of areas and volumes. In this case small squares of area δ^2 and cubes of volume δ^3 are used, respectively, according to the equations:

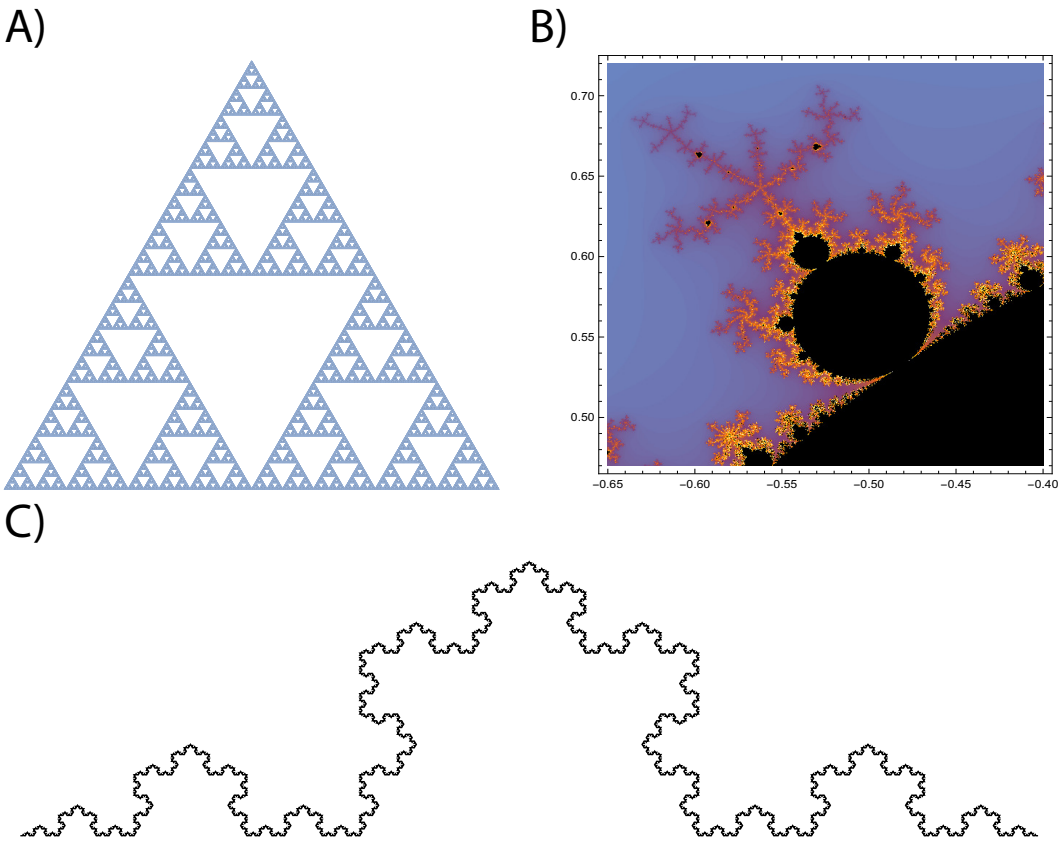

Figure 1.2: Examples of fractal objects. A) Sierpiński triangle (Hausdorff dimension $D \approx 1.585$ (Falconer, 2003)). B) Mandelbrot set (boundary $D = 2$ (Shishikura, 1998)). C) Koch curve ($D \approx 1.263$ (Feder, 1988)). The Haudorff dimension is described in section 1.2.3.

$$A = N(\delta)\delta^2 \tag{1.2}$$

$$V = N(\delta)\delta^3 \tag{1.3}$$

These three equations allow us to define a general rule to represent a measure M_d of a generic object, in which the d exponent represents the dimension of the measurement object (line segments $d = 1$, squares $d = 2$, cubes $d = 3$, etc.) (Feder, 1988):

$$M_d = \lim_{\delta \to 0} N(\delta)\delta^d \tag{1.4}$$

The number of elements used to cover the object can be generalised in the following way, where the K represents a proportionality constant and D is a scaling exponent:

$$N(\delta) = K\delta^{-D} \tag{1.5}$$

A final expression for measure can be obtained by replacing $N(\delta)$ with the previous definition, giving:

$$M_d = \lim_{\delta \to 0} K\delta^{(d-D)} \tag{1.6}$$

This final expression presents, varying d, a transition point when $d = D$ in a similar way to the formal definition of the Hausdorff dimension ($dim_H F$). This transition point also represents the Hausdorff dimension value, i.e. the value for the dimension of the measurement object in which M_d is different from 0 and ∞.

$$M_d(\delta) = \begin{cases} 0, & d > D \\ \\ \infty, & d < D \end{cases} \tag{1.7}$$

For a fractal object, the Hausdorff dimension D can be determined by the expression for M_d by counting the number of segments needed to cover a figure and is called the box counting dimension (Feder, 1988).

A formal definition of the Hausdorff dimension is given by Falconer (2003) and is based on topology and set theory. Taking U as a generic, non-empty, subset of n-dimensional $\mathbb{R}^n$ space, its diameter can be defined as $|U| = \sup\{|x-y| : x,y \in U\}$. If $\{U_i\}$ is a countable collection of sets of diameter δ that cover F (a subset of $\mathbb{R}^n$), it is called a δ-cover of F:

$$F \subset \bigcup_{i=1}^{\infty} U_i \quad \text{with} \quad 0 \leq |U_i| \leq \delta \tag{1.8}$$

Supposing F is a subset of $\mathbb{R}^n$ and S is a non-negative number, for any $\delta > 0$, it is possible to define:

$$\mathscr{H}^S_\delta(F) = \inf\left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\} \tag{1.9}$$

$$\mathscr{H}^S(F) = \lim_{\delta \to 0} \mathscr{H}^S_\delta(F) \tag{1.10}$$

The quantity $\mathscr{H}^S(F)$ is called the S-dimensional Hausdorff measure, in which S represents the dimension of the measurement object and will impact on the obtained estimate (Falconer, 2003). For any set $F \subset \mathbb{R}^n$ and $\delta < 1$, it is possible to assert that:

$$\sum_i |U_i|^t \leq \sum_i |U_i|^t \frac{|U_i|^S}{|U_i|^S} \tag{1.11}$$

$$\sum_i |U_i|^t \leq \sum_i |U_i|^{t-S} |U_i|^S \tag{1.12}$$

$$\sum_i |U_i|^t \leq \delta^{t-S} \sum_i |U_i|^S \tag{1.13}$$

And taking the infimum

$$\inf \left\{ \sum_i |U_i|^t \leq \delta^{t-S} \sum_i |U_i|^S \right\} \tag{1.14}$$

$$\mathscr{H}^t_\delta(F) \leq \delta^{t-S} \mathscr{H}^S_\delta(F) \tag{1.15}$$

Taking the limit when $\delta \to 0$, it is possible to conclude that, if $\mathscr{H}^S(F) < \infty$, $\mathscr{H}^t(F) = 0$ for $t > S$. There is a critical limit for S (S_c) in which $\mathscr{H}^S(F)$ changes from ∞ to 0 as S varies. To this value is given the name of the Hausdorff dimension, or the Hausdorff-Besicovitch dimension, $S_c = dim_{\mathscr{H}} F$. As shown in figure 1.3.

1.2.4 Estimating the Hausdorff dimension: the box counting method

The general equation for the number of measurement objects used can be transformed such that it becomes a linear relationship, with D being the slope of the linear function:

$$\log N(\delta) = \log K \delta^{-D} \tag{1.16}$$

$$\log N(\delta) = -D \log \delta + \log K \tag{1.17}$$

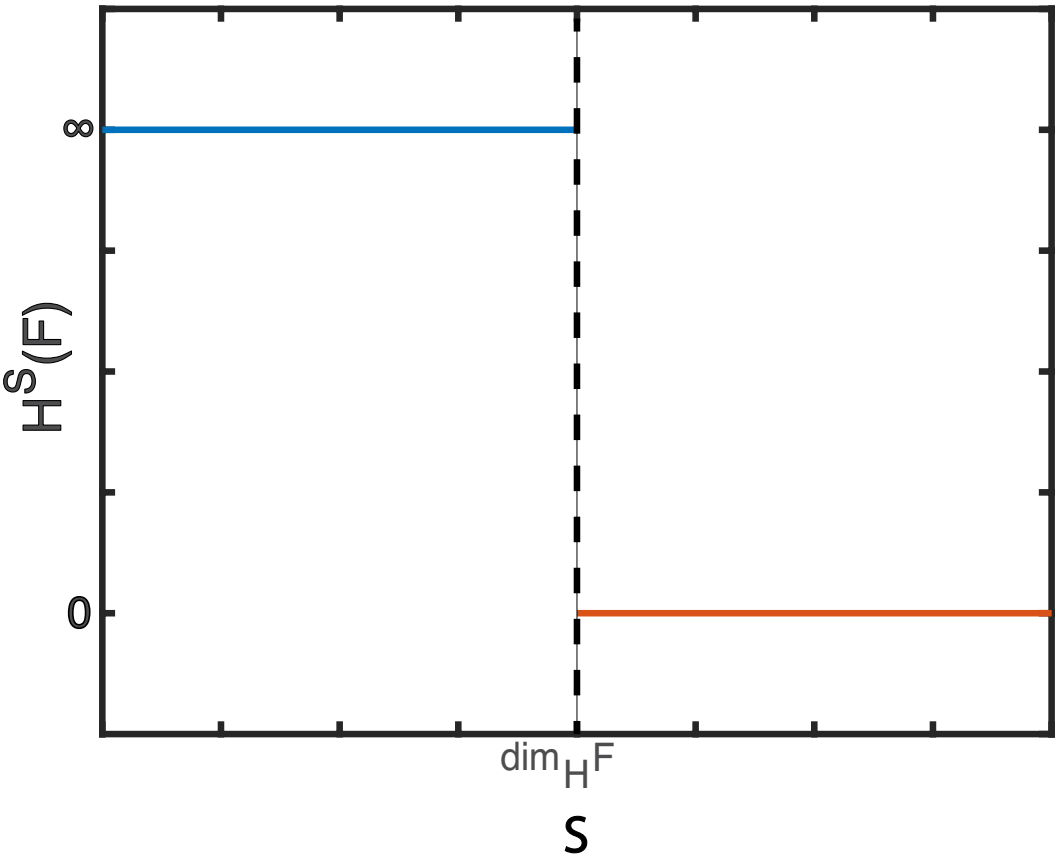

Figure 1.3: Transition on the value of the Hausdorff measure $\mathscr{H}^S(F)$ as the measurement object's dimension S changes. The plot exhibits a transition in which the value of the Hausdorff measure goes from ∞ to 0 for $S_c = dim_{\mathscr{H}} F$.

The box-counting method was devised from this expression and consists of a box with equal dimensions, incorporating the figure. This box is further split into smaller ones keeping equal side lengths. The number of boxes occupied by parts of the figure is then counted and registered. The splitting procedure is repeated for n iterations as shown in figure 1.4.

By counting the number of boxes $N(\delta)$ needed to cover the object, and plotting it against the size of the boxes (δ) in log space, the slope of the plot indicates the Hausdorff dimension.

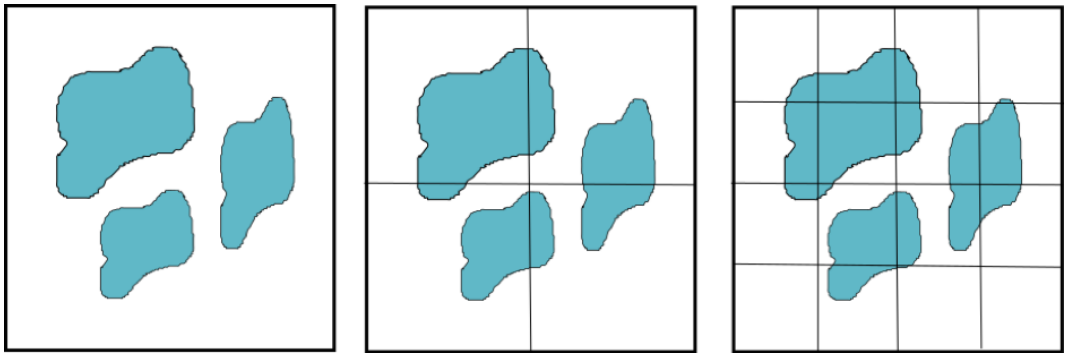

Figure 1.4: Illustration of the box counting procedure used for estimating the Hausdorff dimension of an object. The number of occupied boxes should present a power law in relation to the boxes' side length. The exponent of this function corresponds to the Hausdorff dimension. A fractal's dimension is equal to its Hausdorff dimension.

1.2.5 Fractals and time series

Many time series obtained from natural phenomena present a scale-invariance behaviour, e.g. rainfall (Andrade et al., 1998) and earthquake (Sornette and Sornette, 1989) recordings. However this invariance is not isotropic at times, i.e. the scale factor changes in different directions, a phenomenon called self-affine structure (Feder, 1988).

While trying to devise the ideal water reservoir, i.e. that should never empty or overflow, in the middle of the twentieth century, the American engineer Harold Edwin Hurst developed a method to estimate scaling properties called rescaled range analysis (R/S analysis) (Feder, 1988; Mandelbrot and Wallis, 1968). In brief, if a water reservoir receives an amount of water $\xi(t)$ in a time period τ, the mean influx can be calculated by the expression:

$$< \xi >_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} \xi(t) \tag{1.18}$$

For an ideal reservoir the amount of water that leaves should be the same as the influx. At any time, t, the accumulated outflow can be expressed as:

$$X(t,\tau) = \sum_{u=1}^{t} \{\xi(u) - <\xi>_\tau\} \tag{1.19}$$

Hurst defined the quantity R (from range) as the difference between the maximum and the minimum of the accumulated outflow:

$$R(\tau) = \max((X(t,\tau))) - \min((X(t,\tau))) \tag{1.20}$$

The standard deviation of the accumulated outflow can be defined as:

$$S = \left(\frac{1}{\tau}\sum_{u=1}^{t} \{\xi(u) - <\xi>_\tau\}^2\right)^{\frac{1}{2}} \tag{1.21}$$

Studying a range of natural phenomena, Hurst found the following empirical relationship between the R function and the standard deviation (Feder, 1988)

$$\frac{R}{S} = \left(\frac{\tau}{2}\right)^H \tag{1.22}$$

Where H is the Hurst exponent, which is characteristic of the phenomenon that generated the time series X. The Hurst exponent is related to the fractal dimension of the time series by the following expression (Feder, 1988):

$$D = 2 - H \tag{1.23}$$

H can vary between 0 and 1. Time series with $H < 0.5$ present an anti-

persistent behaviour, while time series with $H > 0.5$ exhibit a persistent pattern, and $H = 0.5$ corresponds to a random process (Feder, 1988). Figure 1.5 shows examples of time series with different values of H (Feder, 1988).

Persistent time series present positive correlation between its elements, as anti-persistent present an opposite behaviour, i.e. increases are followed by decreases and vice versa. Random processes have no correlation between their elements and are known as processes without memory (Feder, 1988).

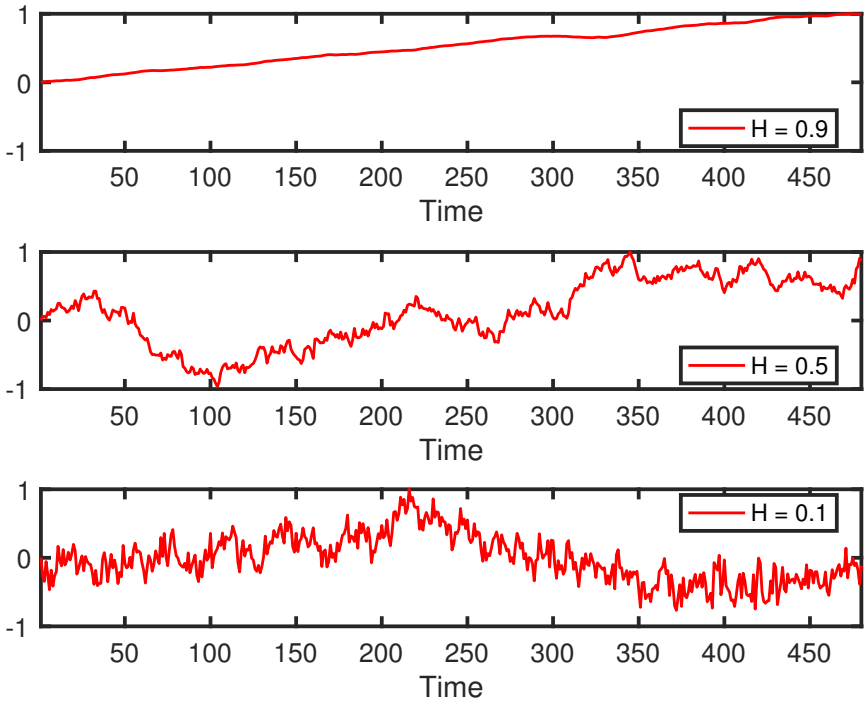

Figure 1.5: Samples of simulated random walks with different values of the Hurst exponent (H). From top to bottom: persistent behaviour ($H = 0.9$ — $D = 1.1$), random ($H = 0.5$ — $D = 1.5$) and anti-persistent ($H = 0.1$ — $D = 1.9$).

Hurst reported that many natural phenomena feature an exponent $H > 0.5$. These values would indicate that natural phenomena follow a persistent behaviour (with memory). For example, in a process with memory, a river discharge would be

affected not only by recent rain levels but also by previous records. Long drought periods would cause the drainage basin to dry and in subsequent periods of higher precipitation would have part of the water absorbed by the same drainage basin. Whereas, in high precipitation years, such a drainage basin would be saturated and not capable of absorbing the same amount of water — resulting in bigger river discharges even for lower precipitations (Feder, 1988).

1.2.6 Multifractals

While fractal analyses brought some insights about the nature of certain phenomena, nonetheless, some structures seem to require more than one-dimension D to be adequately characterised. For example, in time series, the value of D can change as a function of time. Objects that require an infinite number of indices to be properly characterised are called multifractals (Stanley et al., 1999), e.g. several phenomena such as turbulence (Meneveau and Sreenivasan, 1987), soil composition (Miranda et al., 2006) and human activity (França et al., 2019).

The Lipschitz-Holder exponent (α) quantifies the singularities of a structure whose Hausdorff dimension is described by the function, $f(\alpha)$, the combination of these two indices result in a graph called the multifractal spectrum (Chhabra and Jensen, 1989; Feder, 1988).

In order to give an intuition of the multifractal spectrum (in a similar way to the physical meaning of the fractal dimension), figure 1.6 shows examples for the three time series, namely a simulated multifractal time series, a surrogate time series (shuffled version of the original multifractal time series), and a generated random

time series.

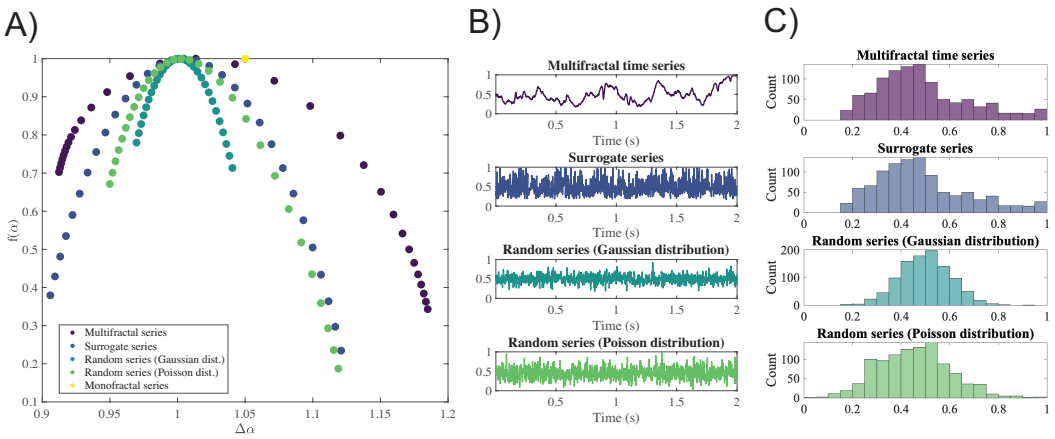

Figure 1.6: Multifractal spectra for different types of simulated time series. A) Spectra obtained for different types of time series. B) Series corresponding to each spectrum. C) Histogram of each time series. The spectra have a similar parabolic shape. The surrogate data and the random series (Gaussian and Poisson) present a reduced spectral width — with the Poisson series showing a wider spectrum than the Gaussian signal. The point labelled as Monofractal series is an idealisation. In a perfect monofractal time series, the spectrum consists of a single point.

Given its parabolic shape the spectrum can be characterised by the difference between its maximum and minimum value $f(\alpha)$ which corresponds to $\alpha(q=0)$ or α_0, and the parabola width, given by the difference between α_{q-} and α_{q+}, corresponding, respectively, to the lowest (negative) and highest (positive) values of q; see figure 1.7 (Miranda et al., 2006). Δf expresses the diversity of Hausdorff dimensions of the different singularities while $\Delta\alpha$ represents the object's diversity of singularities; the quantities $\alpha_0 - \alpha_{q-}$ and $\alpha_{q+} - \alpha_0$ are also used, providing specific information about small or large (respectively) singularities and can be used to characterise the spectrum's symmetry.

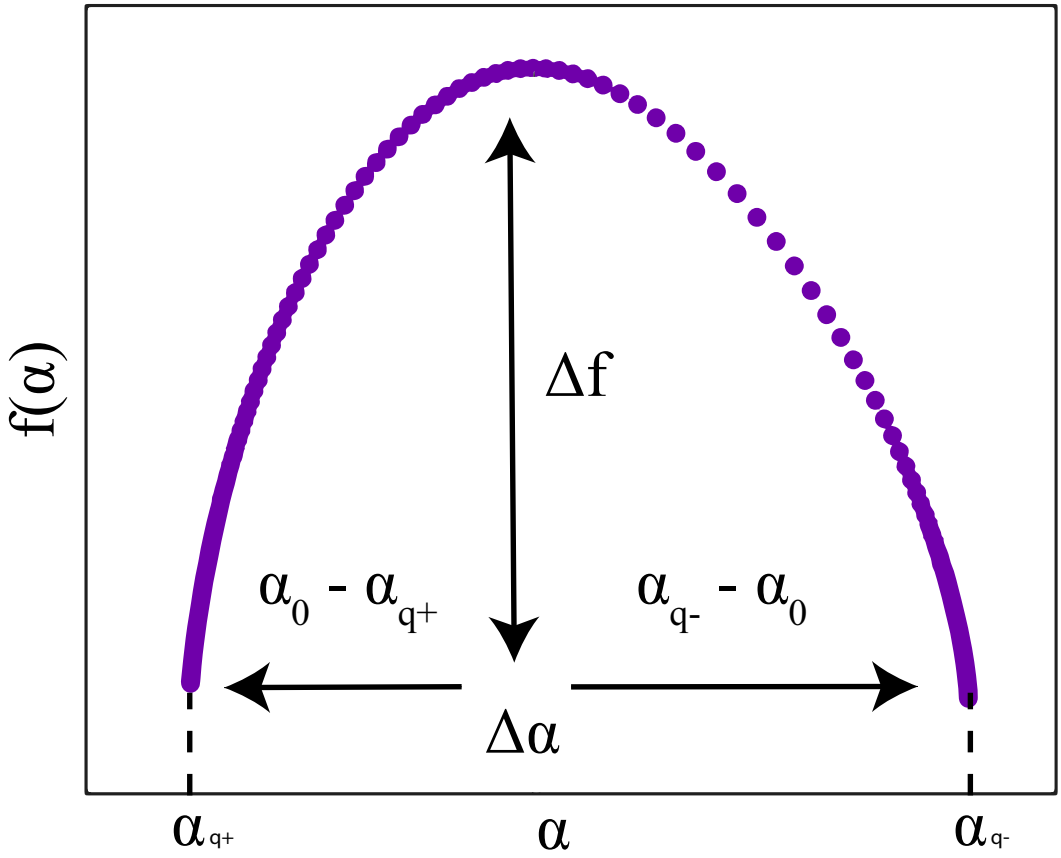

Figure 1.7: Multifractal spectrum parabola characteristics. $\Delta\alpha$ gives a measure of diversity of singularities in the profile, as Δf represents the different Hausdorff dimensions of each singularity on the profile. The two half widths ($\alpha_0 - \alpha_{q-}$ and $\alpha_{q+} - \alpha_0$) provide information about the diversity of singularities on each positive and negative q values.

1.2.6.1 Multifractal theory: The q factor and generalised dimensions

The scale invariance analysis in a monofractal object is applied only to the second statistical moment. This evaluation can be expanded to the q^{th} order moment. If the scaling properties change with q, the object is multifractal (Biswas et al., 2012). For a generic timeseries, the mass exponent $\tau(q)$ can be defined as:

$$< [\Delta Z(x)]^q > \propto X^{\tau(q)} \tag{1.24}$$

$\Delta Z(x)$ represents a generic measure of the time series and x the length of the analysed window. If the graph of $\tau(q)$ versus q presents a single slope (straight line), the series can be described with only one index and consequently is a monofractal (Biswas et al., 2012). An example can be seen in figure 1.8.

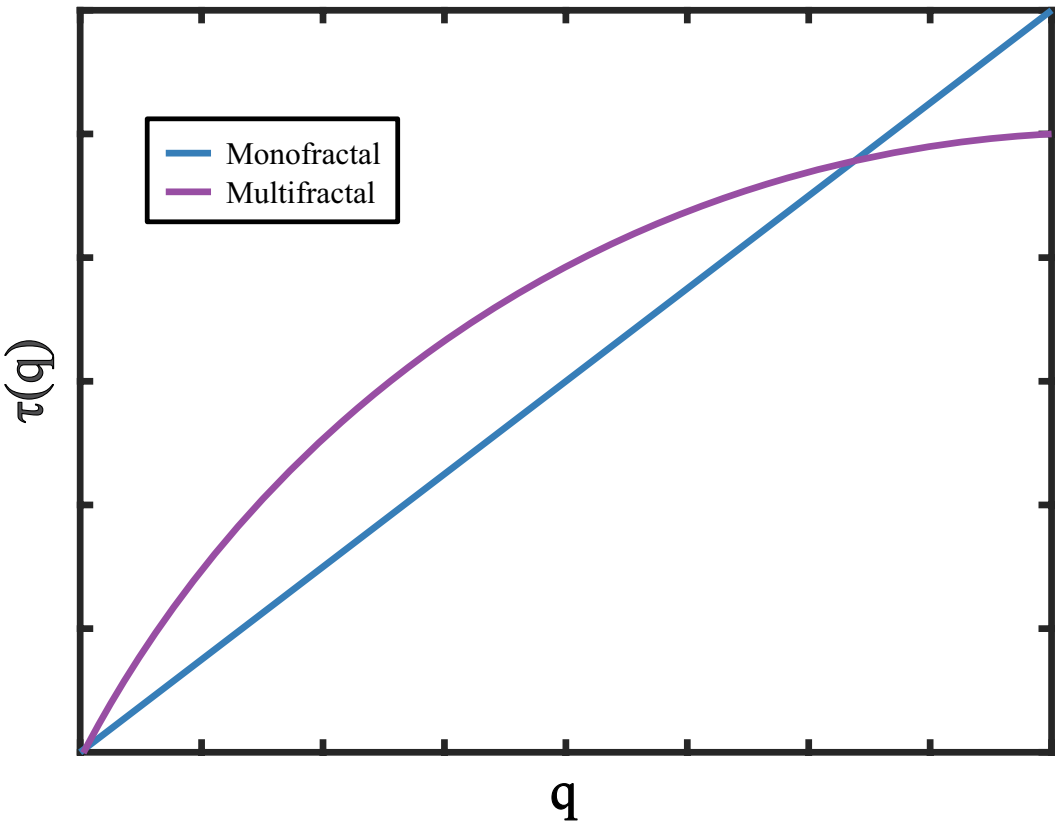

Figure 1.8: Mass exponent graph, or $\tau(q)$ curve, of a monofractal and a multifractal object. As it can be seen, the monofractal curve is a straight line (only one slope value on the whole length).

It is possible to define a generalised dimension (a generalisation of the fractal dimension), or fractal dimensions' spectrum D_q, based on the q parameter (Feder, 1988; Chhabra and Jensen, 1989) as follows:

$$D_q = \frac{1}{(q-1)} \lim_{L \to 0} \frac{\log \sum_i P_i^q(L)}{\log L} \tag{1.25}$$

P_i represents the probability (integrated measure) of finding a point in a box of size L (Chhabra and Jensen, 1989). The equation can be rewritten as:

$$D_q = \frac{1}{(q-1)}\tau(q) \tag{1.26}$$

The mass exponent $\tau(q)$ can then be defined as:

$$\tau(q) = \lim_{L\to 0} \frac{\log \sum_i P_i^q(L)}{\log L} \tag{1.27}$$

According to Chhabra and Jensen (1989), the index q can be conceived as a microscope, probing the object's singularities at various scales (Chhabra and Jensen, 1989). For $q > 0$, the more singular regions are amplified and therefore contribute to a greater degree to D_q, by contrast, for $q < 0$ the less singular regions are accentuated. Figure 1.9 presents a generalised dimensions' spectrum for a time series; it presents two regions of convergence and an inflexion point.

1.2.6.2 Multifractal theory: The Lipschitz-Hölder exponent and the multifractal spectrum

Another approach to multifractal analysis consists of the analysis of the multifractal spectrum of the object. The Lipschitz-Hölder exponent, α, can be defined as:

$$P_i(L) \sim L^{\alpha} \tag{1.28}$$

A normalised measures' family can be built up with the parameter q, in this case, $\mu_i(q,L)$ representing the probabilities in boxes of size L, defined as:

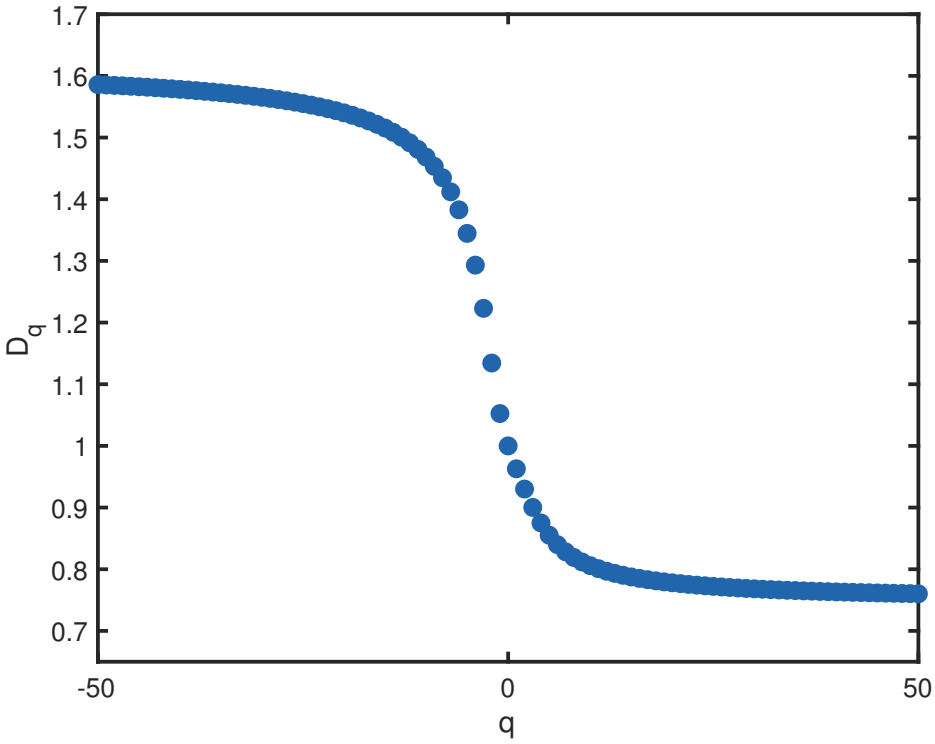

Figure 1.9: Example of generalised dimensions' spectrum for a time series. In this case, q was chosen to vary from -50 to 50.

$$\mu_i(q,L) = \frac{P_i(L)^q}{\sum_j P_j(L)^q} \tag{1.29}$$

$P_i(L)$ represents the integrative measure cited previously in equation 1.24 and consists of the normalised sum of the time series values in a window. From the normalised measures $\mu_i(q,L)$, Chhabra and Jensen (1989) derived the multifractal spectrum that can be obtained from:

$$\alpha(q) = \lim_{L\to 0} \frac{\sum_i \mu_i(q,L)\log P_i(L)}{\log L} \tag{1.30}$$

and

$$f(\alpha) = \lim_{L \to 0} \frac{\sum_i \mu_i(q,L) \log \mu_i(q,L)}{\log L} \tag{1.31}$$

1.2.6.3 Multifractal theory: The relations between generalised dimensions and the multifractal spectrum

The variables $f(\alpha)$, $\tau(q)$, α and q are thermodynamically conjugated, which means that they can be related by a Legendre transform (Gu and Zhou, 2010). This way, it is possible to rewrite them using the following expressions:

$$\alpha(q) = \frac{d\tau(q)}{dq} \tag{1.32}$$

$$f(q) = q\alpha - \tau(q) \tag{1.33}$$

1.3 On fractals and epilepsy

The fractals framework has been widely applied in different studies of human brain dynamics in health (Pereda et al., 1998; Linkenkaer-Hansen et al., 2001; Papo et al., 2017; Gong et al., 2003) and disease (Gómez et al., 2009; Zappasodi et al., 2014; Esteller et al., 1999). These efforts provided some knowledge on critical behaviour of the brain and putative changes provoked by disease processes in conditions such as Alzheimer's disease (Gómez et al., 2009), stroke (Zappasodi et al., 2014), and epilepsy (Esteller et al., 1999).

In refractory mesial temporal lobe epilepsy cases, fractal analysis of electroen-

cephalography signals has shown changes at seizure onset. These signals were obtained from intracranial depth and strip electrodes in patients undergoing epilepsy surgery evaluation. In these signals the fractal dimension increases at the onset of the epileptic seizures suggesting more complex dynamics during the ictal phase — hypothesised to be connected to the rhythms that generate the epileptic seizure (Esteller et al., 1999, 1995). Another study applied fractal metrics to EEG recordings from rats and has shown three different patterns for seizure onset (Li et al., 2005).

Fractal analysis has also been employed in studies on seizure detection; these studies hypothesised that if fractal metrics are able to show differences in time series, maybe they could also differ between interictal and ictal segments (Polychronaki et al., 2010; Päivinen et al., 2005; Yuan et al., 2012). Another study has shown changes in fractal properties of local field potential signals in rats after a brain injury (Spasic et al., 2005).

Polychronaki et al. (2010), Päivinen et al. (2005), and Yuan et al. (2012) employed fractal methods as a feature of the EEG signals containing ictal activity. Polychronaki et al. (2010) evaluated scalp EEG signals from 8 patients with a total of 55 seizures. These signals were registered at 400 Hz. Katz's and Higuchi (described in details in section 3.2.1.1) methods were used as features for a k-nearest neighbours (k-NN) algorithm as a classifier. The authors report a sensitivity of 100% and a False Prediction Rate (FPR) of $0.42h^{-1}$. Yuan et al. (2012) evaluated data from 21 patients recorded at 256 Hz — with a total of 65 epileptic seizures. The authors used fractal properties of the intracranial EEG signals as features of a single-layer neural network classifier. The studies reported a sensitivity of 93.85%

and a FPR of $0.35h^{-1}$. Päivinen et al. (2005) applied the Higuchi method to signals obtained with implanted electrodes in rats. The data was recorded at 256 Hz and the fractal features of the signal were evaluated with discriminant analysis.

Further evidence shows that brain dynamics are in fact multifractal (Ihlen and Vereijken, 2010; Ciuciu, 2012; Zhang et al., 2015; Suckling et al., 2008; Zorick and Mandelkern, 2013; Papo et al., 2017). A breakdown of the monofractal power-law pattern in brain dynamics suggests that additional statistical moments, not reflected by a single statistical moment such as in monofractal characterisation, may be necessary to model brain signals (Fraiman and Chialvo, 2012). Moreover, it is known that interacting processes with different fundamental time scales, such as those observed in the brain, can generate multifractal patterns (Argoul et al., 1989; Suckling et al., 2008).

The multifractal formalism was previously applied to a seizure detection study (Zhang et al., 2015). The authors reported a significant reduction in the multifractal spectral width for the ictal segments, as well as the capability of classifying different segments as ictal or interictal based on the multifractal spectra geometrical measures and employing a Support Vector Machine classifier. Nevertheless, the authors did not analyse in depth the variation of the multifractal indices as a seizure emerges or the dependency of such markers in relation to clinical features, such as the type of epilepsy, and the study was limited to introduce the classifier.

Considering the information about the brain function provided by multifractal analysis in previous studies (Fraiman and Chialvo, 2012; Argoul et al., 1989; Suckling et al., 2008), it is reasonable to assume that such methodologies could

yield relevant information to the study of ictal activity — as it has been suggested by (Zhang et al., 2015). There is, however, evidence that some non-linear metrics would not provide any additional information from EEG recordings when compared to linear metrics, e.g. variance (McSharry et al., 2003). This has not been evaluated for multifractal approaches thus far; however these metrics are more likely not to present such issues due to the multiple statistical moments employed during the derivations of the multifractal spectra. This topic is discussed in detail in chapter 3.

This thesis focusses on the development of an approach to the study of multifractal properties of EEG signals recorded from individuals with epilepsy. The analyses shown in the subsequent chapters evaluate the time-varying properties of the brain's critical dynamics. Differently from the work developed by Zhang et al. (2015), this thesis does not aim to present a classifier that exploits multifractal properties of EEG signals for seizure detection or prediction algorithms. It is, nonetheless, a study on the temporal variation of such properties and how these features can be used to understand epileptic seizures and even classify groups of similar epileptic discharges.

Chapter 2

Aims and hypotheses

The overall aim of this work was to use multifractal analysis to develop a framework that could be widely applied to study brain electrophysiological signals, especially in patients with intractable focal epilepsies. Within this aim, specific hypotheses were tested.

There is considerable evidence that many advanced methods for studying the EEG during epileptic seizures are biased by changes in the variance of the recording, common to many seizures, leading to the following hypothesis:

- *Scaling analysis methods, such as multifractal spectrum estimation approaches are affected by changes in the standard deviation of the time series, which can bias seizure analysis applications (Chapter 3).*

Recordings for pre-surgical planning in refractory epilepsies feature durations of days — or weeks — including sleep. Moreover, putative closed-loop devices for controlling seizures should also consider the sleep patterns of the patient. Hence, it is important to evaluate how the different sleep stages would impact scaling properties of the recording. Therefore, the following hypothesis was tested:

- *Sleep phases impact multifractal estimation in human long-term electrographical recordings for epilepsy surgery planning (Chapter 4).*

Although there are many studies on methods to detect and/or predict seizures, it is largely unknown how sampling frequency can affect such methods.

- *Focal epileptic seizure detection exhibits optimal results at specific sampling frequencies (Chapter 5).*

To date, there is no objective electrographic classification scheme for focal epileptic seizures. A classification of this type of epileptic seizures could not only provide insights into the onset and dynamics of such events but also suggest optimal therapeutic interventions. Therefore, the following hypothesis was tested:

- *Focal epileptic seizures can be classified into families using multifractal analysis (Chapter 6).*

Lastly, following a brain insult (e.g. prolonged seizure, traumatic brain injury, stroke), there are changes in brain circuitry that result in spontaneous epileptic seizures (this process is termed epileptogenesis).

- *Changes that occur during epileptogenesis can be detected by the multifractal analysis approach. An adequate application of multifractal analysis can identify animals that could develop spontaneous seizures after an initial trauma based on the analysis of local field potential signals. (Chapter 7)*

The framework presented in this thesis should be generic and allow different applications. It should also provide insights into the genesis of brain signals and how epileptic seizures — and their emergence — change these recordings.

Chapter 3

An unbiased multifractal evaluation approach for use as a signal feature for machine learning in clinical applications

3.1 Introduction

Drug-resistant epilepsies impose a significant burden to patients and healthcare systems around the world (Carney et al., 2011). Over the last decades, researchers have sought alternative approaches. Many of these, such as response-mode stimulation, depend upon seizure detection and seizure prediction (Mormann et al., 2007).

As for any type of forecasting, the power of predicting seizures relies on the existence of early, preceding signs of the ictal state. The existence and identification of a pre-ictal state is a prerequisite fore response-mode interventions that prevent

seizure occurrence.

Any pre-ictal state should, *a priori*, be characterised as a gradual change, or cascade of changes, in EEG dynamics (Mormann et al., 2007). These changes should therefore be detected by an adequate analytical methodology. Once pre-ictal features have been demonstrated to exist, it would hopefully be possible to devise a therapeutic intervention based on the pre-ictal signal.

Seizure detection would, in contrast, be defined by the capacity of inferring the occurrence of an epileptic seizure at the ictal onset. The prediction differs from detection, as forecasting, intrinsically, suggests some prior information on what is about to emerge, and detection is status reporting of the current behaviour. Prediction can thus be seen as a generalised case, in which, when the warning time prior to the seizure is zero or less, it becomes a detection approach. For prediction studies, the horizon, i.e. time interval between the warning and the ictal onset, must be greater than zero and should provide an interval that makes a possible intervention feasible (Snyder et al., 2008).

In the literature, this horizon is presented in three distinct ways: a set amount of time, which should be satisfied in order to consider a warning as a correct prediction (Shufang Li et al., 2013; Eftekhar et al., 2014; Esteller et al., 1995; Gadhoumi et al., 2013); a statistical value estimated from the mean of the interval times for a certain approach (Haddad et al., 2014; Navarro et al., 2005; Zandi et al., 2011, 2013); or a range of values presented by the apparatus (Bedeeuzzaman et al., 2013; Gadhoumi et al., 2015; Mormann et al., 2003).

Nonetheless, the literature on seizure prediction and detection is extensive and

a more detailed description of the field is beyond the scope of this thesis. It is important to emphasise, though, the relevance of non-linear metrics in the development of seizure prediction approaches. (Iasemidis et al., 1990) has proposed the existence of a pre-ictal transition, based on the largest Lyapunov exponent analysis of EEG signal of patients with epilepsy. These signals exhibited a reduced exponent for the recordings prior to the seizures. The reduced value for the exponent suggests a less chaotic process, and, consequently, characterises a seizure onset as a transition to a more ordered state (Iasemidis et al., 1990).

Another study revealed long-term trends in epileptic seizure emergence and a non-linear transition prior to the ictal onset in patients with mesial temporal lobe epilepsy and an identified focus (Martinerie et al., 1998). Such phenomenon — the change in non-linear measures prior to seizure-onset — was also observed later with a Kolmogorov entropy analysis (van Drongelen et al., 2003).

As was also observed by Iasemidis et al. (1990), the behaviour of the Lyapunov exponent over the different ictal phases, including the possible pre-ictal phase, indicates that the system (epileptic brain) enters a less chaotic state after several transient decreases in chaoticity. The existence of such non-linear features preceding the seizure supports the idea of predictability.

A major concern, however, is that complex methods of analysis may not offer a significant advance over simpler, already established methods — this is crucial for extracting putative features in machine learning applications, e.g. seizure prediction and/or detection (Baldassano et al., 2017; Brinkmann et al., 2016; Freestone et al., 2015; Karoly et al., 2017; Mormann et al., 2007; Kuhlmann et al., 2018a,b;

Varatharajah et al., 2018). For example, in the analysis of the electroencephalogram of epileptic seizures, complex non-linear methods were found to actually reproduce patterns detected by simpler measures such as variance of the signal (McSharry et al., 2003).

It is therefore essential to understand how these approaches, such as (mono- and multi-) fractal measures, relate to more traditional markers, and if new features can be obtained from the signal by applying a mono- or multi- fractal formalism.

This chapter introduces an approach for multifractal assessment of EEG signals, avoiding bias caused by changes in the variance of the signal. The approach is evaluated on both real and simulated time series and compared to classic fractal evaluation methods.

3.2 Methods

This section describes the methods employed in this study. Firstly, the monofractal and multifractal analysis approaches are introduced. Secondly, two methods to produce time series featuring scaling properties are introduced, namely fractional Brownian motion and p-model. The remaining sections describe the proposed approach to remove bias caused by the variance of the time series and concepts of feature redundancy and stability of different multifractal analysis methods.

3.2.1 Monofractal methods

Two established monofractal estimation approaches — Higuchi method (Higuchi, 1988) and Detrended Fluctuation Analysis (Peng et al., 1994) — were used. These methods are widely applied in the literature and aim to capture the features of a time

series in a single scaling exponent.

Mandelbrot (1983) defined fractals as self-similar structures with fractal dimensions (FD) that are between their topological and embedding dimensions T and E, and an established relationship of $FD+H=E$, where $E=T+1$, $T=1$ in the case of a time series, and H is the Hurst exponent. When assuming this self-similarity, one can measure both FD and H in EEG time series as alternative ways of estimating the fractal dimension. However, more generally, the fractal dimension FD and the Hurst exponent H do not necessarily reflect the same property of the time series (Gneiting and Schlather, 2004). Nevertheless, for this thesis, the two established methods were applied to estimate FD and H, respectively.

3.2.1.1 The Higuchi method

The Higuchi method measures the fractal dimension FD of a time series. It consists of constructing series with elements of an original time series and measuring their lengths (Higuchi, 1988). Given a time series with N time points $X(1),X(2),...,X(N)$, the equation 3.1 shows a rule for reconstructing smaller time series with elements of the original recording. The lengths of the time series can be assessed according to equation 3.2. The brackets $\lfloor\rfloor$ represent Gauss notation, i.e. the rounded integer of the division (Higuchi, 1988). The variable d represents a down-sampling factor of the original time series.

$$X(m),X(m+d),X(m+2d),...,X\left(m+\left\lfloor\frac{N-m}{d}\right\rfloor d\right) \quad \text{where} \quad m=1,2,...,d \tag{3.1}$$

$$L_m(d) = \frac{\left\{ \sum_{i=1}^{[(N-m)/d]} |X(m+id) - X(m+(i-1)d)| \frac{N-1}{\lfloor (N-m)/d \rfloor d} \right\}}{d} \tag{3.2}$$

If the average curve length $< L_m(d) >_m$ over d sets follows a power law, according to equation 3.3, the time series has scaling properties, with a fractal dimension FD_{Hig}.

$$< L(d) > \propto d^{-FD_{Hig}} \tag{3.3}$$

The routine used in the estimation of Higuchi fractal dimension FD is available at `https://uk.mathworks.com/matlabcentral/fileexchange/50290-higuchi-and-katz-fractal-dimension-measures`.

3.2.1.2 Detrendred Fluctuation Analysis

The Detrended Fluctuation Analysis (DFA) method is an alternative approach (Peng et al., 1994, 1995), which estimates the Hurst exponent H in time series data instead of the fractal dimension. The method consists of the following steps:

Initially the time series with N time points $X(1), X(2), ..., X(N)$ is integrated as follows:

$$y(k) = \sum_{i=1}^{k} (X(i) - \langle X \rangle) \tag{3.4}$$

Where $X(i)$ represents the ith element of the time series and $< X >$ denotes the mean over the whole recording. The second step consists of dividing the time

series into N_l windows of length l, then the mean square root of the integrated series is subtracted from the local trend, in every window (Peng et al., 1995), as shown in equation 3.5.

$$F(l) = \sqrt{\frac{1}{N_l} \sum_{k=1}^{N_l} [y(k) - y_l(k)]^2} \tag{3.5}$$

The local trend ($y_l(k)$) is obtained from a linear regression over the time series in the window, and number N_l represents the total number of windows. In the following step, equation 3.5 is obtained for several window lengths (l). The relation between $F(l)$ and l is described by a power law, according to equation 3.6, where H is the Hurst exponent.

$$F(l) \propto l^H \tag{3.6}$$

The code used here is available in the Physionet repository (`https://www.physionet.org/physiotools/dfa/`) (Goldberger et al., 2000; Peng et al., 1995).

3.2.2 Multifractal methods

In this section, three multifractal spectrum estimators are described: Multifractal Detrended Moving Average (Gu and Zhou, 2010), Multifractal Detrended Fluctuation Analysis (Ihlen and Vereijken, 2010; Ihlen, 2012; Kantelhardt et al., 2002), and Chhabra-Jensen (Chhabra and Jensen, 1989), as these are the most established methods used in the literature.

Multifractal properties are represented as spectra, where essentially the fractal

scaling properties, or more precisely Hausdorff dimensions (often noted as $f(\alpha)$), are measured over a range of different singularities (α). Formally, the singularity spectrum is a function that describes the Hausdorff dimension of subsets of the time series $X(t)$ with a specific Hölder exponent, according to:

$$f(\alpha) = D_F\{X(t_s), H(X(t_s)) = \alpha\} \tag{3.7}$$

Essentially, $f(\alpha)$ is the Hausdorff dimension (DF) of the subset (t_s) of the time series $X(t_s)$ that has a Hölder exponent α (van den Berg, 1999; Murcio et al., 2015).

To characterise the function, or singularity spectrum $f(\alpha)$, usually, the width ($\Delta\alpha$) and height (Δf) — differences of maximum and minimum values of α and $f(\alpha)$ respectively — of the spectrum are used. $\Delta\alpha$ indicates the range of singularities present in a time series; this is also the most commonly used measure of how multifractal a time series is. The spectrum height Δf indicates the range of Hausdorff dimensions present in the time series. See Fig. 3.1 for an exemplary singularity spectrum plot.

3.2.2.1 MF-DMA

Multifractal Detrended Moving Average (MF-DMA) is one of the most commonly used methods for the estimation of multifractal measures. The method of calculation consists of the following steps (Gu and Zhou, 2010). Given time series $X(t)$ with time points $X(1), X(2), ..., X(N)$, the cumulative sum time series is obtained:

$$y(t) = \sum_{t=1}^{N} X(t) \tag{3.8}$$

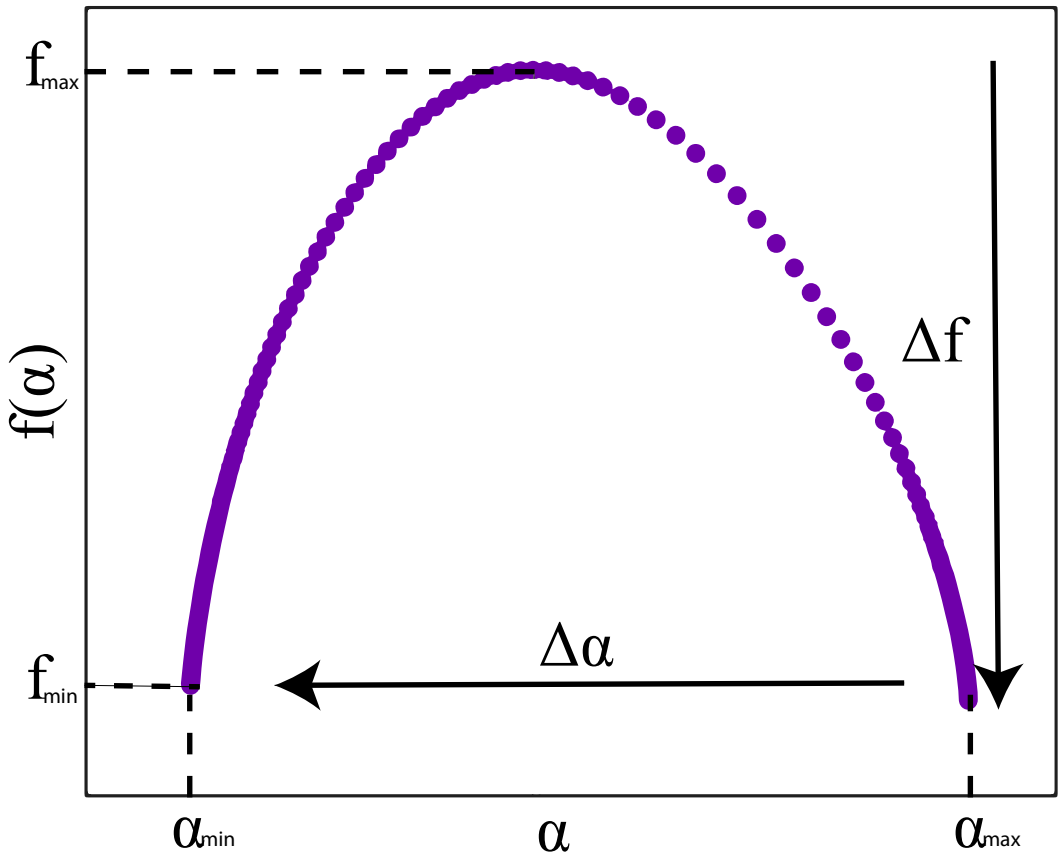

Figure 3.1: Multifractal singularity spectrum, $f(\alpha)$, with a characteristic parabolic shape. The spectrum width ($\Delta\alpha$) and height (Δf) measures are indicated by the arrows.

The moving average over time window of length l is then calculated:

$$\tilde{y}(t) = \frac{1}{l}\sum_{z=0}^{l-1} y(t-z) \tag{3.9}$$

A detrended version of the signal is obtained by the subtraction:

$$\varepsilon(i) = y(i) - \tilde{y}(i) \tag{3.10}$$

The resulting series is then divided in N_l disjoint sets of points of size l and a root mean-square function is obtained for each epoch ν via:

$$F_\nu(l) = \left\{ \frac{1}{l} \sum_{i=1}^{l} \varepsilon_\nu^2(i) \right\}^{\frac{1}{2}} \tag{3.11}$$

A generalised qth-order overall fluctuation function can be obtained from:

$$F_q(l) = \left\{ \frac{1}{N_l} \sum_{\nu=1}^{N_l} F_\nu(l)^q \right\}^{\frac{1}{q}} \quad q \neq 0 \tag{3.12}$$

and

$$\ln F_0(l) = \frac{1}{N_l} \sum_{\nu=1}^{N_l} \ln F_\nu(l) \quad \text{for} \quad q = 0 \tag{3.13}$$

It is possible to find a power-law relationship between $F^q(l)$ and the window length, or scale l by:

$$F^q(l) \propto l^{\alpha(q)} \tag{3.14}$$

The multifractal mass exponent introduced in equation 1.24 (Biswas et al., 2012) can be defined as:

$$\tau(q) = q\alpha(q) - D_f \tag{3.15}$$

where D_f is the fractal dimension of the support measure. For a single-channel time series, $D_f = 1$. The spectrum, $f(\alpha)$, can be obtained with a Legendre transform (Gu and Zhou, 2010).

$$\alpha(q) = \frac{d\tau(q)}{dq} \tag{3.16}$$

$$f(q) = q\alpha - \tau(q) \tag{3.17}$$

It is important to note that Legendre transform is known to cause problems in multifractal spectra derivations if some heterogeneities are present in the signal, as has been reported elsewhere (Chhabra and Jensen, 1989; Mukli et al., 2015).

3.2.2.2 MF-DFA

The Multifractal Detrended Fluctuation Analysis (MF-DFA) method is essentially a generalisation of the DFA approach (Ihlen, 2012; Kantelhardt et al., 2002). The time series is first rebuilt according to eq. 3.18. It is then divided into $N_l = \frac{N}{l}$ non-overlapping epochs ν of length l. The variance of the detrended series l is calculated as follows:

$$F_\nu^2(l) = \frac{1}{l}\sum_{k=1}^{n}(y((\nu-1)l+1) - y_\nu(k))^2 \tag{3.18}$$

where y_ν represents the fitting in the epoch ν obtained via linear regression. The overall q-th order fluctuation functions can be obtained as:

$$F_q(l) = \left\{\frac{1}{N_l}\sum_{\nu=1}^{N_l}(F_\nu^2(l))^{\frac{q}{2}}\right\}^{\frac{1}{q}} \tag{3.19}$$

A log-log plot of $F_q(l)$ versus l for different values of q should present a linear curve defined by the power law in equation 3.14. Similarly to the MF-DMA method,

the multifractal scaling exponent can be defined as in 3.15 and the spectrum $f(\alpha)$ can be determined in the same way as in the MF-DMA approach.

3.2.2.3 Chhabra-Jensen

Multifractal spectra can be obtained in a more direct way, without the need for the Legendre transform using the Chhabra-Jensen (CJ) method (Chhabra and Jensen, 1989; França et al., 2019; Murcio et al., 2015; Vidal-Vázquez et al., 2013; Paz Ferreiro et al., 2010; Paz-Ferreiro et al., 2010; Miranda et al., 2006; Zeleke and Si, 2006; Vázquez et al., 2008; Xu et al., 2017). Considering a time series as a distribution over time, the approach consists of calculating a family of generalised measures by covering the time series with windows. These are probabilistic measures with an emphasis factor q that accentuates different singularities depending on its value. More singular regions are emphasised by $q > 1$ whereas less singular regions will have a higher weight with $q < 1$ (Chhabra and Jensen, 1989). First, a family of generalised measures can be defined:

$$\mu_i(q,l) = \frac{P_i(l)^q}{\sum_j P_j(l)^q} \tag{3.20}$$

where $P_i(l)$ represents the cumulative probability of a window i. l corresponds to the size of the window in which the generalised measures are obtained. The window epochs are indexed by the variables i and j. Then the multifractal spectra can be obtained directly from:

$$\alpha(q) = \lim_{l \to 0} \frac{\sum_i \mu_i(q,l) \log P_i(l)}{\log l} \tag{3.21}$$

and

$$f(q) = \lim_{l \to 0} \frac{\sum_i \mu_i(q,l) \log \mu_i(q,l)}{\log l} \tag{3.22}$$

A numerical approximation to the equations above is provided by the measures $M\alpha$ and Mf functions in eq. 3.23 and 3.24.

$$M\alpha = \sum_i \mu_i(q,l) \log P_i(l) \tag{3.23}$$

$$Mf = \sum_i \mu_i(q,l) \log \mu_i(q,l) \tag{3.24}$$

α and $f(q)$ can then be obtained as the slopes by regressing these two measures against the scales l: $M\alpha \sim l$ and $Mf \sim l$.

The algorithmic summary of the Chhabra-Jensen method consists of the following steps:

- The algorithm has as input the time series, a range of q values to which the spectrum will be evaluated, and epoch length l that vary in a dyadic scale.
- The time series is divided into non-overlapping epochs of length l and the generalised measures are estimated according to equation 3.20.
- The measures $M\alpha$ and Mf are obtained from the generalised measures.
- α and $f(q)$ in eq. 3.21 and 3.22, respectively, are obtained with a linear regression procedure: $\log(M\alpha)$ is regressed against $\log(l)$ and $\log(Mf)$ is regressed against $\log(l)$, they give α and f respectively as the slopes.

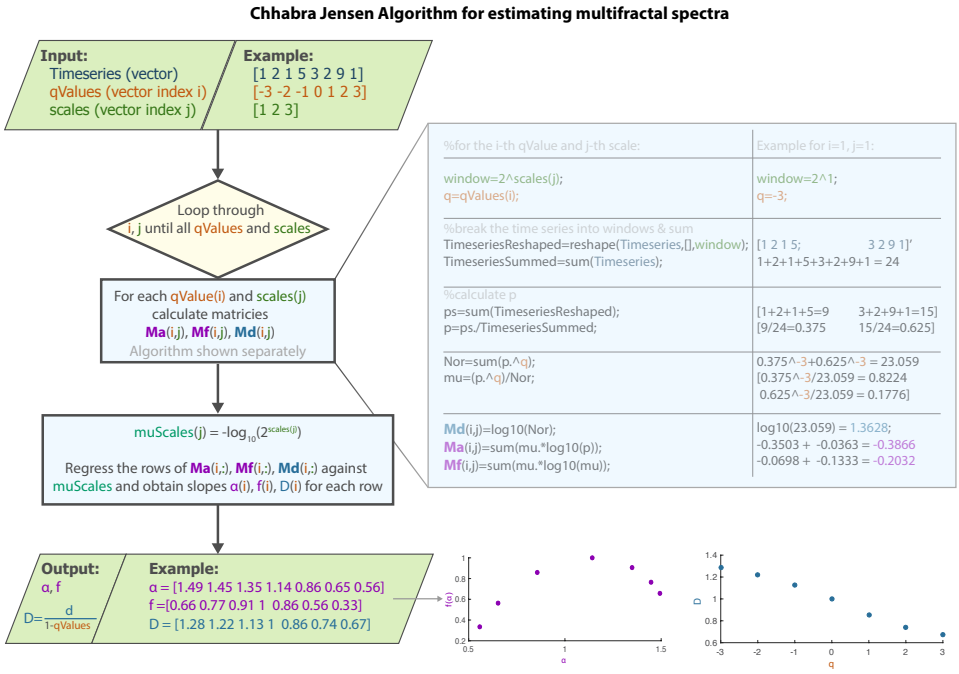

Figure 3.2: Flow-chart diagram of the algorithm for Chhabra-Jensen approach.

- A rejection criterion is also used, where all q exponent values with $R^2 < 0.9$ in the linear regression are not considered to avoid overfitting.

The code used in this study to calculate the multifractal spectrum is available at: `https://github.com/lucasfr/chhabra-jensen`. A flow-chart diagram of the algorithm is available in figure 3.2.

The direct estimation via the Chhabra-Jensen (CJ) equations makes the calculus of the multifractal spectra easier and more precise. The method presents an advantage, compared to others (such as MFDFA or MFDMA), for the precise and direct determination $f(\alpha)$ spectrum, without a Legendre transform (Chhabra and Jensen, 1989). The correct use of the method still depends, though, on the correct choice of the analysis parameters, as described below.

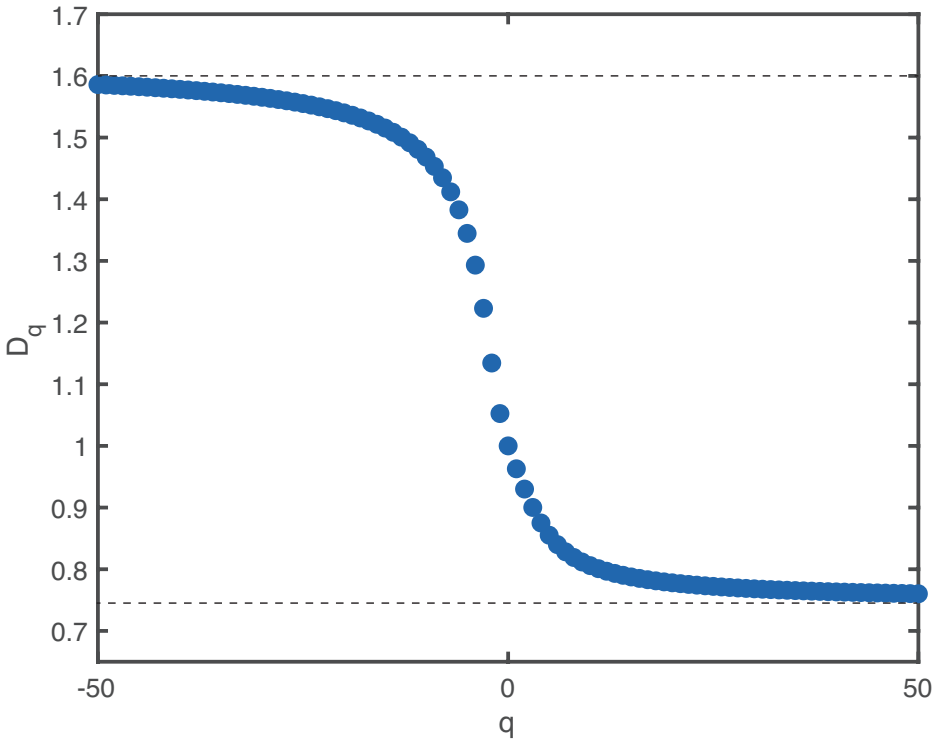

Figure 3.3: Convergence in generalised dimensions' spectrum. The dashed lines show the convergence value and the limits for convergence are, in this case, the values -50 and 50. The limits of convergence should be chosen as parameters for the multifractal analysis.

The first parameter, the range of q, can be defined based on the generalised dimensions plot, figure 3.3, i.e., the points at which the value of D_q converges.

The scales (l) range can be evaluated based on the scaling properties of the partition function, which should present a linear behaviour when on a logarithmic scale for all the q values, as shown in figure 3.4, for all the scales used in analysis. In this example, the last two points, for negative values of q, do not fit properly on the linear fit; these two scale sizes were therefore excluded from analysis.

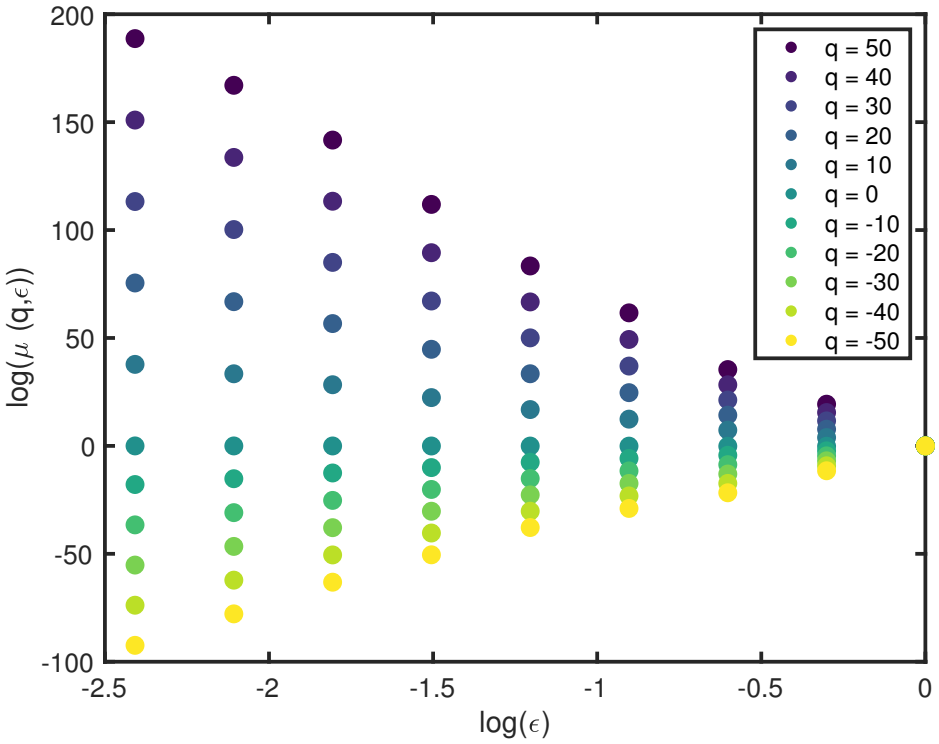

Figure 3.4: Partition function by scale value linearised by the logarithm method. The chosen scales must form a straight line. In this case, the two smallest scales would be removed.

3.2.3 Modelling series with scaling properties

3.2.3.1 Modelling monofractal data

To fully test methods of estimating the monofractal dimension from signals, time series that are known to be fractal were computationally produced. Fractional Brownian motion (fBm) (Mandelbrot and Van Ness, 1968) profiles/time series were generated using a novel modified version of the Wood-Chan or circulant embedding approach (Kroese and Botev, 2015; Shevchenko, 2015), that allows us to change the variance of the signal over time, in order to evaluate its influence on the fractal estimation. This novel modulated fBM approach uses a modulating function, $M(t)$, which produces a signal that has an amplitude varying over time.

A fractional Brownian motion is a continuous zero-mean Gaussian process that can be described by the function in equation 3.26 (Kroese and Botev, 2015). In order to simulate such properties, it is sufficient to simulate the increments on the profile, as described in equation 3.25.

$$\xi_1 = B_1^H, \xi_2 = B_2^H - B_1^H, \xi_N = B_N^H - B_{N-1}^H \tag{3.25}$$

These values form a stationary sequence of standard Gaussian variables with a covariance described in equation 3.26. The vector $\xi = (\xi_1, ..., \xi_N)^T$, composed by the simulated increments, is called fractional Gaussian noise and is a centred Gaussian array with covariance matrix Ω (Shevchenko, 2015).

$$\rho_H(n) = E[\xi_1 \xi_{n+1}] = \frac{1}{2}\left((n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right), \quad n \geq 1 \tag{3.26}$$

$$\Omega = \mathrm{Cov}(\xi) = \begin{pmatrix} 1 & \rho_H(1) & \rho_H(2) & \cdots & \rho_H(N-2) & \rho_H(N-1) \\ \rho_H(1) & 1 & \rho_H(1) & \cdots & \rho_H(N-3) & \rho_H(N-2) \\ \rho_H(2) & \rho_H(1) & 1 & \cdots & \rho_H(N-4) & \rho_H(N-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_H(N-2) & \rho_H(N-3) & \rho_H(N-4) & \cdots & 1 & \rho_H(1) \\ \rho_H(N-1) & \rho_H(N-2) & \rho_H(N-3) & \cdots & \rho_H(1) & 1 \end{pmatrix}$$

In order to solve this problem, one can, alternatively, convert the covariance matrix into a circulant matrix. The circular matrix (Σ) has a known eigenvalues structure, making it easier to solve the problem. Taking the relationship $M = 2(N-1)$, a circular matrix Σ is defined based on the relations in equation 3.27.

$$\begin{aligned} & c_0 = 1, \\ & c_k = \begin{cases} \rho_H(n), & n = 1,2,...,N-1 \\ \rho_H(M-n), & n = N,N+1,...,M-1 \end{cases} \end{aligned} \tag{3.27}$$

$$\Sigma = circ(c_0, c_1, ..., c_{M-1}) = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{M-2} & c_{M-1} \\ c_{M-1} & c_0 & c_1 & \cdots & c_{M-3} & c_{M-2} \\ c_{M-2} & c_{M-1} & c_0 & \cdots & c_{M-4} & c_{M-3} \\ \cdots & \cdots & \cdots & \ddots & \cdots & \cdots \\ c_2 & c_3 & c_4 & \cdots & c_0 & c_1 \\ c_1 & c_2 & c_3 & \cdots & c_{M-1} & c_0 \end{pmatrix}$$

The aim is a factorisation described in the equation 3.28 (Kroese and Botev, 2015). The increments in the simulated Brownian motion time series can be derived from the eigenvalues $\lambda = (\lambda_1, ..., \lambda_{4N^2})^T$ arranged as a $2N$x$2N$ matrix Λ. The variable P is the Kronecker product of two discreet Fourier transform matrices given by $P = F \bigotimes F$ (Kroese and Botev, 2015).

$$\Sigma = P^* \Lambda P \tag{3.28}$$

$$F_{jk} = \frac{1}{\sqrt{2N}} e^{-2\pi i jk/2N}, \quad \text{where} \quad j,k = 0,1,...2N-1 \tag{3.29}$$

The algorithm used to derive the fBm time series is available at `https://github.com/lucasfr/ModfBm` and can be summarised in the following steps:

- First, the covariance matrix is calculated (Ω) (step 1).
- Then a $2N$x $2N$ circulant matrix (Σ) is built (step 2).
- For this operation, one just needs the first row of the matrix Σ (step 3).
- The eigenvalues of matrix Λ can then be obtained via Fourier transformation. Both real and complex parts will present a Gaussian profile, however, it is enough to consider the real part of the eigenvalues (step 4).
- The following step consists of multiplying the output of step 4 by random complex numbers and applying the inverse Fourier transformation (step 5).
- This step was modified, and an element of a function M has been added, such element will be multiplied by these complex random values before applying the inverse Fourier transformation. From now on, this modified version will be denoted as Modulated fractional Brownian motion (ModfBm).
- The last step consists of accumulating the values generated by the previous operation and multiplying by a constant scaled by the Hurst exponent H. The generated time series should present scale properties according to the chosen H exponent and, for the ModfBm, feature changes on its standard deviation.

A time series of 1,843,200 points (1800 windows of 1024 points) was generated, corresponding to a recording of an hour duration with a sampling rate of 512 Hz — similar to the clinical intracranial EEG segment from the patient NHNN1, used on the analysis of real EEG data. The time series was simulated with a Hurst exponent $H = 0.7$, the value was chosen due to its persistent features, i.e. it generates a time series with memory. Additionally, a modulating function M, for every 2 seconds window w, described by equation 3.30 and shown in Fig. 3.7(A) was used to simulate the Modulated fBm.

$$M(w) = \begin{cases} 1, & w < 450 \\ 1 + (w - 449)0.01, & 450 \leq w \leq 900 \\ 1, & w > 900 \end{cases} \tag{3.30}$$

Note that there are alternative methods to generate monofractal time series (Eke et al., 2002; Mukli et al., 2015; Nagy et al., 2017; Davies and Harte, 1987). However, as the aim was not to compare generative models of monofractal time series, but rather simply to demonstrate that the effects observed in EEG signals could be more general. The above-mentioned chosen approach serves as an example demonstration.

3.2.3.2 Modelling multifractal data

This section presents a computational procedure to generate time series that are known to be multifractal based on the p-model, which was developed to reproduce features observed in turbulence experiments known to have multifractal properties

(Meneveau and Sreenivasan, 1987). This is a simple model, having a single fraction p_1 as its only input and is often mentioned in the literature (Pechlivanidis and Arheimer, 2015; Consolini et al., 1996; Davis et al., 1997; Kestener and Arneodo, 2003; Lipa and Buschbeck, 1989; Meneveau and Sreenivasan, 1987, 1991; She and Leveque, 1994; Sreenivasan and Antonia, 1997; Zhou, 2008).

Briefly, the algorithm works as follows: from an interval of length L and height $\varepsilon_L = c$ (is a constant), two segments of length L/2 are created. Based on the input parameter p_1, it is possible to establish a second fraction in which a second parameter will be given by $p_2 = 1 - p_1$. The heights of each interval will thus be given by $y = 2p_{1L}$, and $y = 2p_{2L}$, respectively. This procedure is repeated for each remaining segment, selecting left or right for p_1 randomly (Meneveau and Sreenivasan, 1987).

The p-model was employed in the simulation of a time series profile with multifractal properties to be evaluated by different estimation methods. It was generated with a code available at `http://www2.meteo.uni-bonn.de/staff/venema/themes/surrogates/pmodel/` (Davis et al., 1997; Venema et al., 2006). Using this algorithm, time series were generated to evaluate the performance of different multifractal estimators with $p = 0.4$. The value was rounded (for simplicity) from the figure used elsewhere ($p = 0.375$) (Davis et al., 1997).

3.2.4 Data

In order to evaluate the effect of EEG signal variance change on multifractal properties, one specific recording where the signal variance changes dramatically over time was chosen. Such a recording comes from one patient (male, 28 years old,

temporal lobe epilepsy, recorded at the National Hospital for Neurology and Neurosurgery (NHNN) ((UCLH NHS Foundation Trust, Queen Square, London, UK), patient ID: NHNN1) near one seizure event. A 60-minute recording segment around the epileptic seizure was used for this analysis. The EEG recording was sampled at 512 Hz.

The seizure onset and offset were marked by expert clinicians, independent of this research project. Note that this segment specifically was used due to the dramatic change in signal variance, which actually occurs before the seizure and evolves over about 15 minutes. This chapter does not make conclusions about the seizure event itself at this stage, but rather uses this recording as an example to illustrate a technical point about multifractal property estimation from EEG.

In order to show that the results obtained in this study are reproduced in other seizures, a further two segments of the same patient and channels were evaluated with the developed methodology. The first of these two segments consisted of an interictal EEG recording of 8380 seconds with no seizures — featuring interictal epileptiform discharges (IEDs), and the second comprises another recording of 9376 seconds containing one epileptic seizure.

For NHNN data, the subjects gave informed written consent, and the study was approved by the Joint Research Ethics Committee of the NHNN (UCLH NHS Foundation Trust) and UCL Queen Square Institute of Neurology, Queen Square, London, UK.

3.2.5 A sigmoid approach

Unless stated otherwise, the same pre-processing and analysis parameters have been applied to the computationally generated time series and the human EEG recordings. Also, the fractal and multifractal estimations were performed on 1024-sample epochs. This chapter is specifically focused on the effect of signal variance on the (multi) fractal estimation, and therefore the results for signal subjected to a standardisation procedure were compared, the novel approach is performed as follows:

$$x' = \frac{X - < X >}{s} \tag{3.31}$$

where $< X >$ is the epoch mean and s the epoch standard deviation of the time series X, resulting in a time series with zero mean and unit standard deviation. The Chhabra-Jensen method requires as input a distribution function over the domain of positive real numbers, which is incompatible with EEG data which contain positive and negative values. Hence, this chapter proposes the use of a sigmoid-transformation here (equation 3.32) to map the time series onto positive values, in order to apply the Chhabra-Jensen method. Example sigmoid functions and correspondingly transformed EEG signal are shown in Fig. 3.5.

$$\sigma(X) = \frac{1}{1 + e^{vX}} \tag{3.32}$$

The parameter v was chosen based on its effect on the estimated multifractal width for three types of time series: icEEG (NHNN1 - channel 1), surrogate EEG (temporally shuffled values of the original time series from NHNN1 - channel 1)

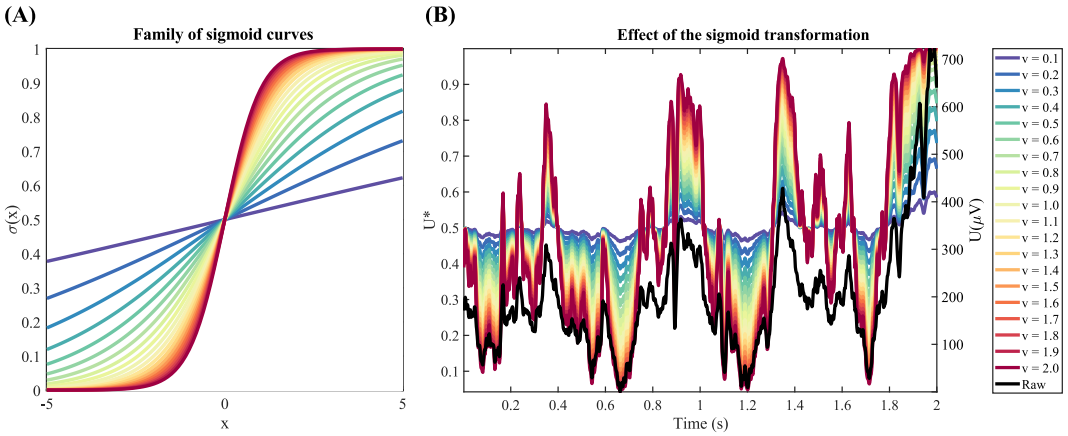

Figure 3.5: Impact of sigmoid transformation on the signal. This function maps the original (raw) time series into a sequence of elements ranging in between 0 and 1, allowing the use of the Chhabra-Jensen method on recordings originally containing negative values. (A) Sigmoid curves with different values of v. This parameter defines how shallow/steep the sigmoid curve will be. (B) Effect of the sigmoid transform on an epoch extracted from an intracranial EEG recording. Smaller values of the parameter v tend to flatten the curve.

and a simulated random series (with the same mean and variance), across the range $v = [0.1, 2.0]$ in steps of 0.1. To find the optimal value for the parameter v, it was necessary to balance the trade-off between the three series in terms of presenting the most distinct $\Delta\alpha$ values (Fig. 3.6 (A)), while showing minimum distortion of the recording, or maximum correlation with the original time series (Fig. 3.6 (B)). The value $v = 1$ was chosen as an acceptable trade-off point.

3.2.6 Multifractal estimation stability

In order to evaluate the stability of multifractal properties in time, a time series exhibiting stable multifractal properties over time was generated; such a series was computationally generated using the p-Model. The time series was then evaluated using an epoch-based approach with the three estimators: MF-DFA, MF-DMA, and Chhabra-Jensen. The stability of the estimator was assessed as the temporal variability of its output.

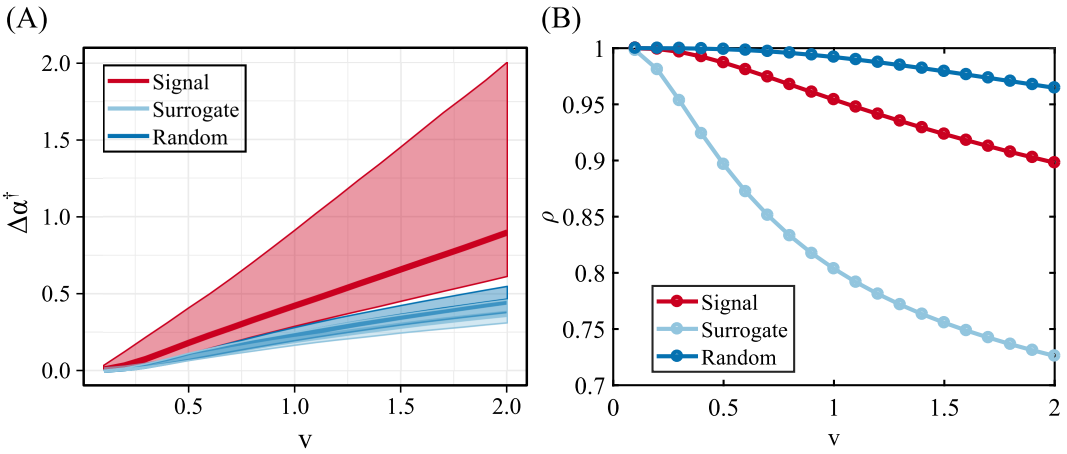

Figure 3.6: Assessment of parameters for sigmoid mapping function. A) Variability of $\Delta\alpha$ (0.1, 0.5 and 0.9 quantiles) as a function of the parameter v. B) Pearson correlation of the original series and the mapped one for the three types of signal. Based on the optimisation criterion — maximum difference and minimum distortion (maximum correlation) — the value chosen was $v = 1$.

3.2.7 Feature redundancy

To assess whether the chosen multifractal metrics contribute any non-redundant features about the signal in addition to more established signal metrics, human EEG signals recorded intracranially were analysed.

Again, an epoch-based approach was employed, and the multifractal metrics were compared to a number of conventional signal metrics (mean, standard deviation, line length, band power) on each epoch. The similarity between signal features was evaluated using Pearson correlation and Mutual Information (Guyon and Elisseeff, 2003; Ince et al., 2017) (the code is available at `https://github.com/robince/gcmi`). Furthermore, monofractal metrics were also included in this comparison, to determine advantages in applying multifractal over monofractal approaches.

Finally, to compare multifractal properties to classical EEG frequency band power, the following definitions for the classical EEG frequency bands were con-

sidered: δ (0.5-4 Hz), θ (4-8 Hz), α (8-15 Hz), β (15-30 Hz), and γ (30-60 Hz).

3.3 Results

This study evaluated the relationship between monofractal measures and signal variance using a simulated time series based on fractional Brownian motion (fBm), where its signal variance is modulated by a modified ramp function. The modulation function is shown in Fig. 3.7(A) and resulting time series in Fig. 3.7(B). The standard deviation of the generated time series indeed tracks the shape of the modulating function (Fig. 3.7(C)).

The monofractal dimension of this simulated signal was estimated using two standard methods: Higuchi and DFA. It is possible to observe that both methods appear to be affected by the changing signal variance (Fig. 3.7(D), (F)). Furthermore, this effect persists after epoch-based standardisation (Fig. 3.7(E), (G)): the monofractal properties and standard deviation correlate with $\rho = 1.00$ and $\rho = 0.99$ for the Higuchi and DFA methods, respectively. A similar effect was observed for a real icEEG recording that contained changes in signal variance over time (Fig. 3.8).

In conclusion, monofractal properties derived for each epoch from DFA and Higuchi methods (with, or without signal standardisation) correlate highly with the signal standard deviation of the epoch. Therefore, in epoch-based approaches (e.g. for applications such as detecting or predicting epileptic seizures), the monofractal properties cannot be regarded as a useful, non-redundant EEG feature when standard deviation of the epoch is readily available.

This text will now denote the epoch-wise estimates of multifractal width $\Delta\alpha$

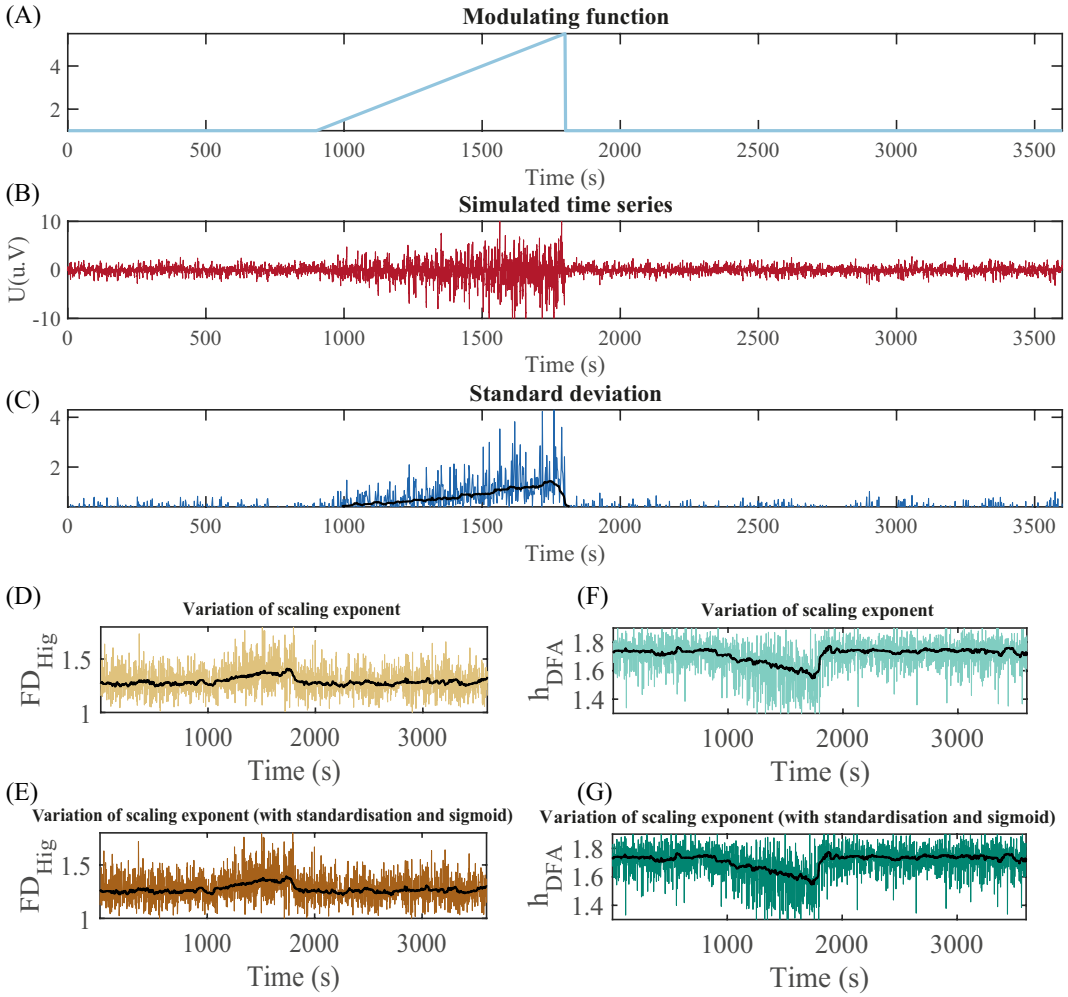

Figure 3.7: Impact of the signal standard deviation on monofractal scaling exponent estimation. (A) Modulation of the standard deviation of the time series over time; (B) Time series simulated using fractional Brownian motion based on modulation in (A); (C) Standard deviation of the simulated signal in (B). (D) Monofractal dimension obtained with the Higuchi method from signal without epoch-based standardisation. (E) Monofractal dimension obtained with the Higuchi method from epoch-based standardised signal. (F) Hurst exponent obtained with the DFA method from signal without epoch-based standardisation. (G) Hurst exponent obtained with the DFA method from epoch-based standardised signal.

and height Δf (and $\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$ for the epoch-based standardised measures). This experiment was designed to assess the reliability of the different multifractal estimation methods over time. In other words, if the multifractal properties of the time series remain constant over different epochs, then one would expect the multifractal

estimation method to show the same output over these different epochs. Note that the accuracy of these methods (i.e. the method outputting the expected multifractal measures of a predefined multifractal object with known multifractal properties) has been demonstrated elsewhere (Chhabra and Jensen, 1989; Gu and Zhou, 2010; Kantelhardt et al., 2002).

Figure 3.9 shows the simulated signal by the p-model and the outputs of the three multifractal spectral estimation methods. In all cases, the magnitude of $\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$ were clearly different from zero. The ($\Delta\alpha^{\dagger}$, $\Delta f^{\dagger}$) output variances over time for the MF-DFA, MF-DMA, and Chhabra-Jensen estimation methods were: (0.018, 0.18), (4.17e-4, 0.0028) and (2.3e-30, 6.5e-30), respectively.

In addition, the MF-DFA output violated the theoretical topological limit of $\Delta f^{\dagger} = 1$, again indicating problems in the MF-DFA method, potentially due to the inversion of multifractal spectrum (Mukli et al., 2015). As the Chhabra-Jensen method shows the lowest variance over time (i.e. most reliable/stable), it will be the multifractal analysis method of choice for the remainder of this work.

Next, the relationship between multifractal signal properties and other widely used conventional EEG measures (such as signal variance) was evaluated. Figure 3.10 shows the results of the multifractal spectrum and conventional measures in comparison. The pattern of multifractal spectrum width without epoch-based standardisation ($\Delta\alpha$) reflects the signal variance closely, in contrast to the estimate for the epoch-based standardised signal ($\Delta\alpha^{\dagger}$). A similar outcome is obtained from another segment with an epilpetic seizure in the same patient, as seen in figure 3.12. Finally, signal line length also shows a very different temporal profile from $\Delta\alpha^{\dagger}$.

The variation of Δf and $\Delta f^{\dagger}$ metrics is available in Fig. 3.11.

Figure 3.13 shows the quantification of similarities of the signals in Fig. 3.10 and 3.11 through a correlation analysis. In summary, a high degree of correlation is present between the signal standard deviation, multifractal spectrum width ($\Delta\alpha$), and detrended fluctuation analysis (monofractal approach) both with and without epoch-based standardisation. The standardisation reduces the correlation between $\Delta\alpha$ and the standard deviation from $\rho = 0.86$ (for $\Delta\alpha$) to $\rho = 0.14$ (for $\Delta\alpha^{\dagger}$). Also note that $\Delta\alpha$ is highly correlated with *DFA* and *DFA*† estimates ($\rho = 0.74$ and $\rho = 0.71$, respectively) while it is markedly reduced for $\Delta\alpha^{\dagger}$ ($|\rho| < 0.3$). The analysis based on mutual information (Ince et al., 2017) rather than correlation showed a similar pattern (Fig. 3.14).

The relationships of the multifractal properties and specific EEG frequency band power are shown in Fig 3.15. In summary, the correlation values between the multifractal measures $\Delta\alpha^{\dagger}$, $\Delta f^{\dagger}$ and signal power in the classical EEG bands are low ($|\rho| < 0.3$).

3.4 Discussion

In this study, the monofractal and multifractal properties of human EEG recordings and simulated data were explored to test the performance of fractal property estimation methods. Although mono- and multi-fractal approaches have been widely employed in the study of physiological signals in humans (Costa et al., 2017; França et al., 2019; Hu et al., 2004, 2009; Ivanov et al., 1999; Stanley et al., 1999), it has been demonstrated that the non-linear measures, such as monofractal dimension,

may be capturing a similar signal feature as the signal variance (McSharry et al., 2003).

When using standardisation to remove the effect of signal variance, it was shown that multifractal measures (estimated by the Chhabra-Jensen method) capture features not contained in widely used conventional signal measures, making it a viable feature for machine learning in clinical EEG applications.

One of the key observations is that monofractal estimators are tightly correlated with signal variance even following epoch-wise standardisation, whereas multifractal properties following epoch-wise standardisation are no longer tightly correlated with signal variance. This non-intuitive observation has not been reported before.

To interpret this observation, it is worth noting the relationship between monofractal and multifractal analyses. Essentially, in multifractal analysis, at the point for which $q = 2$, the corresponding $f(\alpha)$ is the so-called correlation dimension, which is an alternative way of estimating the monofractal dimension (Murcio et al., 2015). The relationship between monofractal dimension and signal variance has been established and explained before (Cannon et al., 1997).

By the same token, signal variance also affects higher statistical moments ($q > 2$ or $q < 2$). However, when analysing the exact effect of variance on the multifractal spectrum (Fig. 3.16), it was observed that the variance particularly impacts the multifractal spectrum width and height but maintains an almost constant value of $f(\alpha)$ for $q = 2$. This explains why epoch-wise standardisation does not impact monofractal dimension but does impact multifractal spectrum width and height.

The mono- and multifractal properties being investigated here are essentially

describing different properties of the multifractal spectrum. Note that through the standardisation procedure, the approach does not abolish multifractality, but only its dependence on signal variance. Future work will need to show mathematically the exact reason for this observation, although intuitively it is understandable that the standardisation procedure (a linear transformation of the signal) changes the $q = 2$ moment least and affects higher moments more.

It has also been observed that the Chhabra-Jensen method is the most reliable out of the three methods. As was pointed out in the original publication (Chhabra and Jensen, 1989), this is most likely due to the fact that the Chhabra-Jensen method avoids a Legendre transform that the other methods require. The Legendre transformation requires smoothing of the D_q curve and can lead to errors. For further advantages of the Chhabra-Jensen method, the reader is referred to the original publication (Chhabra and Jensen, 1989). A recent development, FMF method (Mukli et al., 2015; Nagy et al., 2017), is an alternative to the approach proposed in this study.

Previous studies reported that the brain is characterised by critical dynamics (Chialvo, 2010, 2012; Eguíluz et al., 2005; Racz et al., 2018). This characteristic, found from microscopic spatial scales (such as neuronal networks) (Beggs and Plenz, 2003, 2004) to the global brain structure (Eguíluz et al., 2005), is thought to facilitate the storage and processing of information. It has been further suggested that more than one scaling exponent would be necessary to properly characterise the brain's critical dynamics (Ciuciu, 2012; Fraiman and Chialvo, 2012; Ihlen and Vereijken, 2010; Papo, 2014; Papo et al., 2017; Racz et al., 2018; Suck-

ling et al., 2008; Zhang et al., 2015; Zorick and Mandelkern, 2013), as departures from the power-law pattern have been frequently observed in brain signals. Hence, it has been proposed that using additional, higher-order statistical moments can better characterise such data (Fraiman and Chialvo, 2012). This work contributes a complementary observation: while monofractal measures of EEG appeared to essentially follow the slow changes of signal variance, multifractal characterisation is capable of revealing new information.

In terms of generative processes that can produce monofractal properties, it has been suggested that a property called Self-Organised Criticality (SOC) (Bak et al., 1987) may play an essential role. SOC describes the capacity of a system to evolve naturally into a critical state (a state in which a minimum perturbation could lead to events of all sizes). Such phenomena display power-law distributions and fractal properties as signatures (Bak and Paczuski, 1995). An example process that displays SOC is the so-called single avalanche or Bak-Tang-Wiesenfeld model (also known as Abelian sandpile model) (Bak et al., 1987).

SOC behaviour has been linked to physiological control mechanisms, such as in human heart rate variability (Goldberger et al., 2002). Similar to SOC, a related regime — termed non-classical SOC — is thought to give rise to multifractal properties (Lovejoy and Schertzer, 2007). The analysis and understanding of the non-classical SOC is, however, still under development.

In this context, the proposed multifractal spectral analyses of human EEG data suggest that cerebral phenomena should not be modelled by a single avalanche model (classical SOC), in agreement with findings in a previous study (Fraiman

and Chialvo, 2012). Moreover, it is hypothesised that brain dynamics are non-ergodic (Bianco et al., 2007), i.e. display preferential states and depend on previous states (Papo, 2014), which are all properties of multifractal processes (Lovejoy and Schertzer, 2007). Thus, multifractal analyses could provide a new paradigm for studying brain function and structure, as previously suggested in other studies of normal (Ciuciu, 2012; Ihlen and Vereijken, 2010; Papo, 2014; Papo et al., 2017; Racz et al., 2018; Suckling et al., 2008; Zorick and Mandelkern, 2013) and pathological brain activity (Zhang et al., 2015).

Furthermore, generative processes displaying multifractal properties could help our understanding of the observed multifractal changes on a mechanistic level. It is important to emphasise that the conclusions from this work are drawn on the basis that slow changes in signal fractal features can be captured by using an epoch-wise feature extraction procedure. It is also evident from a feature redundancy perspective that there is a suggestion for the need of multifractal approaches over monofractal measures.

This work does not dispute the usefulness of monofractal measures in other general applications. In this work, a feature selection procedure was performed using correlation and mutual information (Guyon and Elisseeff, 2003). Then the differences between signal features were compared on an epoch-wise basis. Feature selection is crucial to obtain faster and cost-effective models and avoids overfitting of the available data. It might also help achieving a deeper insight into the nature of the studied phenomena (Blum and Langley, 1997; Guyon and Elisseeff, 2003; Liu et al., 1998; Liu and Yu, 2005; Saeys et al., 2007).

Furthermore, the slow temporal changes in multifractal dynamics need to be characterised in a systematic way. Using epileptic seizures as an example, Fig. 3.10 shows that dramatic changes in multifractal properties can sometimes be seen before an epileptic seizure. This observation requires further investigation to address questions such as: are all epileptic seizures characterised by pre-ictal changes in multifractal properties? Do other physiological processes, such as sleep, influence this finding? To answer these questions, a large scale future prospective study will most likely be necessary. Finally, it is well recognised that epileptic seizures are spatio-temporal processes (Wang et al., 2014, 2017), and the current approach of only focusing in the temporal aspect in one location will need to be expanded. Data-driven unsupervised approaches, such as dimensionality reduction, may help summarise spatial aspects. Additionally, the challenge will be to develop a spatio-temporal multifractal analysis approach that can also deal with the challenges of low spatial sampling resolution.

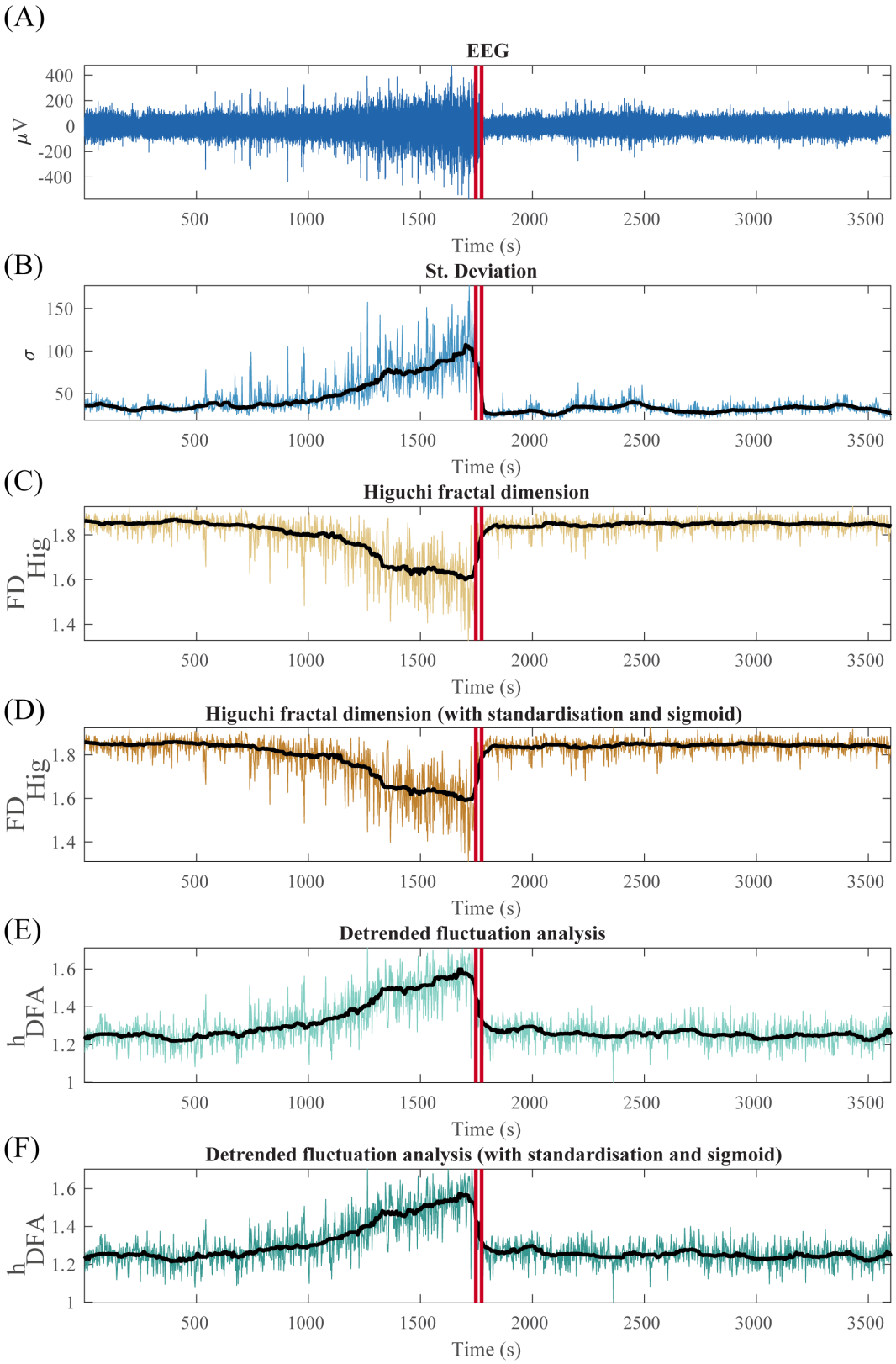

Figure 3.8: Monofractal analysis of seizure segment. Both Higuchi and DFA approaches exhibits similar patterns to the standard deviation of the signal, in accordance with the simulated data presented on the main text of this article. The signals from intracranial EEG shows a $H = 1.2$ (non-stationary) for interictal segments in patients with epilepsy.

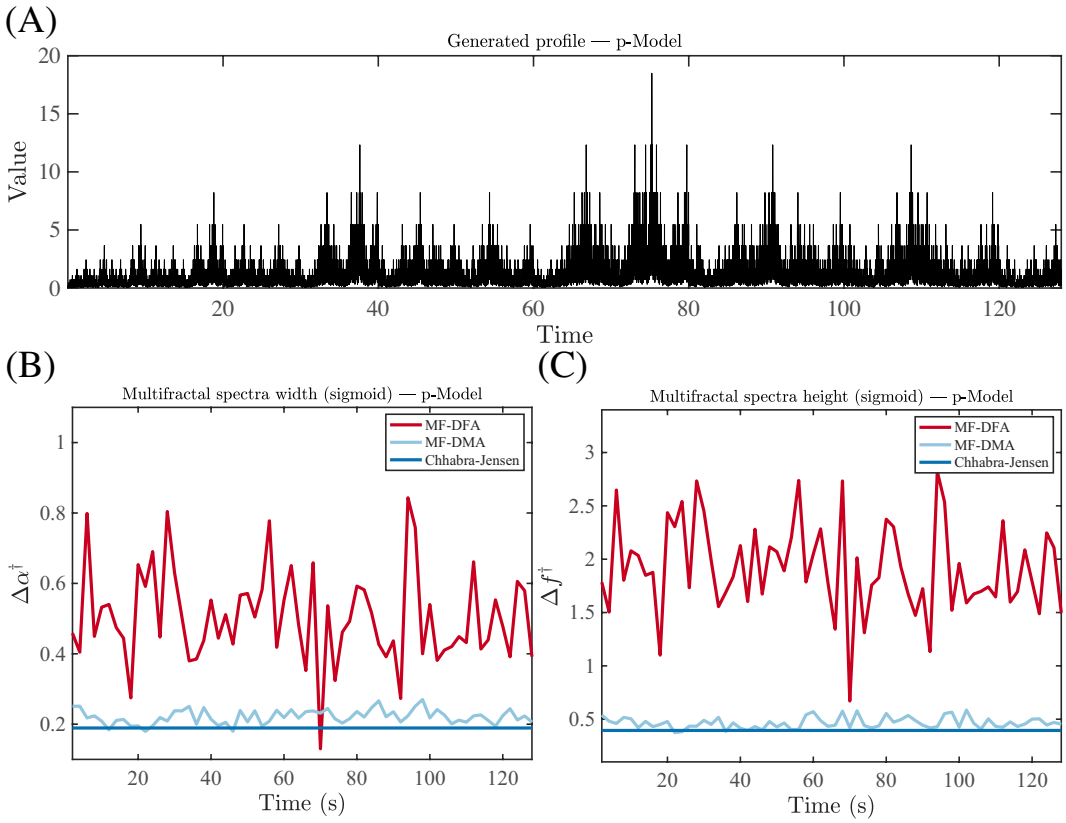

Figure 3.9: Comparison of three multifractal spectrum estimation methods (MF-DFA, MF-DMA, and Chhabra-Jensen) for p-Model simulated time series. (A) Time series simulated for $p = 0.4$. (B) Estimated multifractal spectra width $\Delta\alpha^{\dagger}$ and (C) height $\Delta f^{\dagger}$.

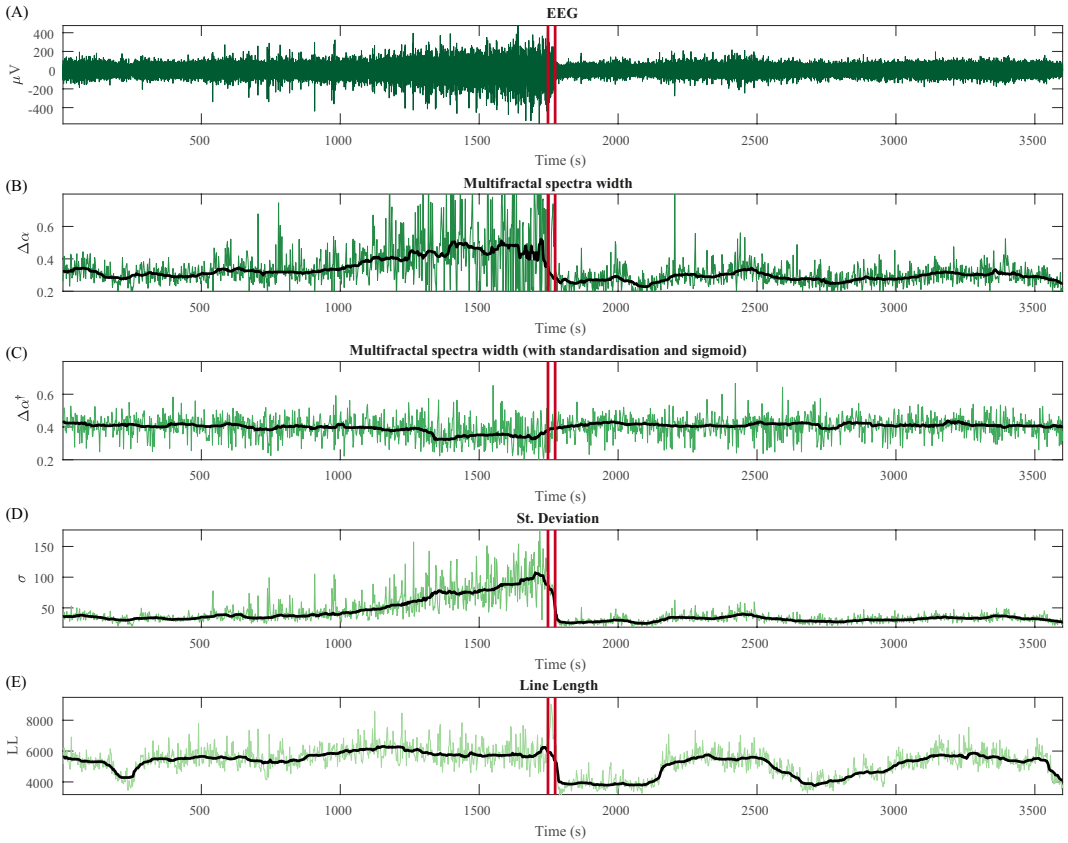

Figure 3.10: Temporal dynamics of multifractal spectrum width compared with conventional measures for human intracranial EEG. (A) Intracranial EEG segment containing a seizure (onset and offset marked with red vertical lines). Note that this recording was chosen because it showed a dramatic change in signal variance during non-seizure periods, not because of any seizure related properties. (B) Variation of multifractal spectrum width without epoch-wise standardisation ($\Delta\alpha$). (C) Multifractal spectrum width based on epoch-wise standardised time series ($\Delta\alpha^{\dagger}$). (D) Standard deviation in each epoch. (E) Line length in each epoch. Black line: moving average of each measure.

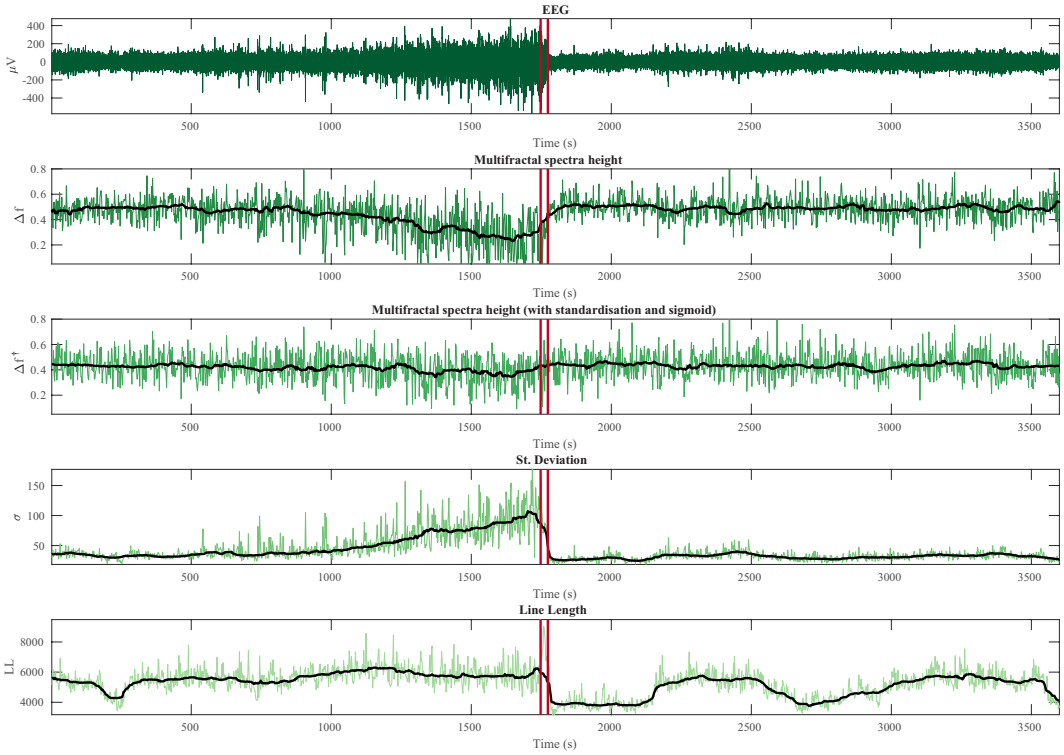

Figure 3.11: Multifractal spectrum height and other signal property changes over time. (A) A single channel intracranial EEG time series segment containing one seizure in the patient NHNN1 (channel 1). The seizure onset and offset are marked by red lines. (B) Variation of multifractal spectrum height (Δf) estimated on epochs of the EEG segment in (A). (C) Epoch-wise normalised multifractal spectrum height ($\Delta f^{\dagger}$). (D) Standard deviation of the time series for each epoch. (E) Line length of the signal for each epoch.

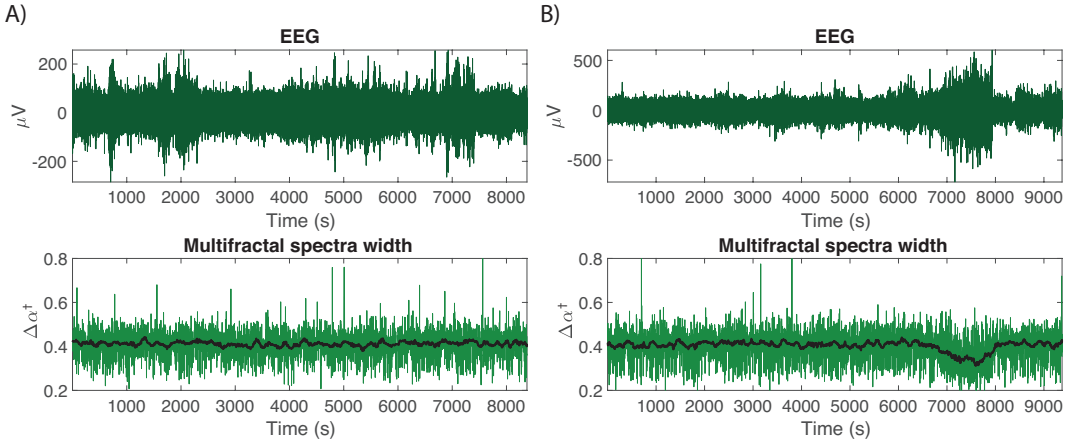

Figure 3.12: Temporal dynamics of multifractal spectrum width in two additional human intracranial EEG recordings. (A) Intracranial EEG segment with interictal activity, including interictal epileptiform discharges (IEDs). The analysis results in multifractal spectra width ($\Delta\alpha^{\dagger}$) with constant average over time. (B) Intracranial EEG segment containing one epileptic seizure. The multifractal spectra width ($\Delta\alpha^{\dagger}$) analysis reports a similar result to the segment shown in 3.10 — with a drop in multifractal spectra width prior to the seizure.

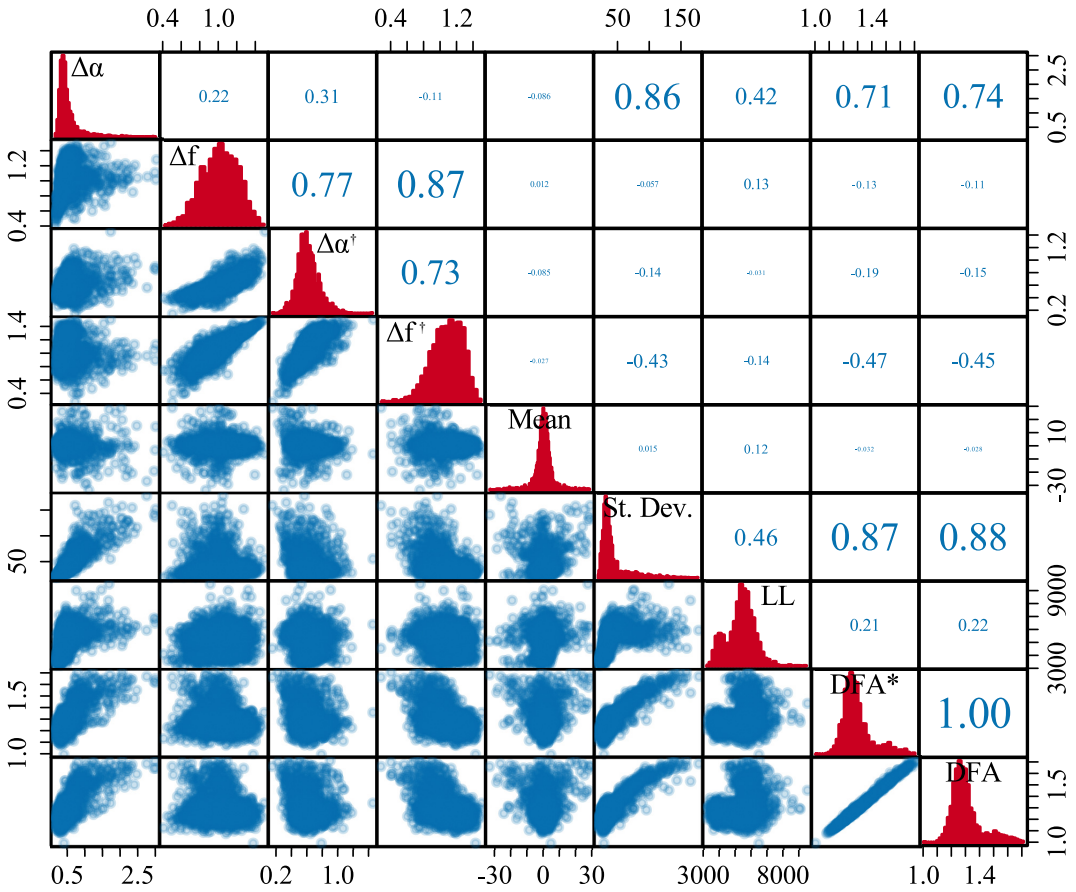

Figure 3.13: Correlation between multifractal spectrum and conventional EEG measures for human icEEG data (from Figure 3.10). The diagonal of the matrix shows the distribution for each measure across epochs. The lower triangle contains the scatter plots for each pair of measures across epochs. The upper triangle shows the Pearson correlation value for each pair of measure, where the size of the font additionally corresponds to the correlation coefficient to provide an additional visual cue.

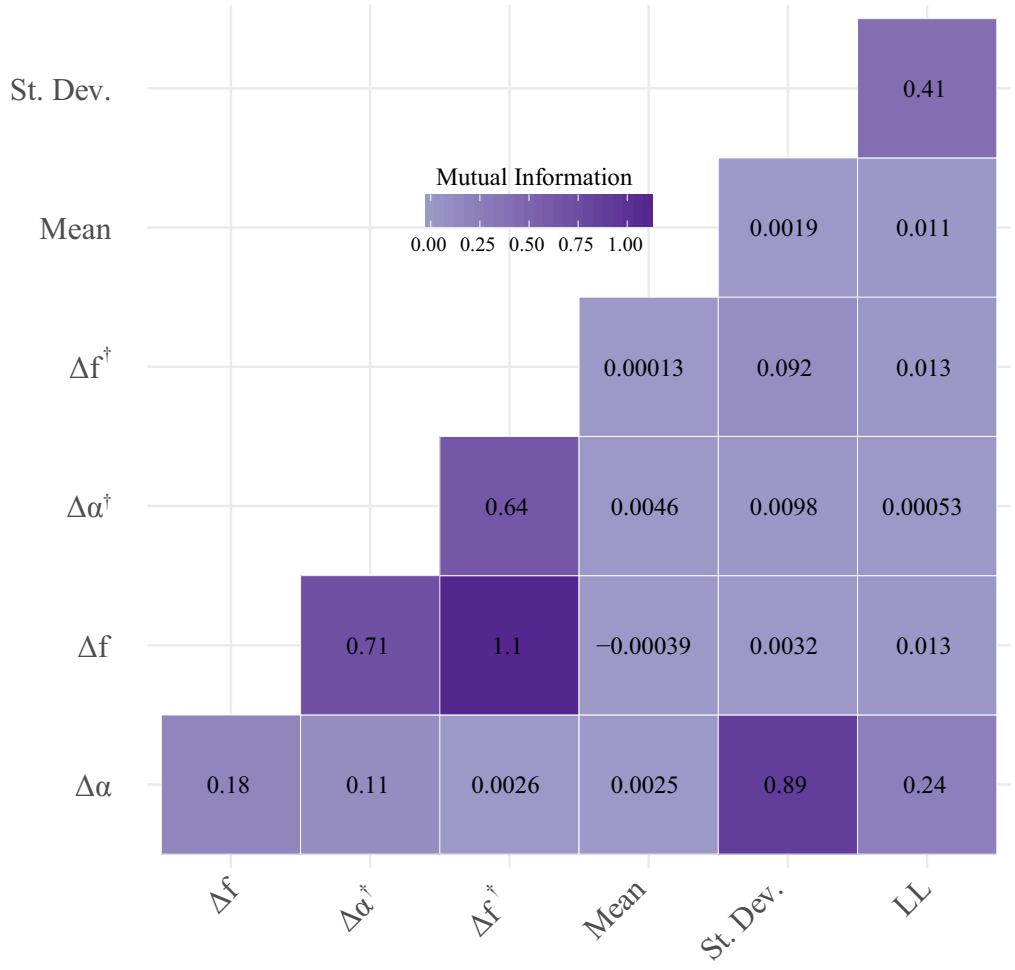

Figure 3.14: Mutual information of different measures. The MI measure shows higher values for the pair $\Delta\alpha$/St. Dev. than to $\Delta\alpha^{\dagger}1$/St. Dev., similar to the result in Fig. 3.13.

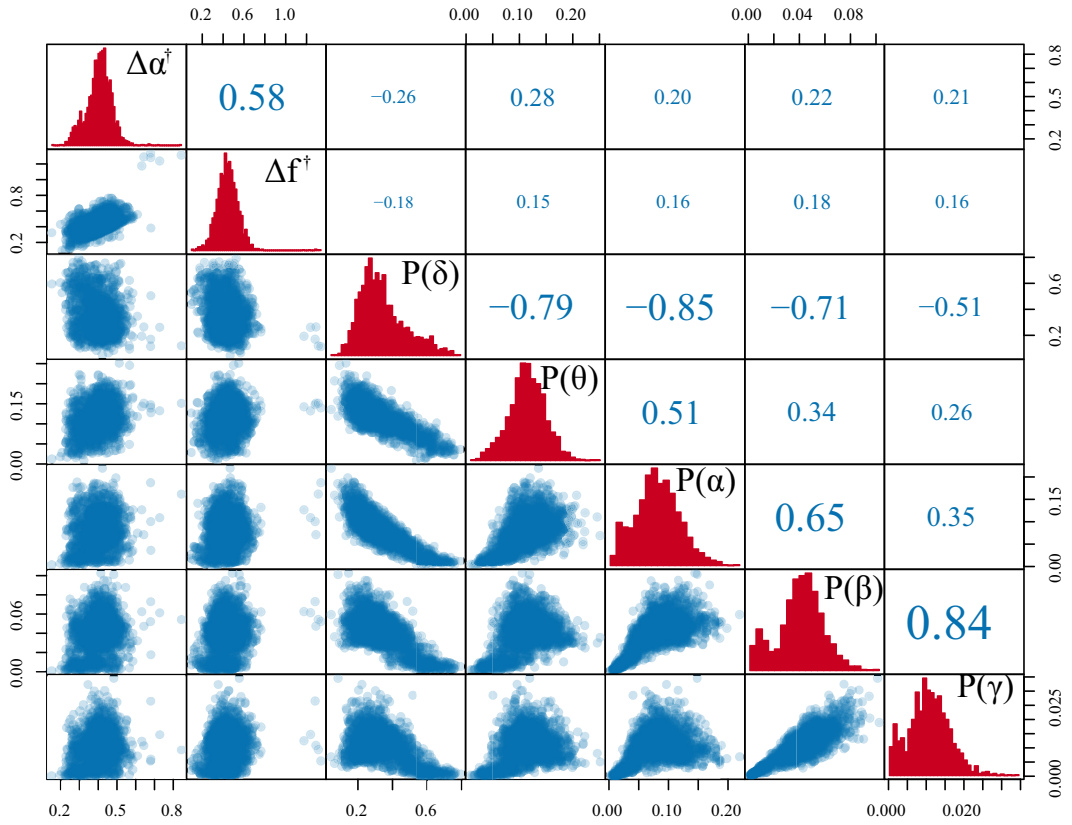

Figure 3.15: Comparison of multifractal measures with classical spectral band power. Scatter plot matrix comparing both standardised multifractal spectrum width and height ($\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$) with the δ, θ, α, β, and γ average band power in each epoch. Each scatter point is derived from a single epoch of the time series. The diagonal of the matrix features the histograms for each measure. The lower triangle contains the scatter plots for each pair of measures. The upper triangle shows the Pearson correlation for each pair of measure, where the size of the font additionally corresponds to the correlation coefficient to provide an additional visual cue. The icEEG data underlying this figure is shown in Figure 3.10 A.

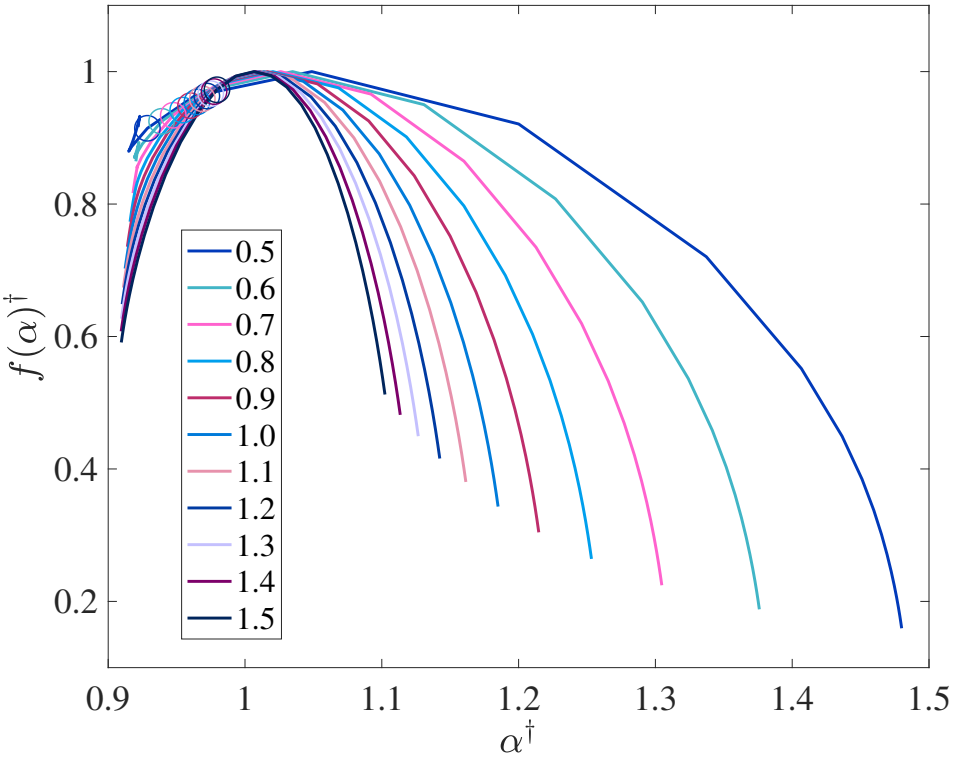

Figure 3.16: Multifractal spectra from a segment obtained with standardisation procedure, varying the standard deviation on sigmoid transformation. The standard deviation of the time series (128.4839) was multiplied by a scaling factor ranging from 0.5 to 1.5 by steps of 0.1. The scaling of the standard deviation changes the spectral features however $f(\alpha)$ for $q = 2$ (correlation dimension that is equivalent to monofractal dimension highlighted by the circles) shows a relatively small change. The little change in the correlation dimension caused by the standardisation procedure might explain why monofractal methods are not affected by the sigmoid standardisation approach and still correlate with the variance of the signal.

Chapter 4

Do sleep phases impact multifractal estimation in human electrographic recordings?

4.1 Introduction

Long term EEG monitoring of epileptic seizures is defined as the process of recording EEG signals for a 'prolonged period', i.e. for more than one hour (Lagerlund et al., 1996). Nevertheless, in clinical practice the duration of these recordings varies from a few hours to over a week (Lagerlund et al., 1996).

Long-term EEG recording of epileptic seizures is often indicated for an accurate diagnosis of the syndrome as well as correct identification of the type of seizures and their onset zone — for focal paroxysms (Lagerlund et al., 1996).

This type of recording has also been exploited by researchers studying brain dynamics, e.g. studies of the occurrence of epileptic seizure (Baldassano et al.,

2017; Brinkmann et al., 2016; Freestone et al., 2015; Karoly et al., 2017; Mormann et al., 2007; Kuhlmann et al., 2018a,b; Varatharajah et al., 2018). Long-term EEG monitoring is necessary to implement novel therapeutic approaches, which depend upon seizure detection/prediction (Iasemidis, 2003).

However, such long-term monitoring imposes several challenges. An example of these problems is systematic and time-varying properties of the brain, e.g. sleep/awake status. Epileptic seizures have also been shown to change accordingly with circadian and ultradian rhythms (Karoly et al., 2017; Baud et al., 2018). Nevertheless, the effect of brain rhythms during distinct sleep phases on EEG markers for seizure detection/prediction has had little attention in the literature.

Sleep stages are conventionally defined using specific power bands — namely δ ($<$ 4 Hz), θ (4-8 Hz), α (8-13 Hz), and β (14-40 Hz) bands (Noachtar et al., 1999), and can be divided in three main groups: waking, non-REM sleep, and REM sleep (Mendelson, 1987). The non-REM sleep stage can be divided in four subsequent types: stage 1, stage 2, stage 3, and stage 4 (Mendelson, 1987). More recently, stages 3 and 4 have been amalgamated into one deep sleep stage, termed N3 (American Academy of Sleep Medicine, 2007).

The waking, resting stage is characterised by an EEG signal predominantly composed of sinusoidal α waves, with some β waves of lower amplitude (Mendelson, 1987). During this stage the individual is in relaxed wakefulness and with their eyes closed (Mendelson, 1987).

For non-REM sleep, or stage 1, the α activity is reduced and there is a mixture of low-amplitude β and θ waves (Mendelson, 1987). Stage 2 mostly exhibits θ

waves and is characterised by two main distinguishing events: sleep spindles and the K complexes (Mendelson, 1987). The former consists of a burst of central rhythmic activity at 14 Hz, whereas the latter is a "high-amplitude negative wave followed by a positive wave" (Mendelson, 1987).

Stages 3 and 4 are marked by high amplitude (slow) δ waves (Mendelson, 1987). These stages are also known as slow-wave sleep or delta sleep (Mendelson, 1987). If δ waves comprise from 20% to 50% of a 30-second epoch, the stage is scores as 3. If δ waves correspond to more than 50% of a 30-second epoch, this characterises the stage 4 (Mendelson, 1987).

Sleep stages' scoring is normally performed by a specialist who does the markings manually. An automatic or semi-automatic approach to facilitate such work could help optimise the scoring of sleep EEG signals by making it quicker and more precise. There have been several studies of computational methods to perform such tasks, often applying machine learning techniques (Liang et al., 2012; Estrada et al., 2005; Smith and Karacan, 1971; Agarwal and Gotman, 2001; Flexer et al., 2005; Koley and Dey, 2012; Güne et al., 2010; Radha et al., 2014; Lajnef et al., 2015).

For this thesis, the studies that apply (multi)fractal and/or scaling evaluation approaches to sleep stage scoring are of greatest interest. In the last decades, some studies focussed on such properties and how these markers relate to sleep stages Koley and Dey (2012); Ma et al. (2018); Lee et al. (2004); Ma et al. (2005).

Fractal and multifractal metrics were capable of showing differences between sleep phases when applied to EEG signals. Overall, the EEG signals exhibited increased complexity (a higher scaling exponent / shifted multifractal spectrum) for

deeper sleep stages (Lee et al., 2004; Ma et al., 2005).

Following these findings, and considering long-term seizure monitoring — in which a patient will likely pass through different sleep stages, this chapter focusses on whether/how sleep could affect multifractal analysis of EEG with the standardised approach. The aim of this chapter is to evaluate whether the approach proposed in chapter 3 is affected by sleep stages and how multifractal features change under such circumstances.

4.2 Methods

This section describes the application of the unbiased multifractal estimation method introduced in chapter 3 to sleep scalp EEG signals obtained from individuals during polysomnography.

4.2.1 Data

This study evaluated signal features that could be connected to sleep stages in scalp EEG signals measured from individuals without epilepsy. The scalp EEG data studied are available at PhysioNet repository (https://physionet.org/) (Goldberger et al., 2000; Kemp et al., 2000; Mourtazaev et al., 1995).

This chapter reports an analysis performed on EEG signals of five of these healthy individuals: ST7011J, ST7022J, ST7041J, ST7052J, ST7061J. These subjects were recruited from the general population in the period 1989-1991 and undertook medical examination conducted by a neurologist/psychiatrist. The individuals did not have any somatic, neurologic or psychiatric conditions (Mourtazaev et al., 1995). Individuals also did not have sleep complaints or took any sleep-related

medication (Mourtazaev et al., 1995).

The signals were digitised at 100 Hz with 12 bits/sample in a polysomnography set-up with two different channels: Fpz-Cz and Pz-Oz (Mourtazaev et al., 1995). Rechtschaffen and Kales stages were manually marked by a specialist in 30-second epochs, following the standard criteria (Mourtazaev et al., 1995).

4.2.2 Multifractal analysis

The multifractal analysis approach presented in chapter 3 was appled to these scalp EEG obtained from subjects as described in previous section, using both Fpz-Cz and Pz-Oz channels.

The Chhabra-Jensen method was applied to the scalp EEG data that was epoch-wise standardised and sigmoid transformed (i.e. the established pipeline introduced in chapter 3). Each epoch was 10.24 seconds long, as the epochs had 1024 points sampling, as the data had been sampled at 100 Hz.

4.2.3 δ-power

δ-power was also obtained for the EEG recordings as reference and to serve as comparison to standard sleep analysis. The delta power is often a proxy for slow wave sleep stages. For δ-power, the recordings were also split in 1024 points epochs - comprising 10.24 seconds. Both Fpz-Cz and Pz-Oz channels were analysed.

4.2.4 Dimensionality reduction

Dimensionality reduction consists of a procedure that converts high-dimensional input data $X = \{x_1, x_2, ..., x_n\}$ to lower-dimensional elements, e.g., two or three dimensions, $Y = \{y_1, y_2, ..., y_n\}$ (Van Der Maaten and Hinton, 2008). Several methods

attempt to tackle such problem with linear approaches, e.g. Principal Component Analysis (PCA) and Multidimensional Scaling (MDS) (Van Der Maaten and Hinton, 2008).

There is, however, a class of methods that perform non-linear operations and try to preserve the local structure of the data points (Van Der Maaten and Hinton, 2008). These methods will be referred to in this thesis as neighbour embedding approaches. This chapter shows the application of t-distributed stochastic neighbour embedding (t-SNE). t-SNE is an improvement on a previous method, stochastic neighbour embedding (SNE) (Van Der Maaten and Hinton, 2008). The technique was applied to visualise the higher-dimensional data from the sleep analysis in a 2-dimensional space.

4.2.4.1 SNE

Stochastic Neighbour Embedding (SNE) is a dimensionality reduction technique that converts high-dimensional Euclidean distances into low-dimensional representations. In SNE, these elements are treated as probabilities and are compared according to their similarities (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

Equation 4.1 indicates the conditional probability that x_i would pick x_j as neighbour — based on their probability density. $p_{j|i}$ is high for nearby points and low for sparse points (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$p_{j|i} = \frac{e^{||x_i - x_j||^2/2\sigma_i^2}}{\sum_{k \neq i} e^{||x_i - x_k||^2/2\sigma_i^2}} \tag{4.1}$$

Where σ_i is the variance of a Gaussian distribution that is centred on data point x_i. The aim of this operation is to look for pair-wise similarities so it is possible to assume $p_{i|i} = 0$ (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

A conditional probability for x_i and x_j low-dimensional counterparts, y_i and y_j, can be defined in a similar way to equation 4.1. In this case, the variance is set $\sigma_i = 0$ (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$q_{j|i} = \frac{e^{||y_i - yj||^2}}{\sum_{k \neq i} e^{||y_i - y_k||^2}} \tag{4.2}$$

Again, the aim is to look for pair-wise similarities so $q_{i|i} = 0$. If y_i and y_j model correctly x_i and x_j, $p_{j|i}$ and $q_{j|i}$ must be equal. A way of quantifying the similarity between two different distributions is Kullback-Leibler (KL) divergence, shown in equation 4.3 (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

SNE minimises the sum of KL divergences over all data points with gradient descent method and cost function given by 4.3 (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$C = \sum_i KL(P_i||Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}} \tag{4.3}$$

P_i is the conditional probability distribution over data points x_i, whereas Q_i is the conditional probability distribution over data points y_i (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

SNE cost function focusses on returning the local structure of the data points, i.e. there is a large cost for mapping nearby points with widely separated representations. However, it has a small cost to map widely separated points and nearby representations. Such a difference is due to KL divergence asymmetry (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

There is still a remaining parameter to be chosen — variance σ_i of the Gaussian distribution centred at each high-dimensional data point x_i. Data points are unlikely to have a unique optimal variance and σ_i should vary according to the density of the data, as smaller values of σ_i are more appropriate to denser regions. SNE searches for a value of σ_i that produces a P_i with a perplexity specified by the user. The perplexity is associated with the number of neighbours of a point x_i and is defined according to equation 4.4 (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$Perp(P_i) = 2^{H(P_i)} \tag{4.4}$$

$H(P_i)$ corresponds to Shannon's entropy measured in bits, as seen in equation 4.5.

$$H(P_i) = -\sum_j p_{j|i} \log_2 p_{j|i} \tag{4.5}$$

The gradient of equation 4.3 — cost function — allows the minimisation of C and the determination of the best low-dimensional representation of x_i. The gradient of the cost function is obtained with gradient descent method and is shown in

equation 4.6 (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$\nabla C = \frac{\delta C}{\delta y_i} = 2\sum_j (p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j) \tag{4.6}$$

The procedure of minimisations shown in 4.7 starts with a random mapping $\lambda^{(t-1)}$. An exponentially decaying $\alpha(t)(\lambda^{(t-1)} - \lambda^{(t-2)})$ is added to accelerate the calculation (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

$$\lambda^{(t)} = \lambda^{(t-1)} + n\frac{\delta C}{\delta \lambda} + \alpha(t)(\lambda^{(t-1)} - \lambda^{(t-2)}) \tag{4.7}$$

Moreover, the two multipliers n and $\alpha(t)$ also feature on the iterative operation. With the former representing the learning rate or step size and the latter indicating the moment of random noise term at iteration t. The output corresponds to $\lambda^{(t)}$ when t reaches the maximum number of iterations defined before the calculation starts (Van Der Maaten and Hinton, 2008; Hinton and Roweis, 2002).

4.2.4.2 t-SNE

SNE is capable of reproducing high-dimensional structures in analogue low-dimensional representations. Nevertheless, the approach suffers from two main problems: the lack of symmetry of the cost function — KL divergence — and the so-called ”crowding problem” (Van Der Maaten and Hinton, 2008). The latter refers to difficulties associated with representing high-dimensional distances in a lower dimensionality map, i.e. if small distances are accurately represented on the map, medium or longer distances might not be reproduced as well (Van Der Maaten and Hinton, 2008).

A new method called t-distributed stochastic neighbor embedding (t-SNE) has sorted both these issues from SNE with a symmetric cost function and using a Student t-distribution for low-dimensional representations (Van Der Maaten and Hinton, 2008). For t-SNE high-dimensional and low-dimensional probabilities are described, respectively according to equations 4.8 and 4.9 (Van Der Maaten and Hinton, 2008).

$$p_{i|j} = \frac{e^{\frac{-||y_i - y_j||^2}{2\sigma^2}}}{\sum_{k \neq l} e^{\frac{-||y_k - y_l||^2}{2\sigma^2}}} \tag{4.8}$$

$$q_{i|j} = \frac{(1 + ||y_i - y_j||^2)^{-1}}{\sum_{k \neq l}(1 + ||y_k - y_l||^2)^{-1}} \tag{4.9}$$

The cost function — KL divergence — and the minimisation (gradient) of this can then be written accordingly, as shown in equations 4.10 and 4.11.

$$C = KL(P||Q) = \sum_i \sum_j p_{i|j} \log \frac{p_{i|j}}{q_{i|j}} \tag{4.10}$$

$$\nabla C = \frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 + ||y_i - y_j||^2)^{-1} \tag{4.11}$$

t-SNE has been successfully applied to a wide variety of research problems, including in the biomedical sciences (Wong et al., 2015; McDonnell et al., 2016; Platzer, 2013; Taskesen and Reinders, 2016). A case to note is the work of Birjandtalab et al. (2016), that applied t-SNE in a seizure detection study.

This chapter shows the application of t-SNE to the six variables collected from

the two channels, comprising multifractal spectra width, multifractal spectra height, and δ-power — from all subjects. For such a procedure, the perplexity was defined as 40 and the maximum number of iterations gradient was set to 1000.

The aim of this procedure is to reduce the dimensionality of the features obtained from sleep EEG — for both Fpz-Cz and Pz-Oz channels — and to look for patterns for the different sleep stages that should suggest that the outcomes of the approach proposed in chapter 3 would be susceptible to sleep states. If this turn out to be the case, the suggested approach could be exploited in the development of an automatic (or semi-automatic) sleep staging system.

4.3 Results

Figure 4.1 shows the variation of the multifractal spectra width (B), the multifractal spectra height (C), and power in δ-band for the different sleep stages shown by the hypnogram (A) for patient 'ST7011J' and both EEG channels: Fpz-Cz and Pz-Oz.

The violin plots show the distribution of $\Delta\alpha^{\dagger}$, $\Delta f^{\dagger}$, and power in the δ band (Fig. 4.1 E-J). A clear drift in the $\Delta\alpha^{\dagger}$ distribution can be observed that slowly changes from awake to REM to S1, S2, S3 and S4 (Fig. 4.1 E and H). This is in contrast to power in the δ band, where the distributions between awake, REM, and S1 are very similar, and then become highly variable in S2, S3 and S4 (Fig. 4.1 G and J).

A repeated measures ANOVA test was applied to the multifractal metrics in both channels to verify whether these provide statistically significant information about the sleep stages. The predicting model evaluated

Table 4.1: F-values and p-values obtained from repeated measures ANOVA hypothesis test for $\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$ on Fpz-Cz and Pz-Oz channels of polysomnography sleep EEG. Both $\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$ show statistically significant differences on FPz-Cz and Pz-Oz channels, suggesting that multifractal measures in long term EEG can be affected by the state of the patient and that these metrics could be employed to devise some (semi) automatic sleep staging method. ** - $p < 0.01$ * - $p < 0.05$.

	$\Delta\alpha^{\dagger}$	$\Delta f^{\dagger}$
Fpz-Cz	$F = 16.310$ $p = 0.016^{*}$	$F = 29.320$ $p = 0.006^{**}$
Pz-Oz	$F = 7.938$ $p = 0.048^{*}$	$F = 19.310$ $p = 0.012^{*}$

is described by the formula: $Sleep\ stage \sim Multifractal\ measure * Time + Error(Patient/(Multifractal\ measure * Time))$. The test reports statistically significant differences for both $\Delta\alpha^{\dagger}$ and $\Delta f^{\dagger}$ across different sleep stages. F-values and p-values for every channel and metric are shown in table 4.1.

Fig 4.2 shows the projection into a two-dimensional space. Note that the information about the sleep stages was not given to the tSNE algorithm. The projection shows one big cluster and two smaller clusters. When applying the information about the sleep stages, it becomes clear that the two smaller clusters are awake states, whereas the one big cluster shows a progression from REM to S4. Note that there are also data points from the awake state in this cluster.

4.4 Discussion

Results from this preliminary evaluation on how multifractal metrics are affected by sleep stages suggest that sleep stages do affect multifractal markers. The outcome is important because it emphasises that future applications of multifractal analysis to EEG signals should take sleep phenomena into account when designing experi-

ments and solutions.

Distributions of multifractal spectra width show an increase towards deeper sleep phases, probably due the appearance of new singularities — related to the sleep stage — in the signal. The variation is also similar to the one exhibited by power in δ-band. In contrast, the multifractal spectra height does not vary with sleep stage.

The output exhibited by the two-dimensional representation of the measures with t-SNE reinforces the results show in figure 4.1, suggesting some continuity of sleep stages. There are also points of wake stage scattered all over the figure. Future work will need to determine if those data points are simply mislabeled, or in any way closer to sleep states. If so, multifractal measures may be useful in devising a classifier for sleep stages in EEG.

The sleep rhythms show different spatial distributions and so different montages and referencing could impact the multifractal analysis and indeed sleep interpretation. American Academy of Sleep Medicine therefore recommends standard montages and referencing (American Academy of Sleep Medicine, 2007). These were used in this study but it is possible that greater information could be obtained with different montages.

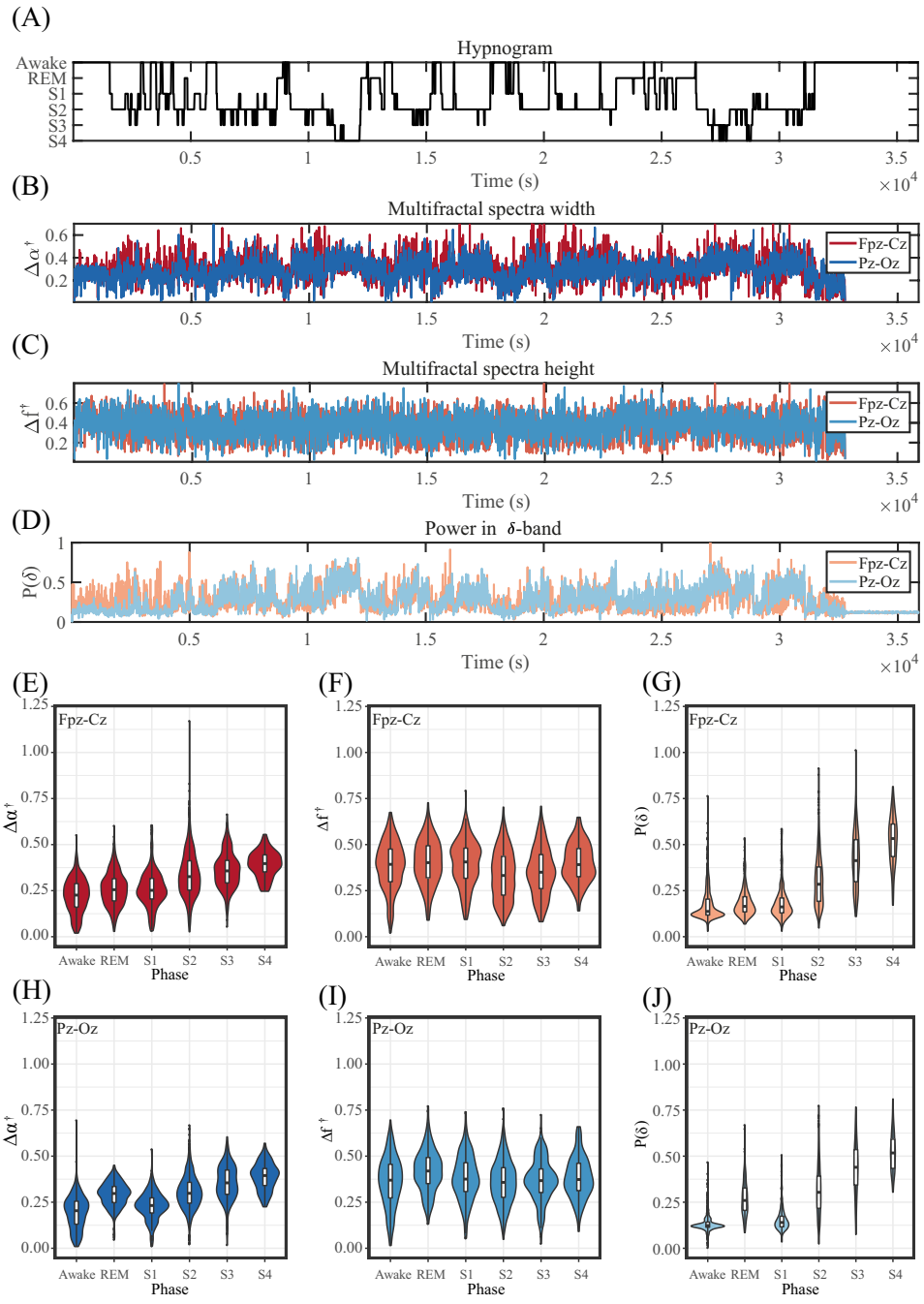

Figure 4.1: Impact of sleep stages on multifractal markers. The figure features measures performed on a scalp EEG recorded in a polysomnography in two montages Fpz-Cz and Pz-Oz for patient ST7011J. (A) Hypnogram with sleep phases marked by a specialist. (B) Variation of the Multifractal spectra width in time. (C) Variation of the Multifractal spectra height in time. (D) Variation of the power of the δ band in time. (E, F, G, H, I, and J) Violin plots of the estimated values of $\Delta\alpha^{\dagger}$ (E, H), $\Delta f^{\dagger}$ (F, I), and $P(\delta)$ (G, J) for both Fpz-Cz (E, F, G) and Pz-Oz (H, I, J) montages. At visual inspection, $\Delta\alpha^{\dagger}$ presents a behaviour similar to $P(\delta)$, following the variations of the hypnogram. The violin plots show an increase in the measures towards more advanced sleep phases for both $\Delta\alpha^{\dagger}$ and $P(\delta)$, in both montages.

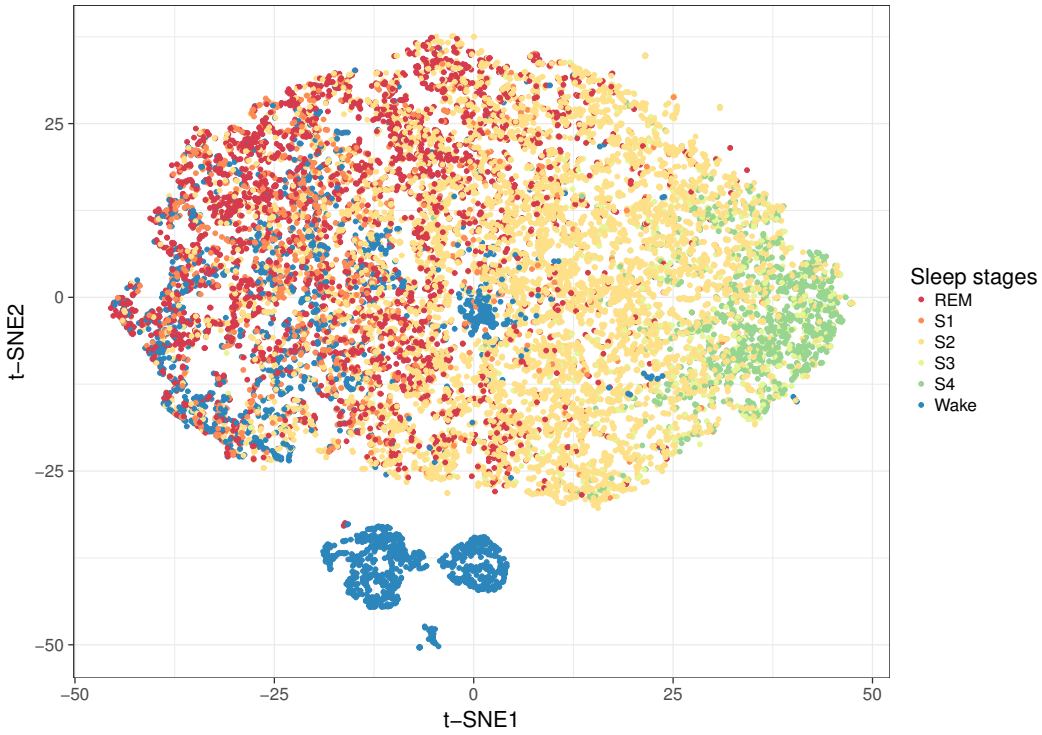

Figure 4.2: Dimensionality reduction with t-distributed stochastic neighbor embedding of the $\Delta\alpha^{\dagger}$, $\Delta f^{\dagger}$, and $P(\delta)$ variables for both Fpz-Cz and Pz-Oz channels in a two-dimensional space. The figure shows all sleep stages. It is possible to observe that the measures provide information about the sleep stage that could be used in order to devise a classifier.

Chapter 5

Do seizures have a characteristic time for detection?

5.1 Introduction

The parameters of analysis, namely sampling rate and epoch size, are also important in evaluating putative methods for seizure detection/prediction. The latter will depend on the chosen methods of analysis and choice for the study; however, the former relies on the data acquisition, which in the case of epilepsy is often decided in a clinical context.

The seizure detection/prediction literature is rather diverse when it comes to the characteristics of the analysed data. Regarding the source of data which has been the subject of studies on prediction/detection studies, it can be divided in three different groups. Most of the data comes from intracranial EEG (iEEG) (Le Van Quyen et al., 2005; Harrison et al., 2005a,b; Feldwisch-Drentrup et al., 2011; Takahashi et al., 2012; Williamson et al., 2012; Naro et al., 2014; Sharafat and Nesaei, 2015;

Eftekhar et al., 2014; Zhang et al., 2014, 2015; Zheng et al., 2014; Niknazar and Nasrabadi, 2015).

The second group is composed of data from scalp EEG, and features a significant amount of different studies (Protopopescu et al., 2001; Drury et al., 2003; Li et al., 2006; Aksenova et al., 2007; James and Gupta, 2009; Wang et al., 2011).

The third and smallest group comprises scalp + intracranial recordings (not acquired from the same patient and/or simultaneously) (Navarro et al., 2002; Hyunchul Kim and Rosen, 2010; Cheng-Yi Chiang et al., 2011; Bandarabadi et al., 2012; Chang et al., 2012).

Different sources of data can affect the capacity of recording epileptiform activity and be more (scalp) or less (intracranial) susceptible to noise. Moreover, of interest to this thesis is the study by Zhang et al. (2015) that applied multifractal analysis to seizure detection and suggests that such a technique can be successfully employed into seizure detection.

The greater number of intracranial EEG data reflects the technique's higher sensitivity and lower noise compared to scalp recordings and serves to localise seizures when non-invasive methodologies are not able to do so. Intracranial EEG are often obtained at higher sampling rates, up to 5000 Hz.

Such difference in parameters of analysis impose difficulties in comparing different detection/prediction approaches. Moreover, a question remains unanswered: are there optimal sampling frequency and epoch length to detect focal epileptic seizures?

This chapter aims to explore this issue, comparing the impact of different

Table 5.1: Duration of the epileptic seizure evaluated in each patient.

Patient	Seizure duration (s)
I001_P005_D01	70.50
Study 040	19.65
I001_P010_D01	39.62
I001_P034_D01	28.31

choices for epoch sizes and distinct sampling rates on multifractal analysis metrics.

5.2 Methods

This analysis consisted of evaluating the sensitivity of multifractal metrics to different sets of parameters of the signal, namely epoch sizes and sampling rates.

5.2.1 Data

Intracranial EEG from four subjects were retrieved from the ieeg.org repository (`http://www.ieeg.org/`) (Wagenaar et al., 2013): I001_P005_D01, I001_P034_D01, 'I001_P010_D01', and Study 040. These subjects were chosen due to the high sampling rate of their recordings (5 kHz), so the impact of sampling frequency on multifractal properties could be evaluated. A single seizure was obtained for every patient.

A 30-minute segment around the seizure was evaluated in each patient for further analysis. The multifractal analysis was performed on channels that were marked as seizure onset channels by a specialist. The duration of the epileptic seizure of each patient is shown in table 5.1.

The anonymised data analysed in this study were recorded in patients undergoing evaluation for epilepsy surgery. The iEEG.org portal provided EEG data

and ethical approval for analysing the data was provided by Mayo Clinic IRB (Brinkmann et al., 2016, 2009).

5.2.2 Downsampling

In order to evaluate the effect of different sampling frequencies, the original signals were downsampled by removing data points. Signals were downsampled to the following sampling frequencies: 4000 Hz, 3000 Hz, 2500 Hz, 2000 Hz, 1000 Hz, 800 Hz, 750 Hz, 600 Hz, 500 Hz, 400 Hz, 300 Hz, and 250Hz.

5.2.3 Multifractal analysis

Finally, the impact of the multifractal estimation parameters in the characterisation of a seizure was evaluated. The multifractal analysis approach presented in chapter 3 was applied to these original and downsampled versions. For this analysis, downsampled versions were analysed with epochs of different sizes.

The Chhabra-Jensen method was applied to the recordings that were epoch-wise standardised and sigmoid transformed (i.e. the established pipeline introduced in 3). Each epoch featured different sizes (1024 points, 2048 points, 4096 points, 8192 points, and 16384 points), and different durations — depending on the downsampled frequency.

5.2.4 Cohen's D effect size

The effect of the pair of parameters — sampling rate and epoch size — over the sensitivity of multifractal metrics to epileptic seizures was evaluated with Cohen's *D* measure in equation 5.1.

The multifractal spectrum width ($\Delta\alpha^{\dagger}$) during the seizure was compared to the

background as the effect size (Cohens D) between the ictal and interictal periods:

$$D = \frac{< \Delta\alpha^{\dagger}_{ictal} > - < \Delta\alpha^{\dagger}_{interictal} >}{s(\Delta\alpha^{\dagger}_{interictal})} \tag{5.1}$$

where $< \Delta\alpha^{\dagger} >$ represents the mean and s denotes standard deviation. Higher values of Cohens D would show a higher sensitivity to epileptic seizure detection and a variation of such values for different sampling rates and epoch sizes could suggest optimal parameters for seizure detection — and maybe a characteristic time.

5.3 Results

The variation of the multifractal spectrum width $\Delta\alpha^{\dagger}$ for different combinations of epoch sizes and sampling frequencies for patient I001_P005_D01 is shown in figure 5.1 (A). On visual inspection, it is clear that there are some combinations of epoch size and sampling frequency that show a clear increase of $\Delta\alpha^{\dagger}$ during the ictal period (marked by the red lines).

To quantify this effect, figure 5.1 (B) shows the Cohen's effect size D of the ictal vs. interictal $\Delta\alpha^{\dagger}$ distributions plotted against epoch duration (in seconds). In this plot, 15 different sampling frequencies were included, and also data from three different EEG channels (all in the seizure onset zone).

A peak in D can be seen at about 1 second (across all sampling frequencies), indicating that the change in $\Delta\alpha^{\dagger}$ during a seizure can be best captured when using 1-second epochs (regardless of sampling frequency, or number of samples). This effect was not found for the sampling frequency or epoch length separately.

Other patients have shown similar results, as can be seen in figures 5.2, 5.3, and

5.4. Nevertheless, some of these individuals show peaks for D at different epoch durations for either some specific channels, e.g. figure 5.3, or for all channels in the seizure onset zone, e.g. figure 5.4. This feature is more pronounced for patients shown in figures 5.1 and 5.2.

5.4 Sampling rate invariance

A visual inspection of Figure 5.1 also suggests that Cohen's D is not directly affected by the sampling, i.e. is invariant to sampling rate. A more careful evaluation suggests that such a property is indeed true.

Figure 5.5 shows the distribution of Cohen's D values for a range of frequencies. It is possible to observe, that for the interval between 250 Hz and 5000 Hz, the value of Cohen's D is constant across different sampling frequencies. Moreover, the colour scales show that some time epoch durations feature higher values for D - across multiple sampling frequencies.

A hypothesis test reports non-significant differences between the Cohen's D values obtained for different sampling frequencies. The ANOVA model reports $F - value = 1.213$ and $p = 0.272$.

5.5 Discussion

This analysis highlighted the importance of choosing an adequate epoch size given a sampling frequency, in order to study events such as epileptic seizures. Moreover, this study has also shown that these features are invariant to sampling frequencies (in a given range between 250 Hz and 5000 Hz) and specific optimal time scales

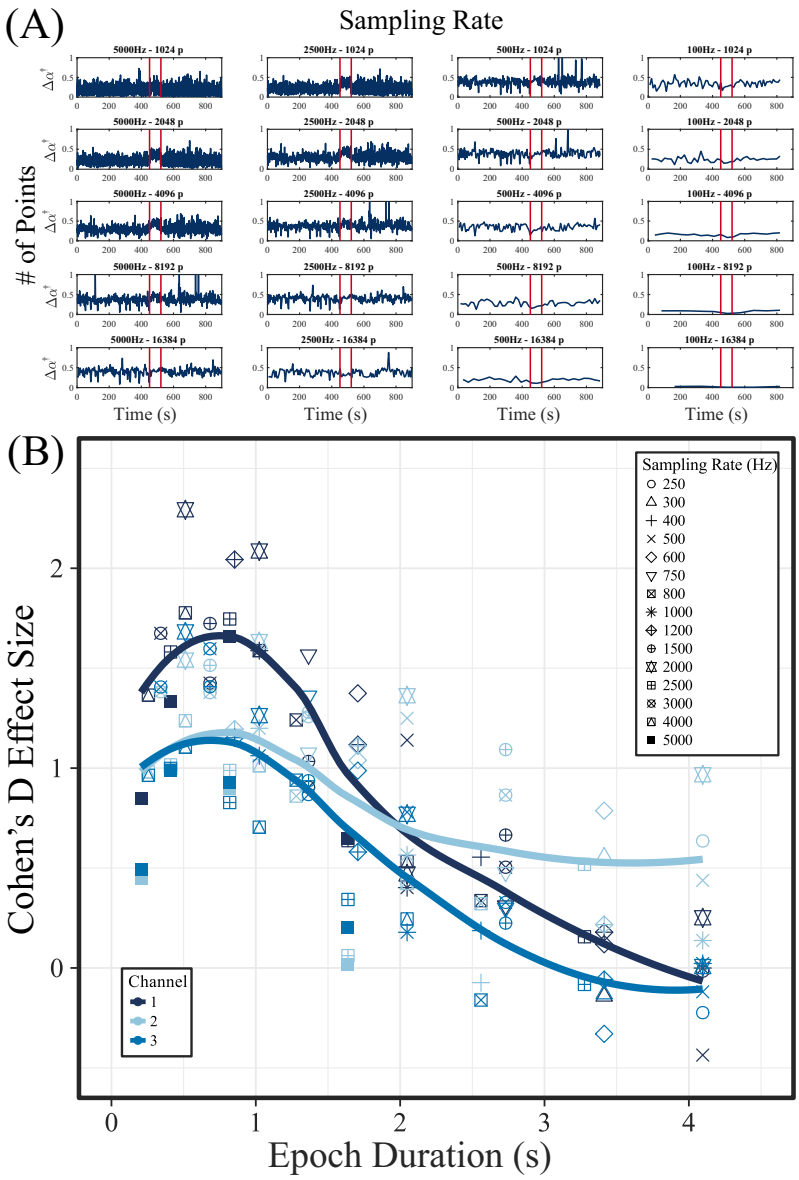

Figure 5.1: Influence of EEG sampling frequency and epoch length on multifractal spectrum width around and during an epileptic seizure. The data used for this figure was obtained from subject 'I001_P005_D01' — channels 1, 2, and 3. (A) Multifractal spectrum width ($\Delta\alpha^{\dagger}$) in a 15-minute intracranial EEG segment containing one seizure (onset and offset marked by the red lines). The signal was initially sampled at 5000 Hz. Each column shows $\Delta\alpha^{\dagger}$ for 5000 Hz, 2500 Hz, 500 Hz and 100 Hz sampling rates. Different epoch sizes were used ranging from 1024 to 16384 samples (in each row). (B) Relationship of effect size D (between the interictal and ictal distribution of $\Delta\alpha^{\dagger}$) and epoch duration in seconds (obtained by dividing the number of sampling points by the sampling rate of the signal). Channel 1 is the data shown in (A). The solid line represents a LOESS curve fitting of the data points, with formula $y \sim x$.

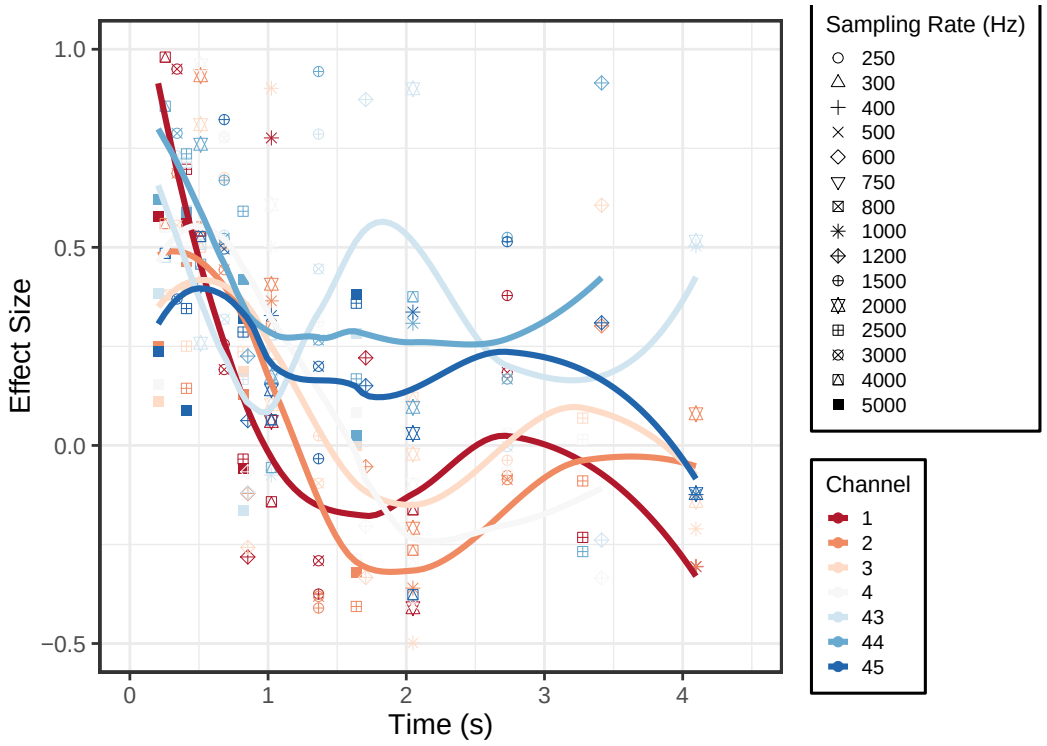

Figure 5.2: Relationship of effect size D (between the interictal and ictal distribution of $\Delta\alpha^{\dagger}$) and epoch duration in seconds (obtained by dividing the number of sampling points by the sampling rate of the signal). The data used for this figure was obtained from subject Study 040 — channels 1-4 and 43-45. The solid line represents a LOESS curve fitting of the data points, with formula $y \sim x$.

are featured across sampling frequencies. However, this study was based on the analysis of ictal vs. interictal epochs, i.e. a hard separation that may not represent continuous phenomena accurately.

Future work should take into account that multifractal properties may be continuously changing over time, and an explicitly time-based approach may be needed. Along similar lines, this finding of an optimal timescale may be due to the non-stationary nature of the multifractal properties. Further theoretical work may have to develop a temporally-resolved multifractal estimator, in order to fully understand this aspect.

A fundamental observation in this work is that optimal time scales may exist

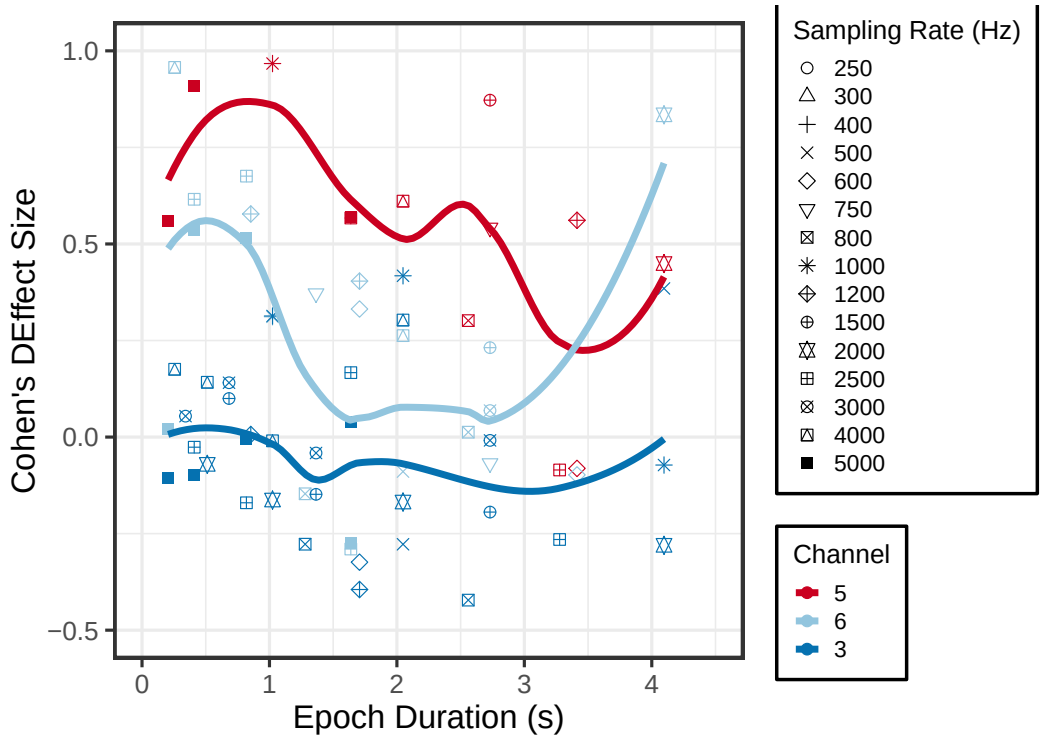

Figure 5.3: Relationship of effect size D (between the interictal and ictal distribution of $\Delta\alpha^{\dagger}$) and epoch duration in seconds (obtained by dividing the number of sampling points by the sampling rate of the signal). The data used for this figure was obtained from subject 'I001_P010_D01' — channels 3, 5, and 6. The solid line represents a LOESS curve fitting of the data points, with formula $y \sim x$.

for specific physiological processes (such as epileptic seizures) in terms of their multifractal dynamics. This result suggests that, at least in an epoch-based study, for any given epileptic seizure in a given patient, the variety of scaling exponents ($\Delta\alpha$) will depend on the length of the epoch analysed. This is further supported by similar findings in monofractal analysis (Eke et al. (2002)). The implications of this observation are that certain scaling exponents will only exist in specific timescales and the diversity of scaling exponents will depend on the duration of the epoch. These results suggest the potential need for tuning, i.e. potentially having to find the characteristic time for every studied phenomenon. If this is indeed the case, a temporally resolved (not epoch-based) multifractal method should be developed in

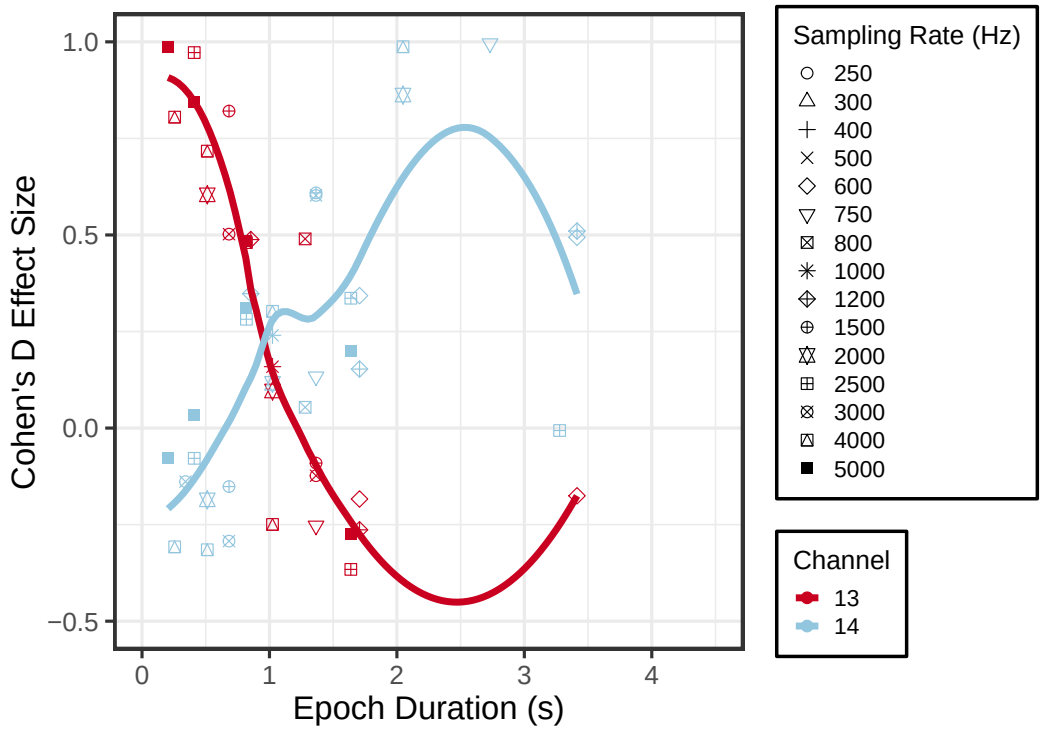

Figure 5.4: Relationship of effect size D (between the interictal and ictal distribution of $\Delta\alpha^{\dagger}$) and epoch duration in seconds (obtained by dividing the number of sampling points by the sampling rate of the signal). The data used for this figure was obtained from subject 'I001_P034_D01' — channels 13 and 14. The solid line represents a LOESS curve fitting of the data points, with formula $y \sim x$.

future to adequately characterise brain dynamics.

The occurrence of peaks for D at different epoch durations suggest that patients with distinct types of seizure could exhibit other characteristic times for detection — again, this would be connected to the dynamics of onset and propagation of the epileptic seizure. A similar rationale would be valid for different channels implanted to a same patient; the dynamics of recruitment would follow local patterns. It is important to highlight that channels not obviously involved in the seizures may nevertheless carry information about the epileptic seizures as part of a wider ictal network (Sinha et al., 2017).

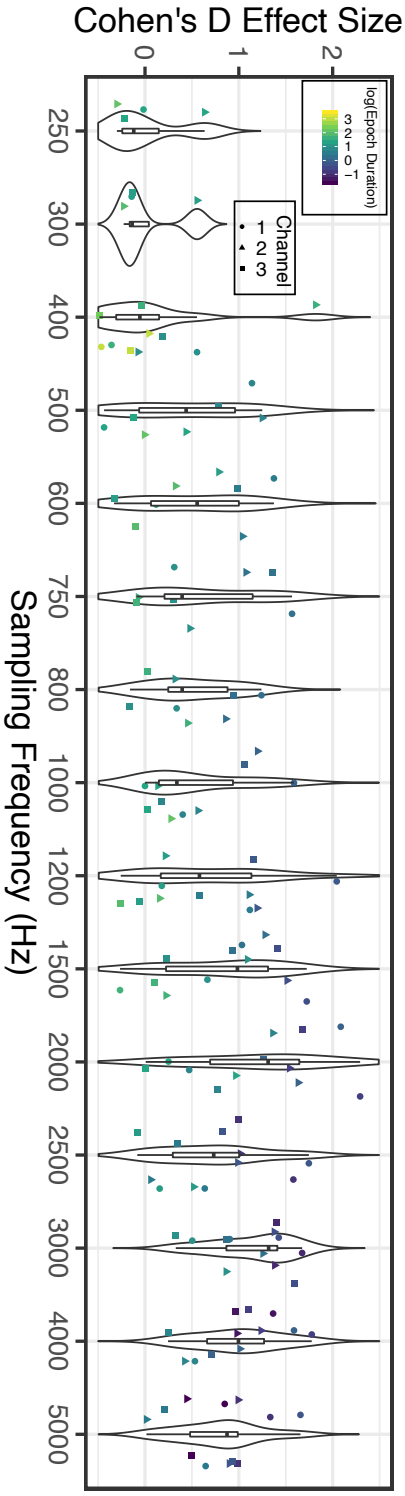

Figure 5.5: Violin plots showing the distributions of effect size D values (between the interictal and ictal distribution of $\Delta\alpha^{\dagger}$) for frequencies ranging from 250 Hz to 5000 Hz for subject 'I001_P005_D01'. The figure shows that D is invariant to sampling frequency in the range between 250 Hz and 5000 Hz. Moreover, the colour scale shows that some timescales feature higher values for D in multiple sampling frequencies.

Chapter 6

How to classify focal epileptic seizures

6.1 Introduction

Over the last decades, different methods for seizure detection/prediction have been proposed. Such approaches employ a range of different techniques in order to make inferences on the occurrence of epileptic seizures from information contained in brain signals, e.g. intracranial EEG (Mormann et al., 2007; Freestone et al., 2015; Kuhlmann et al., 2018b). Most of these methods have been applied to small groups of patients though.

Since the appearance of seizure detection studies in literature, only one clinical trial on seizure prediction has been conducted (Cook et al., 2013, 2016); in this study, nine patients underwent the whole course of the trial (Kuhlmann et al., 2018b). The system consisted of implanted electrodes connected to an advisory system with indicating lights — that reported the risk of an epileptic seizure onset.

This algorithm used by the device was developed for every individual patient in order to account for patient specific patterns and changes in the signal (Kuhlmann et al., 2018b).

The clinical trial resulted in good outcomes for three of the nine patients with sensitivities of 86% and 100%; however, the other six individuals have shown lower sensitivities — ranging from 54% to 71% (Kuhlmann et al., 2018b). Further studies on the occurrence of seizures are still necessary to help determine why some patients do not show desirable results with the advisory system; such knowledge could help in the understanding and classification of epilepsy (Kuhlmann et al., 2018b). However, some factors have already been shown to confound seizure prediction. These features are: high frequency of seizures, temporal variations in the durations of seizures, and multimodal seizure durations and intervals between events (Cook et al., 2014, 2016; Kuhlmann et al., 2018b).

These issues seem to arise from the fact that the seizure onset — and subsequent dynamics — is a complex phenomenon, i.e. does not feature monomorphic properties (Grinenko et al., 2018). Moreover, epileptic seizures seem to exhibit different patterns and properties in both intra- and inter- patient comparisons (Cook et al., 2016; Kuhlmann et al., 2018b; Schroeder et al., 2019). Essentially, there might be different types of seizures, exhibiting different dynamics, even within the same patient. Such a difference can be explained by both the ictal onset zone and brain network in which the seizure spreads across the brain (Schroeder et al., 2019).

The existence of distinct types of seizures implies that there may be a method of classifying such episodes in groups — that could provide insights into epilepto-

genesis and epilepsies. Machine learning techniques have been extensively applied to comparisons of signals/time series and, in some cases, to medical applications (De Fauw et al., 2018). Such techniques could be applied to compare epileptic seizures and to devise an objective classification of such episodes.

However, according to Kuhlmann et al. (2018b), "machine learning does not readily reveal which physiological aspects underlie the predictive characteristics of EEG", as previously applied to some seizure prediction/detection studies (Cook et al., 2016; Kuhlmann et al., 2018b). Nevertheless, if explored with care, these techniques can be useful to extract information from seizure data, providing insights on a putative classification scheme of focal epileptic seizures.

An example of such an application was proposed by Schroeder et al. (2019). In this study the authors applied Dynamic Time Warping to classify epileptic seizures based on their network of propagation.

This chapter presents an approach to classify epileptic seizures based on MF methods introduced in chapter 3, Dynamic Time Warping and hierarchical clustering.

6.2 Methods

This section describes the approach developed in this study to classify focal epileptic seizure with machine learning techniques for dimensionality reduction and clustering.

Table 6.1: List of patients featuring number of channels and seizures. This study evaluated a total of 137 epileptic seizures across all patients.

Patient	# of channels	# of seizures	Surgical outcome
Study 004-2	56	3	Bad
Study 016	64	7	Bad
Study 019	96	36	Bad
Study 020	56	8	Bad
Study 021	108	13	Good
Study 022	56	7	Bad
Study 023	88	4	Good
Study 026	96	20	Good
Study 028	96	9	Good
Study 029	64	3	Bad
Study 033	128	17	Bad
Study 038	88	10	Good

6.2.1 Data

Intracranial EEG segments from 12 subjects were retrieved from the ieeg.org repository (`http://www.ieeg.org/`) (Wagenaar et al., 2013). The segments contain a seizure each and start 60 seconds before seizure onset, finishing 10 seconds after. Table 6.1 shows the number of seizures and channels for each patient.

Seizures with artefacts were not included in the analysis. The twelve patients had 137 epileptic seizures selected in total. These segments feature various durations, from seconds to minutes.

6.2.2 Multifractal analysis

The multifractal analysis approach presented in chapter 3 was applied to these intracranial EEG data. The Chhabra-Jensen method was applied to time data that was epoch-wise standardised and sigmoid transformed (i.e. the established pipeline introduced in chapter 3). Each epoch was 2.048 seconds long, as the epochs had 1024

points sampling, as sampling frequency of the data was 500 Hz.

6.2.3 Principal components analysis

Principal Components Analysis (PCA) is a method often used to extract low-dimensional subsets of features from high-dimensional data (James et al., 2013). Given a vector X of p variables $X = \{x_1, x_2, ..., x_p\}$, it may be challenging to study the variances and correlations of p - unless this vector has a very simple structure (Jolliffe, 2002).

One can alternatively look for the elements of larger variance. Principal components analysis aims to describe the data with a linear function $\alpha_1' X$ of X, where α_1 is a vector of p constants $\{\alpha_{11}, \alpha_{12}, ..., \alpha_{1p}\}$. $\alpha_1' X$, in order to maximise the inherited variance. $\alpha_1' X$ is the first component, PC1 (Jolliffe, 2002).

$$\alpha_1' X = \alpha_{11} x_1 + \alpha_{12} x_2 + ... + \alpha_{1p} x_p = \sum_{j=1}^{p} \alpha_{1j} x_j \tag{6.1}$$

The next step consists of finding another function uncorrelated with $\alpha_1' X$ and with maximum variance, $\alpha_2' X$. The following steps are similar until the k-th iteration and the linear function $\alpha_k' X$ that must be uncorrelated with $\alpha_1' X, \alpha_2' X, ..., \alpha_{k-1}' X$ (Jolliffe, 2002). A two-dimensional example is shown in figure 6.1. The green line (PC1) represents the direction of greatest variance and the projection of the data points in the function $\alpha_1' X$ will represent the first component's scores (James et al., 2013).

In this study, PCA was applied to reduce the dimensionality of the multifractal spectra width obtained for each of the EEG channels.

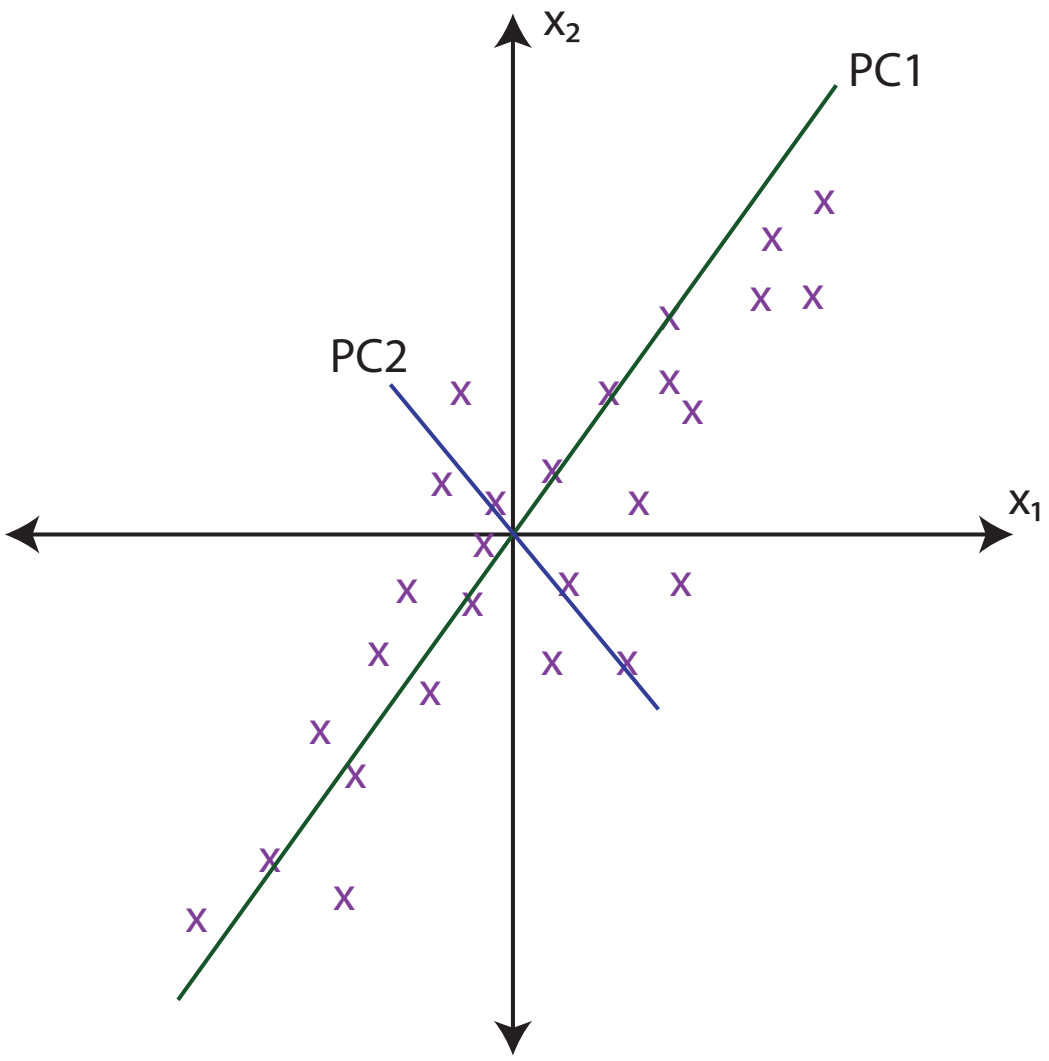

Figure 6.1: Illustration of Principal Component Analysis in a two-dimensional case. PC1 and PC2 represent the functions $\alpha_1' X$ and $\alpha_2' X$, respectively

6.2.4 Dynamic time warping

It has been previously shown that Euclidean distance is not an adequate measure to compare differences between two time series (Petitjean et al., 2011; Sakoe and Chiba, 1978). In order to avoid such limitations, a new method — called Dynamic Time Warping (DTW) — has been developed (Petitjean et al., 2011; Sakoe and Chiba, 1978; Berndt and Clifford, 1994).

Given two sequences $A =< a_1, ..., a_i, ..., a_I >$ and $B =< b_1, ..., b_j, ..., b_J >$, and

a distance between elements (or coordinates) δ, it is possible to define the Euclidean distance of two sequences as shown in equation 6.2 (Petitjean et al., 2011; Sakoe and Chiba, 1978; Berndt and Clifford, 1994). This distance, however, is not capable of capturing flexible similarities (Petitjean et al., 2011; Sakoe and Chiba, 1978; Berndt and Clifford, 1994).

$$D(A,B) = \sqrt{\delta(a_1,b_1)^2 + ... + \delta(a_i,b_j)^2 + ... + \delta(a_I,b_J)^2} \tag{6.2}$$

DTW tries to solve this limitation by finding an optimal alignment (or coupling) between the two sequences by obtaining the dissimilarity according to equation 6.3, i.e. the approach tries to bypass the timing differences between the two sequences by matching them on a plane shown in figure 6.3 (Petitjean et al., 2011; Sakoe and Chiba, 1978; Berndt and Clifford, 1994).

The timing differences between the points in sequences A and B are described by $C = (i, j)$, that constitutes a series shown in equation 6.3. Such a series corresponds to a Function F that approximately maps sequence A onto sequence B. For the case of two identical sequences, the warping function coincides with the diagonal $i = j$ and consists of the canonical Euclidean distance (Petitjean et al., 2011; Sakoe and Chiba, 1978; Berndt and Clifford, 1994). Figure 6.3 illustrates the differences between Euclidean distance and DTW.

$$F = C_1, C_2, ..., C_k, ..., C_K \quad : \quad C_k = (i(k), j(k)) \tag{6.3}$$

DTW has been initially developed to solve problems in speech recognition

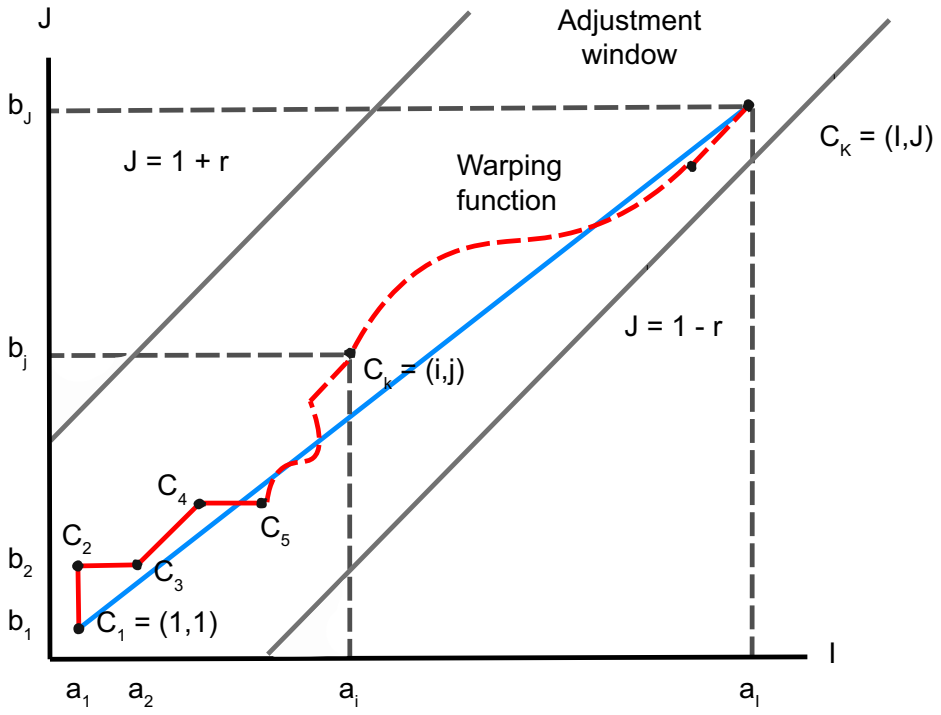

Figure 6.2: Warping function $C(k)$ (red line) for optimal alignment of the sequences A and B. The devised warping functions follows the boundary conditions established — these are indicated by the lines J, as well as the monotonicity (subsequent values must be equal or bigger than previous measure) and continuity restrictions. The blue line indicates a pair-wise comparison between sequences A and B, i.e. without Dynamic Time Warping approach. The values of $C(k)$ indicate the warping of the sequence before obtaining the difference between the two sequences. Figure adapted from Sakoe and Chiba (1978).

methods. These applications used to feature inaccuracies as words can be pronounced at different paces. The technique was established under two main conditions:

Condition 1: The two segments feature similar sampling rates. However, these segments can, and at times do, feature distinct number of elements (length) (Sakoe and Chiba, 1978).

Condition 2: The researchers did not posses any *a priori* knowledge on segments or on embodied information and its distribution across the recording (Sakoe

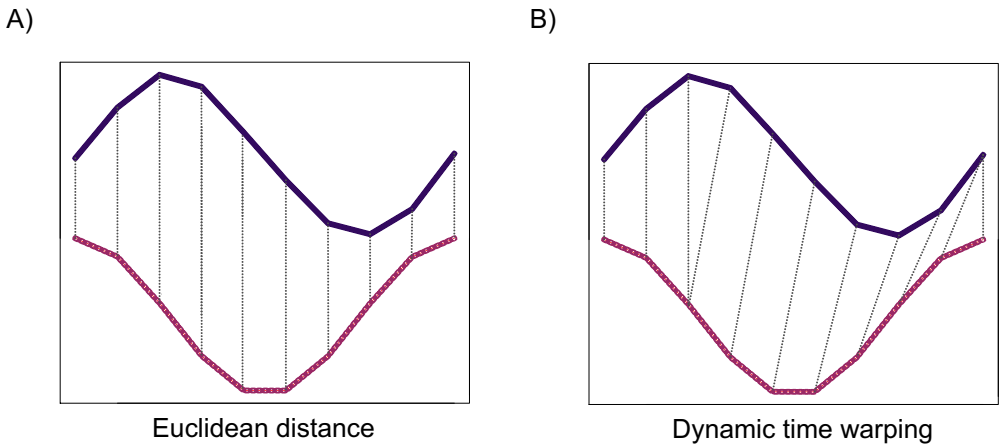

Figure 6.3: Euclidean distance (A) and Dynamic time warping (DTW) (B): Two distinct methods of estimation of similarities between two time series. DTW consists of warping the time series in order to obtain an optimal alignment of both recordings and then calculating the distance between both, whereas traditional Euclidean distance just compares each of the elements in sequences in a pair-wise fashion. Moreover, DTW coincides with Euclidean distance if the two recordings already exhibit optimal alignment.

and Chiba, 1978).

In order to optimise the computational solution of this problem, some restrictions are imposed for the warping function. Besides having boundary conditions defining the beginning and end points ($C_1 = C(1,1)$ and $C_K = (I,J)$, where K is the length — number of elements — of the warping series), the warping function has two main restrictions. It must also be a monotonic function, i.e., $a_{i-1} \leq a_i \quad \forall a \in A$ and $b_{j-1} \leq b_j \quad \forall b \in B$; and the function must show continuity conditions according to: $a_i - a_{i-1} \leq 1 \quad \forall a \in A$ and $b_j - b_{i-1} \leq 1 \quad \forall b \in B$. These conditions result in equation 6.4 (Sakoe and Chiba, 1978).

$$D(A_i,B_j) = \delta(a_i,a_j) + \min \left\{ \begin{array}{c} D(A_{i-1},B_{j-1}) \\ D(A_i,B_{j-1}) \\ D(A_{i-1},B_j) \end{array} \right\} \tag{6.4}$$

Figure 6.4 shows the previously mentioned restrictions to the warping function. An example of use of DTW for two distinct time series is shown in figure 6.5. In this study, DTW has been applied to the first component of multifractal spectra width from the set of epileptic seizures described earlier in this chapter.

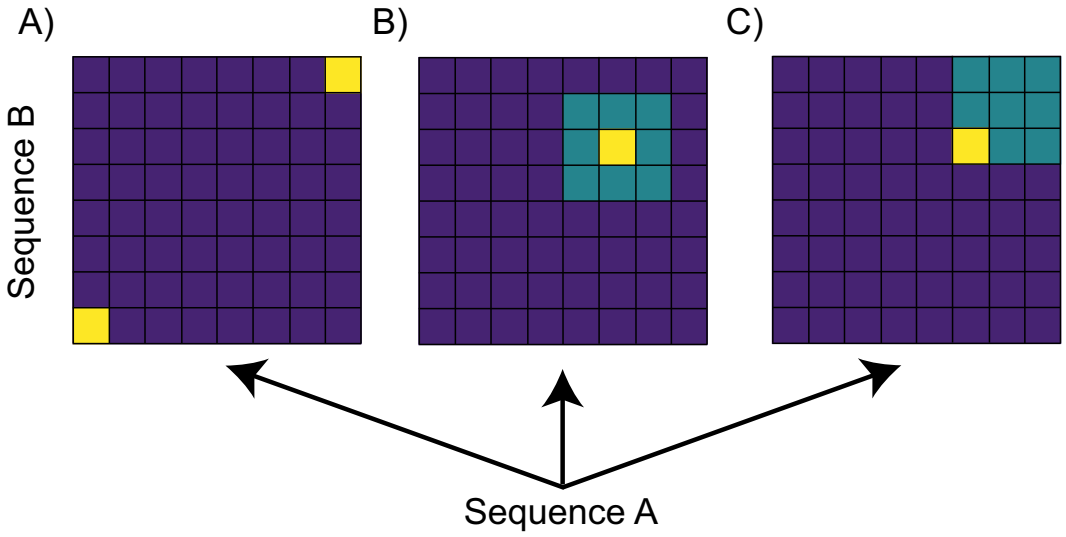

Figure 6.4: Restrictions to optimise DTW algorithm. A) Boundary conditions ($C_1 = C(1,1)$ and $C_K = (I,J)$, where K is the length - number of elements - of the warping series). B) Continuity condition ($a_i - a_{i-1} \leq 1 \quad \forall a \in A$ and $b_j - b_{i-1} \leq 1 \quad \forall b \in B$). C) Monotonicity condition ($a_{i-1} \leq a_i \quad \forall a \in A$ and $b_{j-1} \leq b_j \quad \forall b \in B$).

6.2.5 Distance matrix

A distance matrix is a square matrix that shows the pair-wise distances, in most applications with Euclidean distance (Gentle, 2007), of a set of elements. Considering a set of points $X = \{x_1, x_2, x_3, x_4\}$, according to 6.2, the Euclidean distance by the points in the set X will be given by equation 6.5.

$$m_{ij} = m(x_i, x_j) = \sqrt{(x_j - x_i)^2} \tag{6.5}$$

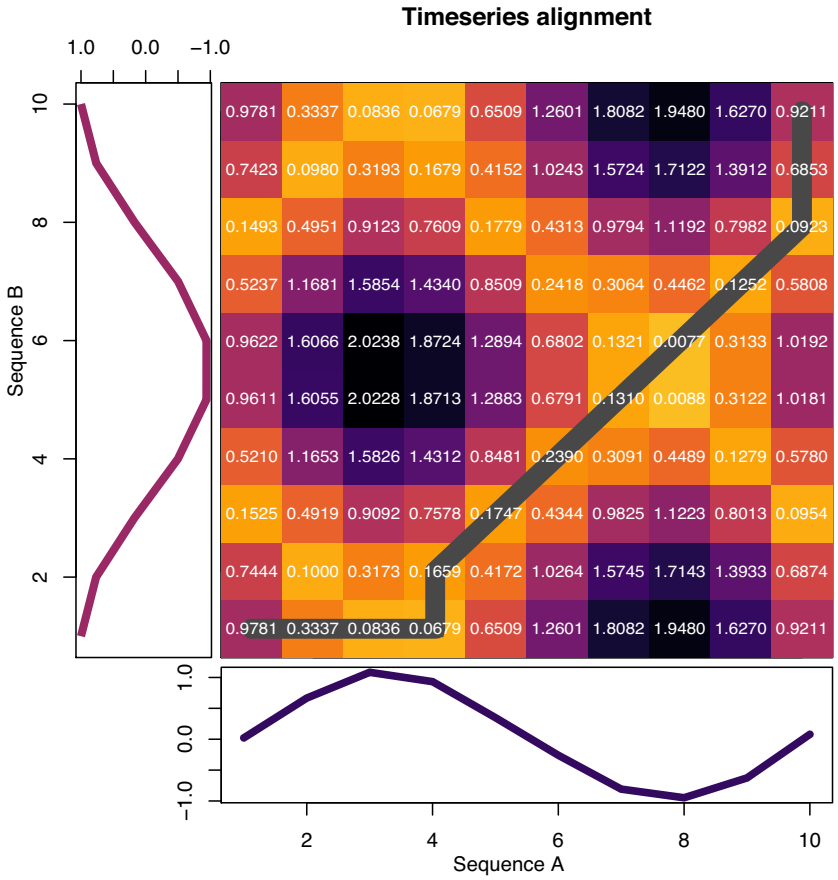

Figure 6.5: Warping function C (grey line) for optimal alignment of the sequences A and B. The devised warping functions follows the boundary conditions established, as well as the monotonicity and continuity restrictions.

A matrix M_{ij} can also be generated from the pair-wise distances between all the elements of the set and has the from shown in the matrix below.

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} \\ m_{31} & m_{32} & m_{33} & m_{34} & m_{35} \\ m_{41} & m_{42} & m_{43} & m_{44} & m_{45} \\ m_{51} & m_{52} & m_{53} & m_{54} & m_{55} \end{pmatrix}$$

Figure 6.6 shows an example of calculation of a distance matrix with a Euclidean metric. However, a similar array to the distance matrix can also be built with different metrics, such as DTW. This chapter reports on results of dissimilarity matrices built from DTW of first components of multifractal spectra width.

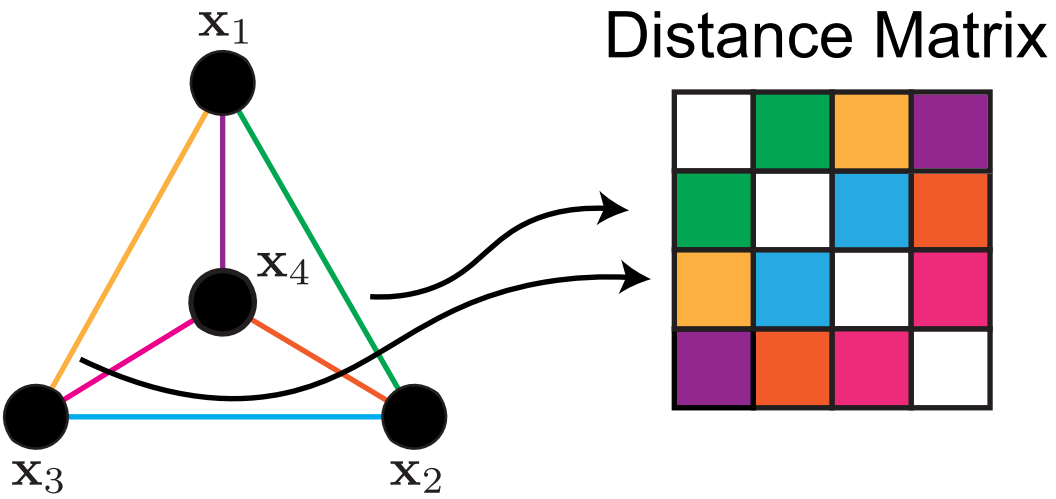

Figure 6.6: Distance matrix. Each element of the matrix shows a pair-wise distance between elements of the set of the figures in the left. Adapted from arXiv:1502.07541v2 (Dokmanic et al., 2015).

6.2.6 Hierarchical clustering

Hierarchical clustering analysis (HCA) is a class of approaches that look to find classes in data and to evaluate their ranking. The outcome of the analysis consists of a dendrogram that shows how smaller structures merge into bigger clusters according to different similarity levels (Camastra and Vinciarelli, 2015).

HCAs can be classified into two different groups: the agglomerative hierarchical clustering methods — that consists of forming smaller clusters and subsequently merging them, and the divisive clustering approaches — that consist of starting with all elements belonging to a single cluster that is subsequently divided in smaller classes (Berndt and Clifford, 1994).

For the agglomerative hierarchical clustering methods, the algorithm consists of considering all the n elements as individual clusters and obtaining a dissimilarity measure, such as Euclidean distance, for all the $n(n-1)/2$ pair-wise combinations (James et al., 2013).

The smallest distance between two of the clusters amongst all pair-wise comparisons is found and these are connected at a height equal to the distance between them. Then the new set of pair-wise dissimilarities is obtained. The procedure is repeated until all clusters are connected (James et al., 2013).

Linkage is the name attributed to the process of obtaining the dissimilarity between cluster groups, i.e. between the new connected branch and the other clusters. There are four main types of linkage: complete — that computes all the dissimilarities between two clusters A and B and chooses the largest of these, single — that is similar to complete but chooses the smallest dissimilarity, average of the dissimilarities of the clusters, and centroid — that is given by the dissimilarities between the centroids of the clusters (James et al., 2013; Berndt and Clifford, 1994).

In this study, the complete-linkage type was applied to a dissimilarity matrix obtained with DTW of the first component of the multifractal spectra width of ictal segments from all patients.

6.2.7 Seizure classification approach

This section describes the procedure for deriving the focal seizures hierarchical clusters. The steps described below apply the techniques previously described in this chapter.

- The EEG signal montage is changed to bipolar to mitigate potential effects in case of artefacts in the reference electrode in the final results.
- Seizure clips starting 60 seconds before the onset and finishing 10 seconds after the termination have their multifractal spectra widths calculated according to the procedure described in chapter 3.
- The above procedures result in spectra widths for all channels of the original EEG signals.
- The spectral widths are standardised to feature a zero mean and standard deviation of 1 with z-scores.
- PCA is applied to the standardised values of the spectral widths.
- The principal components scores are compared with DTW and a dissimilarity matrix is obtained.
- The hierarchical cluster for the focal seizures is then derived from the dissimilarity matrix.
- A similar procedure is performed with line length measure, also described in chapter 3, in order to compare multifractal measures to linear approaches.

6.3 Results

6.3.1 How many principal components are necessary?

One of the challenges involved in applying dimensionality reduction to data, and, in this case, in principal component analysis, is to choose how many variables will

describe the evaluated data.

The number of principal components employed to describe the phenomenon changes from application to application and the according to the nature of the data analysed.

In order to evaluate the number of principal components a scree plot for every patient was plotted. Figure 6.7 shows how much information is contained in every principal component for the first ten components. The aim is to minimise the amount of variables and to increase the information contained in the reduced dimensionality data.

On visual inspection, figure 6.7 shows that the first component features enough variance — illustrated by the sudden drop. Only patient id 'Study 022' shows a smoother decay in the relevance of the subsequent principal components.

6.3.2 First components

The first components show the projection of the data in the direction of maximum variance. In this chapter, only the first component was considered, given that it comprehends most of the variance - as seen in figure 6.7.

Figure 6.8 shows the first components of both MF and LL measures for all seizures across patients, and the duration of each seizure epoch. MF measures show different patterns to the ones exhibited by LL measures.

Moreover, differently from LL measures, MF measures of distinct seizures from the same patient show different features. For LL measures they often look similar and feature a peak after the seizure onset.

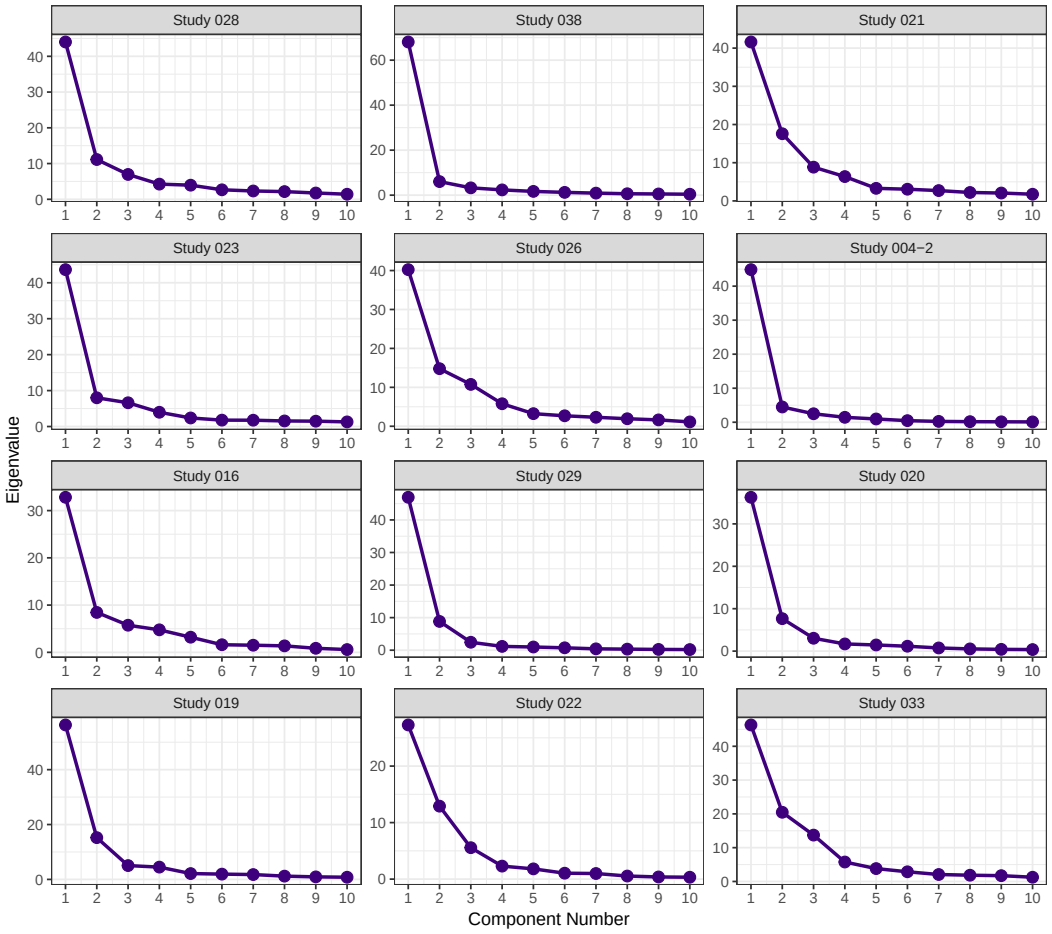

Figure 6.7: How much variance the principal components feature. The scree plots show the amount of variance explained by each of the ten first components. The first component comprehends most of the variance.

6.3.3 Dissimilarity matrix

Dissimilarity matrices were estimated from the principal components of both MF and LL measures of all seizures with DTW. The matrices show smaller differences between seizures of different patients, suggesting a cluster structure for the data. Moreover, LL measures show more pronounced differences between seizures than the equivalent MF time series.

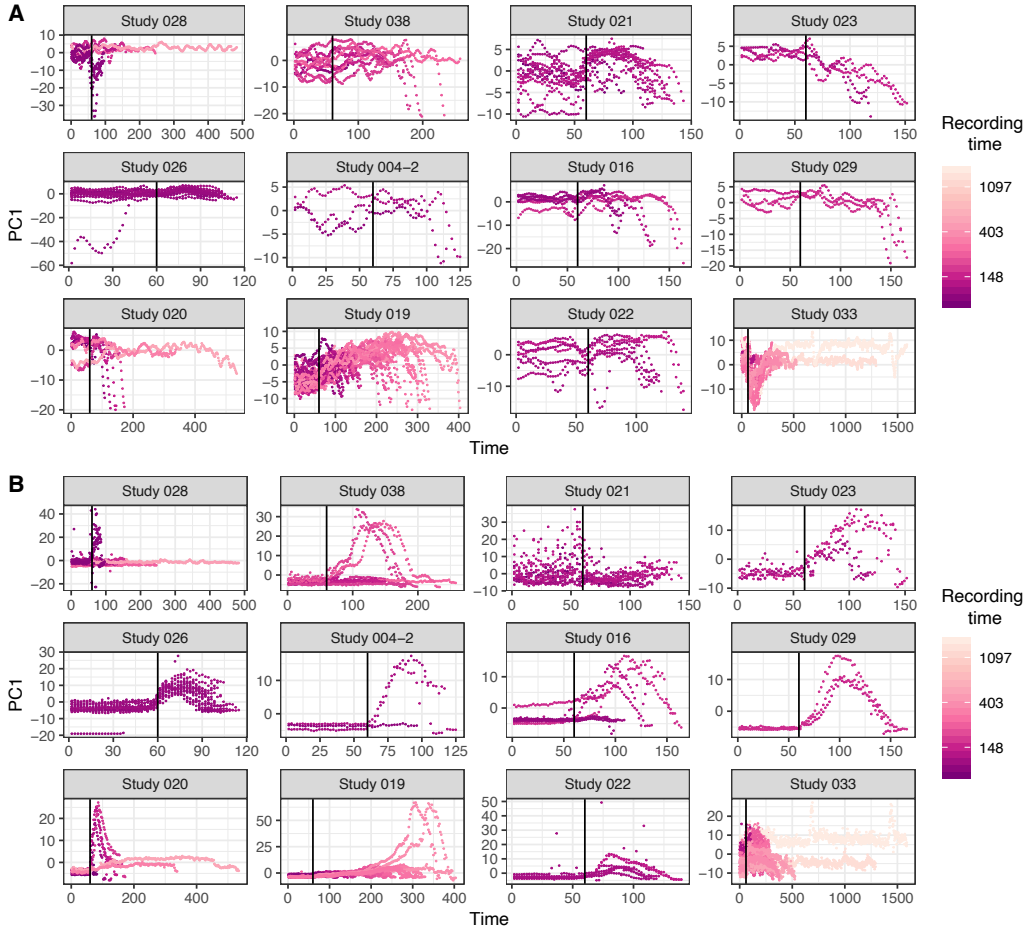

Figure 6.8: First components in **A** MF and **B** LL measures. The figure shows the variation of the first component in time for all 12 patients. The black vertical line marks the seizure onset — 60 seconds after the beginning of the segment. Differently from MF measures, LL curves tend to show differences late after the seizure onset.

6.3.4 Seizure clusters

A hierarchical cluster structure was obtained from the dissimilarity matrix for the first components of the epileptic seizures. The dendrogram showing the structure is available in figure 6.10. The hierarchical cluster shows an outlier that branches out from the remaining branches above the height of 6500. Overall, the other elements split at lower heights — with a maximum of below 2500.

When it comes to bad surgical outcome patients, their seizures seem, on vi-

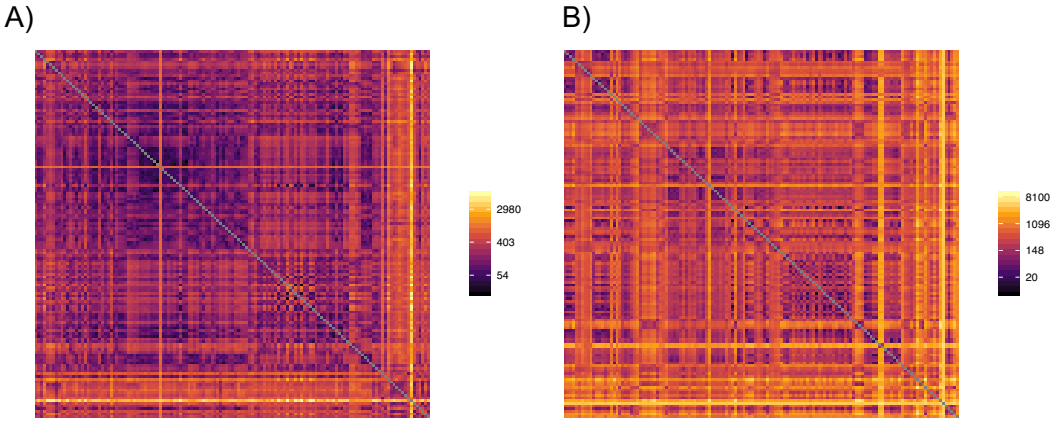

Figure 6.9: Dissimilarity matrices in A) MF and B) LL measures obtained with dynamic time warping (DTW). The two dissimilarity matrices are different and suggest that MF measures provide additional information on the similarities between different seizures when compared to linear metrics such as LL.

sual inspection, spread across the whole dendrogram. However, the first half of the dendrogram shows a higher concentration of seizures from bad outcome cases. Figure 6.11 exhibits a zoomed plot of the dendrogram with labels coloured according to the surgical outcome cases — with blue and red for good and bad outcomes, respectively.

In order to evaluate how the different seizures group into classes, the hierarchical structure and the dendrogram must be cut at a specific height. The choice of height in which the dendrogram is cut varies according to the application and the nature of the data analysed.

In this chapter, four different levels for dendrogram cutting are shown ($\sim$1000, $\sim$700, $\sim$550, and, $\sim$400) and the final height value is chosen according to the similarity between the curves in each seizure group. The curves are warped into a common seizure that belongs to the group with DTW.

At the first height level — $\sim$1000 — the dendrogram features 5 groups of

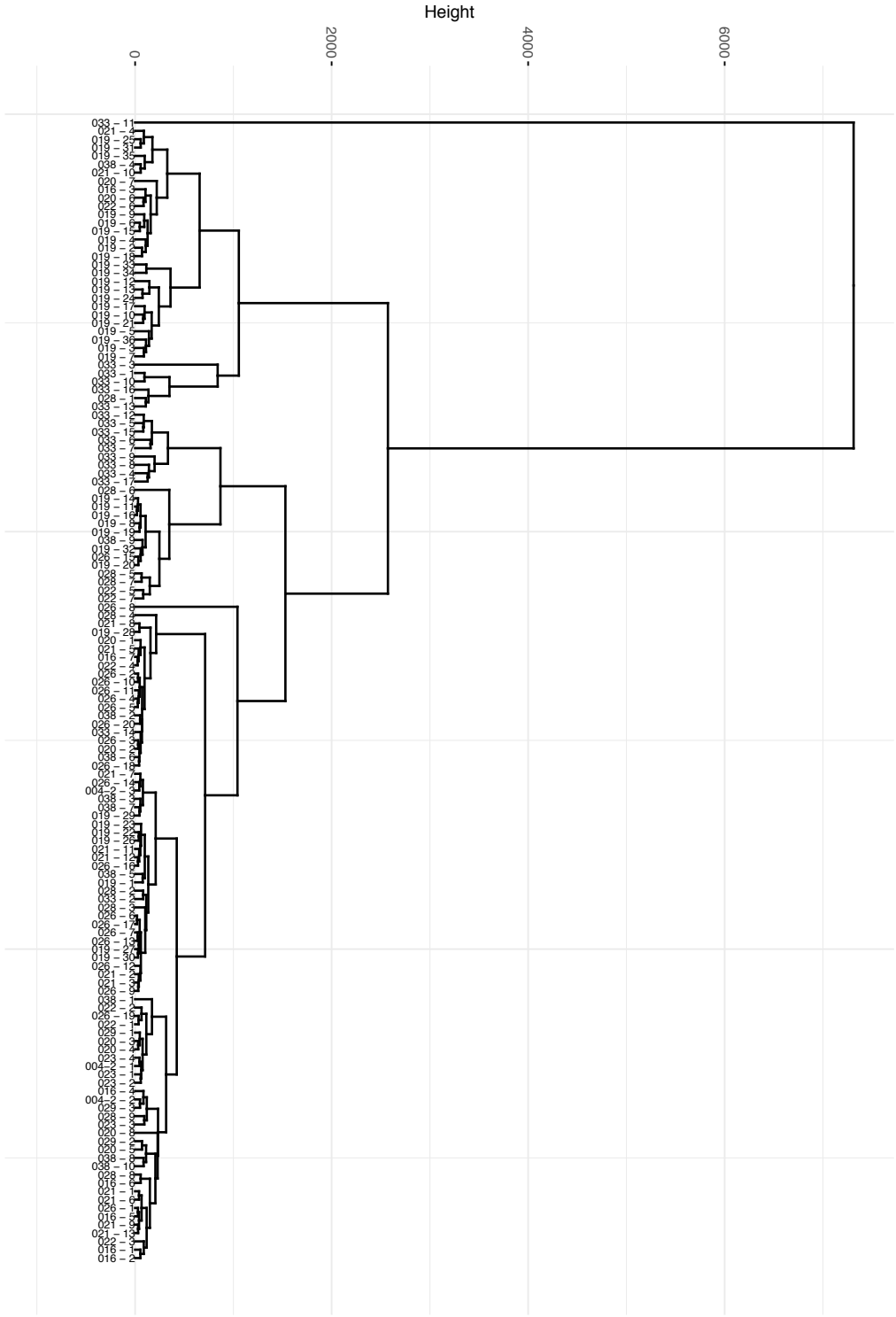

Figure 6.10: Hierarchical cluster for all seizures obtained with dynamic time warping. The dendrogram only features one outlier (033-11).

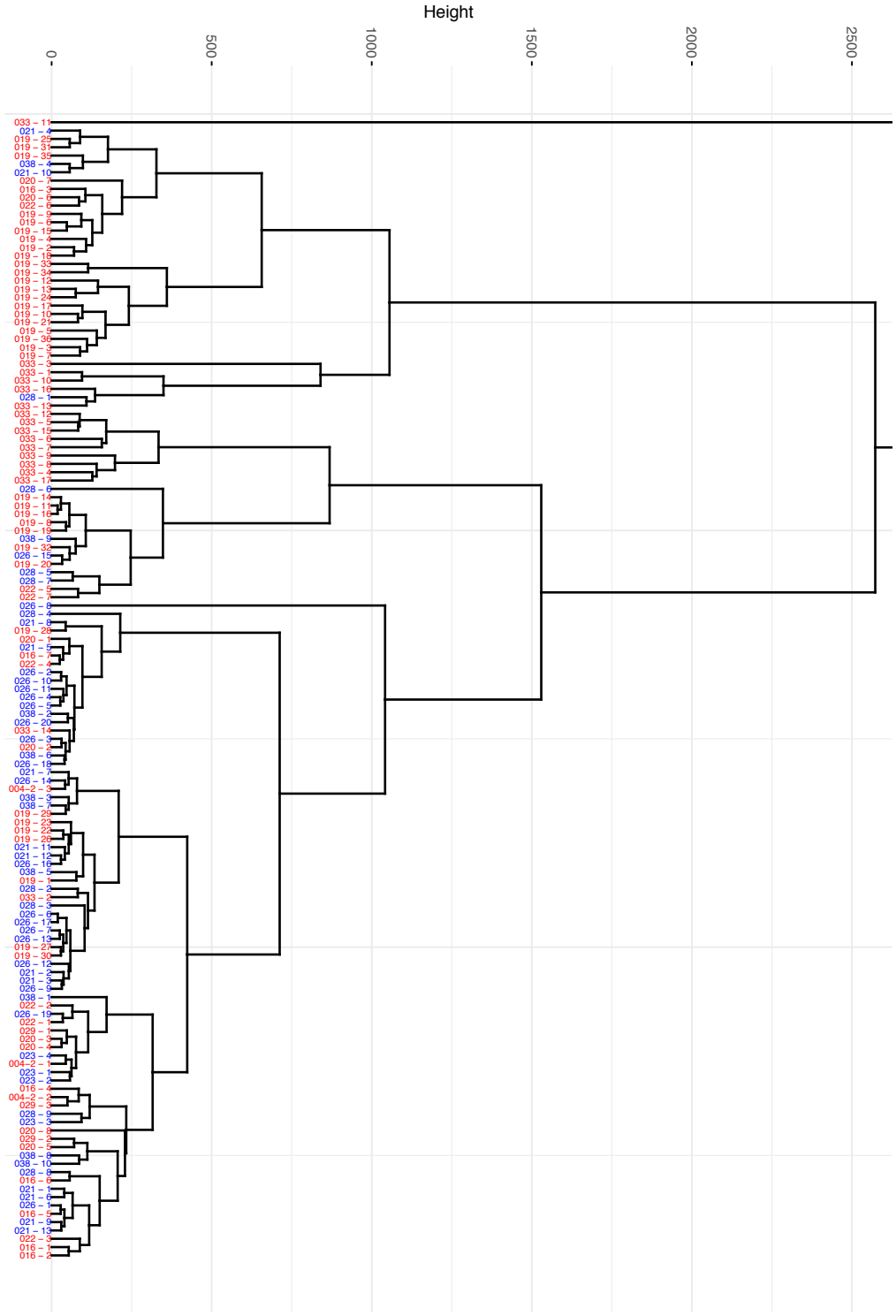

Figure 6.11: Hierarchical cluster for all seizures obtained with dynamic time warping and surgical outcomes. Bad surgical outcome patients seem to feature seizure in different branches; however, some areas of the dendrogram show a higher concentration of those seizures.

seizures, as seen in figure 6.12. One of these groups features a single seizure only — the outlier mentioned earlier. Overall, the curves in the groups suggest that a more refined separation would provide better results.

At the second height level — ∼700 — the dendrogram features 9 groups of seizures, as seen in figure 6.13 — with three single-seizure groups. The curves show a more refined aspect when compared to the previous plots; however, there is still room to make some of them (blue and purple) less noisy.

The next level — ∼550 — holds yet another improvement to the aspect of the curves, as shown by 6.14. Although the dark purple group still features a not defined look. This cut level features 10 seizure groups — with three single-seizure classes.

The last level — ∼400 — features 11 groups, with three single-seizure classes. In this cut the curves look more defined and less noisy than the previous tries. From now on all analyses will consider a cutting at ∼400.

Figure 6.16 is an alternative version of figure 6.15 that additionally features coloured labels to indicate surgical outcomes (red for bad and blue for good outcome cases), patient colour-coded dendrogram leaves, and a colour bar that shows the length of the evaluated epoch.

The groups contain seizures from different individuals, except for the cluster #2 and #4 that feature seizures from patients 'Study 019' and 'Study 033' respectively. Moreover, clusters #1, #2, and #3 show longer durations for epileptic seizures.

Patients have seizures featured in distinct number of clusters, from one as seen in 'Study 023' and 'Study 029', to five as seen in 'Study 019', 'Study 028', and 'Study 038'. The number of clusters do not seem to be related to the surgical out-

come of a patient as seen in figure 6.17-**A**.

When it comes to the seizure duration, clusters #1-4 tend to feature longer seizures than the clusters #5-8. However, these groups still show seizures that are either shorter or longer, as seen in figure 6.17-**B**.

At visual inspection of Figure 6.16, different groups seem to feature distinct ratios of bad/good surgical outcome patient seizures. Figure 6.17-**C** shows the occurrence of either-outcome patient seizures for all the identified groups. The first four clusters show a substantive higher count of bad outcome cases, whereas the remaining four clusters do not show any clear trends for either outcomes.

Figures 6.16, 6.17-**B**, and, 6.17-**C** suggest that bad surgical outcome patients would feature longer seizures. The violin plots show the distribution of durations for seizures of patients with both outcomes and in all clusters. Bad-outcome patients seizures show longer durations and feature a bimodal distribution, whereas good-outcome patients seizures present a reduced duration without a bimodal character of durations distribution. These findings can be seen in figure 6.17-**D**.

To evaluate whether there are statistically significant differences between bad and good outcome seizures durations or not, a non-parametric test (Wilcoxon) was performed and returned a $p < 0.001$ ($p = 2.9e-7$) showing a highly significant distinction between the two distributions. The Wilcoxon test was chosen due to the bimodal character of the bad-outcome seizures durations distribution and consequent need for a non-parametric test.

A similar dedrogram was generated for LL measures (and seizures dissimilarity matrix) and is shown in figure 6.18. The structure of the dendrogram/clusters is

different from those shown by MF measures in yet another piece of evidence evidence that MF is capable of providing additional information on seizure phenomena and their dynamics.

The dendrogram was also evaluated in a systematic way with the silhouette method (Rousseeuw, 1987), which measures the quality of fit for clusters. In this metric, the clusters obtained with the MF measure feature higher values for average silhouette width compared to LL up to five groups — when the average silhouette is closer to 1 (> 0.4). For the subsequent measures of average silhouette width (for # of clusters > 6), MF performs slightly worse or similarly to LL.

6.4 Discussion

This chapter introduces a putative approach for classification of focal epileptic seizures. It is important to emphasise that, to date, there is no objective way of classifying such epileptic seizures, as they do not fall into sets of natural classes (Berg et al., 2010).

The approach introduced here tries to devise this potential classification scheme from insights obtained with machine learning techniques that could help the understanding of seizure phenomena and consequently epilepsy — as suggested by Kuhlmann et al. (2018b).

Multifractal spectra width obtained with the approach introduced in chapter 3 were studied with principal component analysis, dynamic time warping, and hierarchical clustering techniques. After selecting parameters and performing the analysis a hierarchical cluster was devised.

The cluster only exhibited one main outlier — a long seizure. The other segments from both good and bad surgical outcomes are spread across the branches of the dendrogram. There is, nevertheless, a concentration of bad outcome cases seizures in one part of the dendrogram. Which might suggest that these branches could be related to different generating or propagation network.

When cut at the height of ~ 400, the seizures fall into eight different groups and three individual seizures (ungrouped). The curves of these eight groups show overall different characteristics and could be connected to different dynamics. The groups also feature seizures from distinct patients — suggesting that there might be some common aspects to seizures of different patients.

The approach employed to evaluate the cluster quality — silhouette method — suggests that MF measures perform better than LL up to 5 clusters. This is also the range in which the average silhouette width is closer to 1 (ideal value). For more clusters MF measures perform similarly or slightly worse than LL. The silhouette method can help diagnose the quality of a fit of a cluster and to define how many different groups exist in the data. However, for focal epileptic seizures this will require clinical studies with bigger datasets.

A possible explanation for bad surgical outcome in epilepsy surgery is the existence of an additional focus for epileptic seizures or an onset that occurs out of the a priori presumed seizure onset zone (Chaudhary et al., 2012). These seizures that arise from distinct regions could show different dynamic properties and be classed in different groups in an application like this. Looking at the count of groups per patient, i.e. in how many of the identified groups each patient features, there is no

difference between good and bad outcome surgical cases. Thus this suggests that the number of groups in which a patient's seizures are in is not predictive of the surgical outcome.

Moreover, it is important to emphasise that groups #1, #2, #3, and #4 feature a higher count of seizures of bad outcome cases. Future studies should investigate whether some groups are predominantly composed of seizures of bad surgical outcomes patients. If that is indeed true, these seizures would show onset and dynamical features that would be inherent to such bad outcome cases. It is also up to further studies to determine what these properties are — these studies could, as suggested by Kuhlmann et al. (2018b) help to understand underlying mechanisms in ictogenesis. The remaining groups, nevertheless, do not show clear trends to either bad or good surgical outcomes.

When it comes to recording time (seizure duration + 70 seconds), groups #1, #2, #3, and #4 concentrate longer seizures, which is most likely related to the fact that the warp of the time series was not normalised according to its length, causing seizure length to determine seizure dissimilarities. Future study normalising by the seizure length will show whether that is the case or not and if the dynamics are indeed connected to the duration of an epileptic seizure. It has been suggested by Cook et al. (2016) that a seizure's origin, propagation network, and clinical manifestations might be connected to its duration. This study suggests that these connections might be even more profound and related, in some level, to the onset and evolution dynamics themselves. Cook et al. (2016) also suggest the need of further studies in order to understand the connections between seizure duration and their

onset/evolution mechanisms.

As bad surgical outcome cases and longer seizure durations seemed to match at visual inspection, an appropriate evaluation was performed in order to check if that is actually a correlated effect. The character of the distributions — including all seizures from all patients — is different for bad and good surgical outcomes. Bad outcome seizure durations feature longer durations and a bimodal distribution — suggesting multiple seizure populations. Bimodal distributions have been found in patients that show poor seizure prediction results in the prediction clinical trial (Cook et al., 2013, 2016). Cook et al. (2016) suggest different onset mechanisms as a potential reason for the multi-modal character of the seizure duration distribution and poor results of prediction algorithms.

Further study might look into connections between bad surgical outcome cases and efficiency of prediction algorithms in those individuals. There is, however, a challenge in performing such a study as intracranial recordings performed in ambulatory patients record for usually a couple of weeks and the prediction clinical trial required a much longer gathering in order to obtain a minimum number of seizures (Kuhlmann et al., 2018b). Nevertheless, the connections between seizure from bad outcome surgical procedures and poor results of prediction algorithms may provide new insights on seizure onset and dynamics.

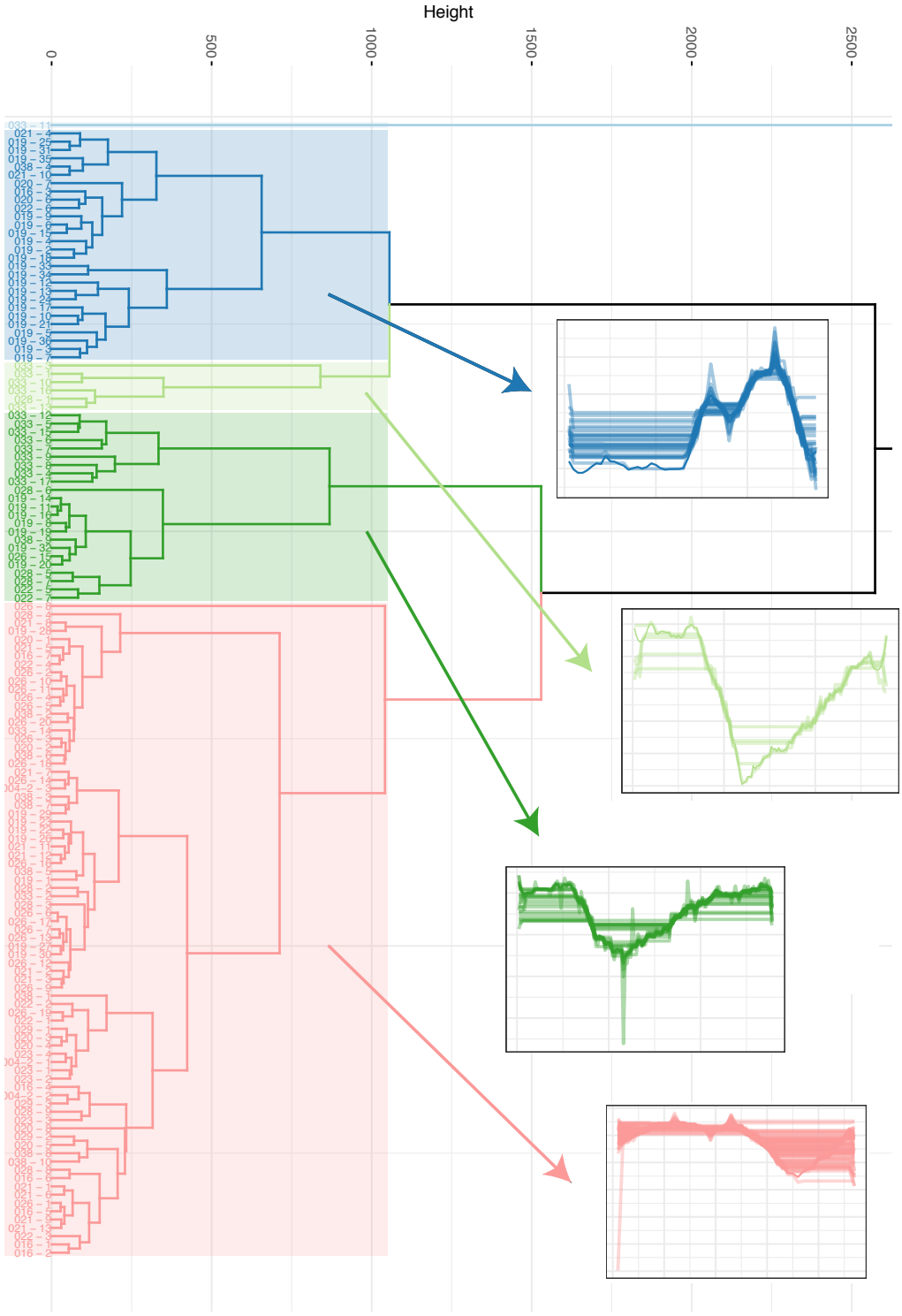

Figure 6.12: Hierarchical cluster for the first components of MF spectral width for all seizures obtained with dynamic time warping cut at the height of ∼1000. With cuttings at this level, the data features 5 different clusters (1 with single seizures). The plots pointed by the arrows show the first components of the seizures in the cluster when warped to a common seizure in the cluster with dynamic time warping. The seizure curves look blurry though, suggesting that the dendrogram should be cut off in a lower level.

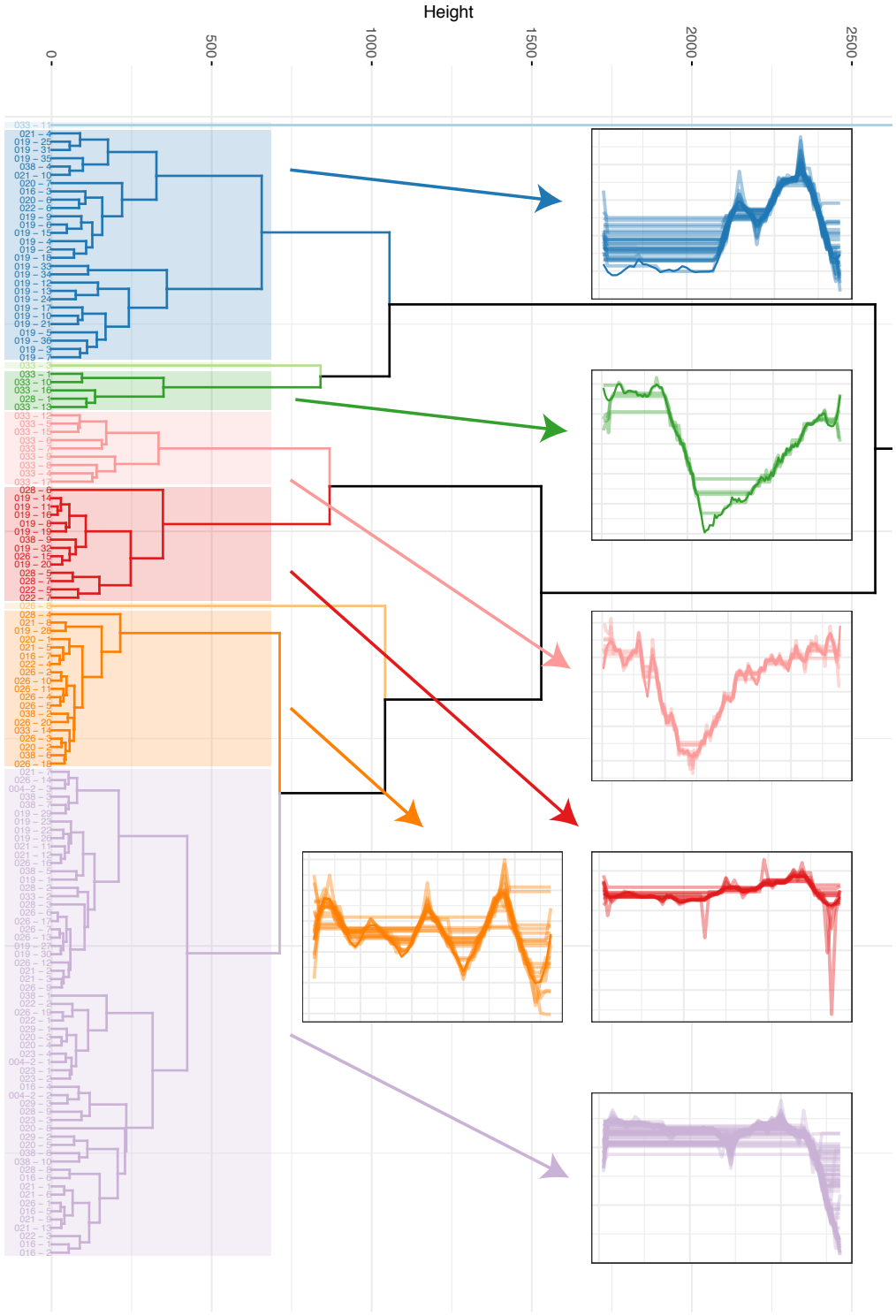

Figure 6.13: Hierarchical cluster for the first components of MF spectral width for all seizures obtained with dynamic time warping cut at the height of ∼700. With cuttings at this level, the data features 9 different clusters (3 with single seizures). The plots pointed by the arrows show the first components of the seizures in the cluster when warped to a common seizure in the cluster with dynamic time warping. Some of the seizure curves look blurry though, suggesting that the dendrogram should be cut off in a lower level.

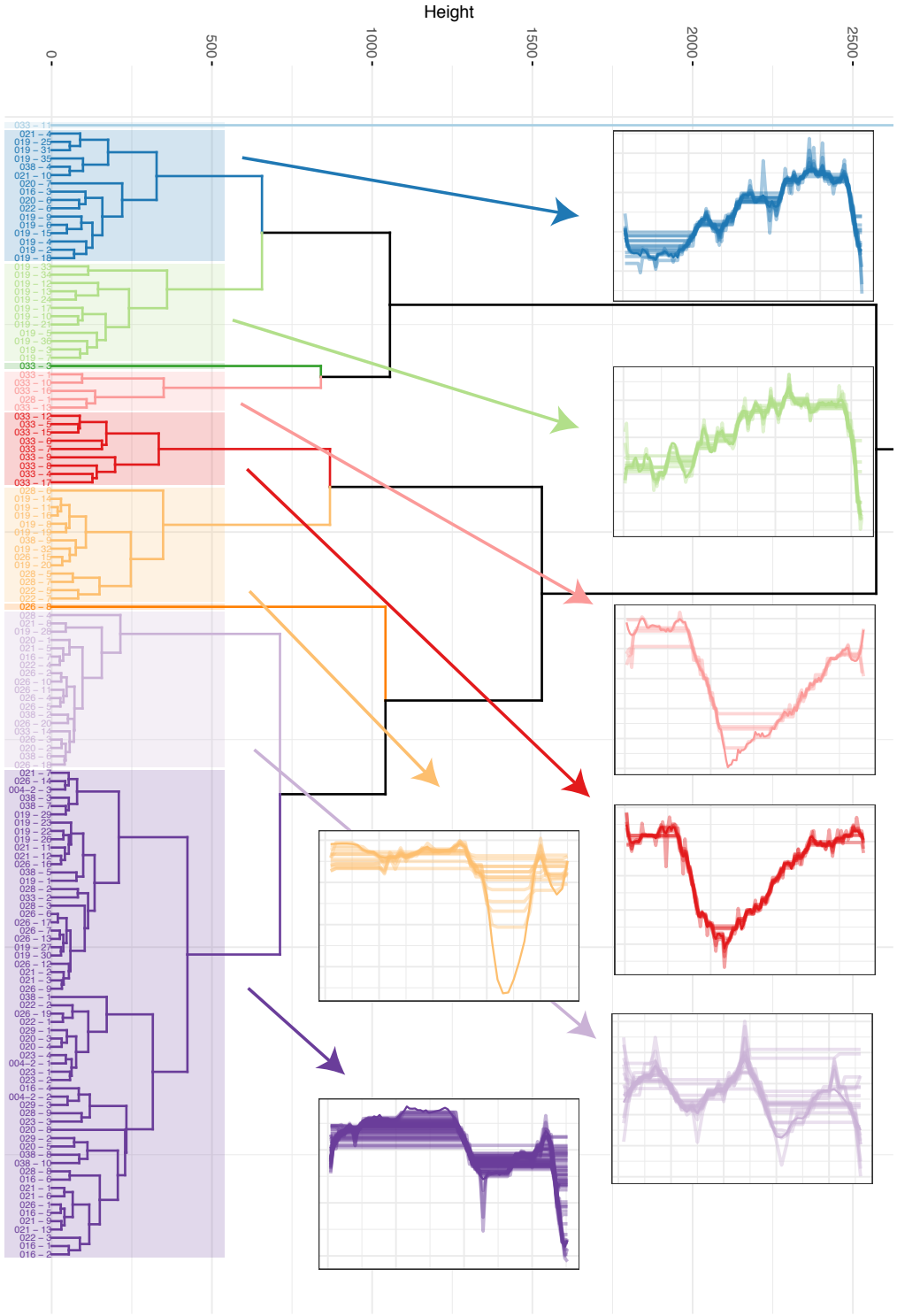

Figure 6.14: Hierarchical cluster for the first components of MF spectral width for all seizures obtained with dynamic time warping cut at the height of ∼550. With cuttings at this level, the data features 10 different clusters (3 with single seizures). The plots pointed by the arrows show the first components of the seizures in the cluster when warped to a common seizure in the cluster with dynamic time warping. Some of the seizure curves look blurry though, suggesting that the dendrogram should be cut off in a lower level.

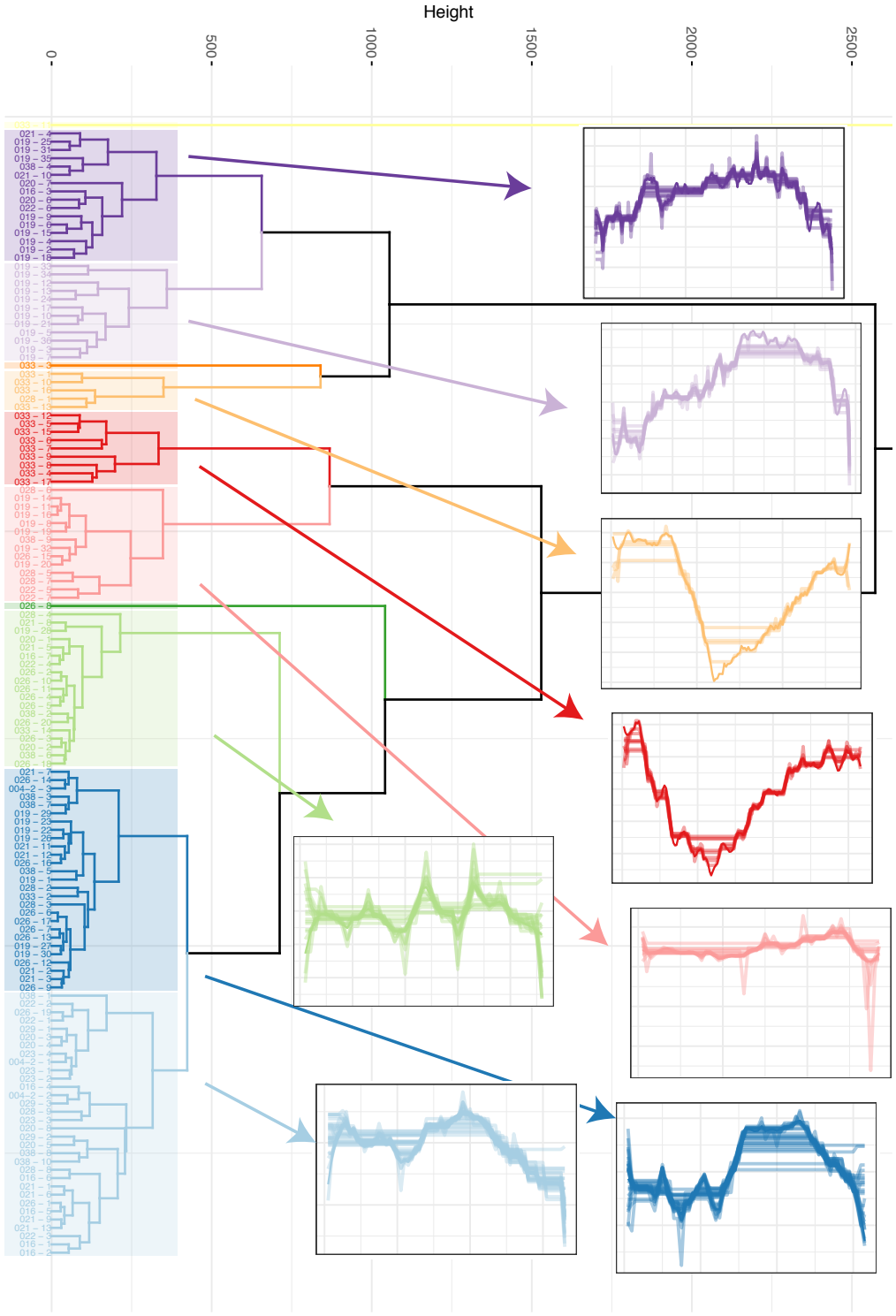

Figure 6.15: Hierarchical cluster for the first components of MF spectral width for all seizures obtained with dynamic time warping cut at the height of ∼400. With cuttings at this level, the data features 11 different clusters (3 with single seizures). The plots pointed by the arrows show the first components of the seizures in the cluster when warped to a common seizure in the cluster with dynamic time warping. At this level, the curves look more defined. This should be the height of the cut employed in the analysis of this chapter.

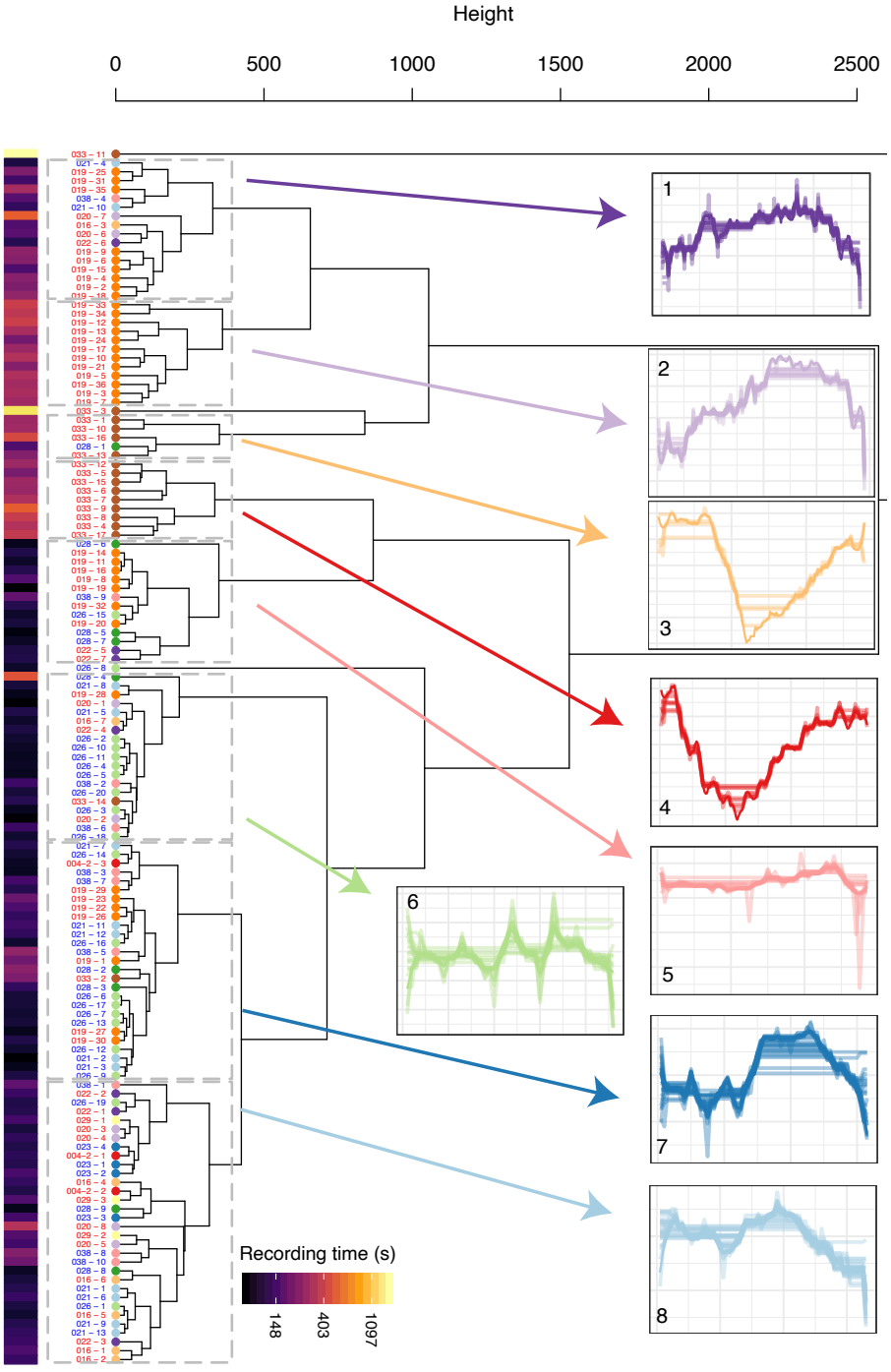

Figure 6.16: Hierarchical cluster, seizure duration and surgical outcome for all seizures. Cluster cut at the height of ∼400. Patient and seizure numbers were colour-coded according to the patient's surgical outcome in red and blue for bad and good outcomes, respectively. Leaves were coloured according to patient and the bar underneath the dendrogram shows the seizure duration for every recording in the cluster. Patients do have seizures spread across different clusters. All clusters show bad outcome patients' seizures, though the number of these elements vary. Some of the clusters seem to feature longer seizures at visual inspection.

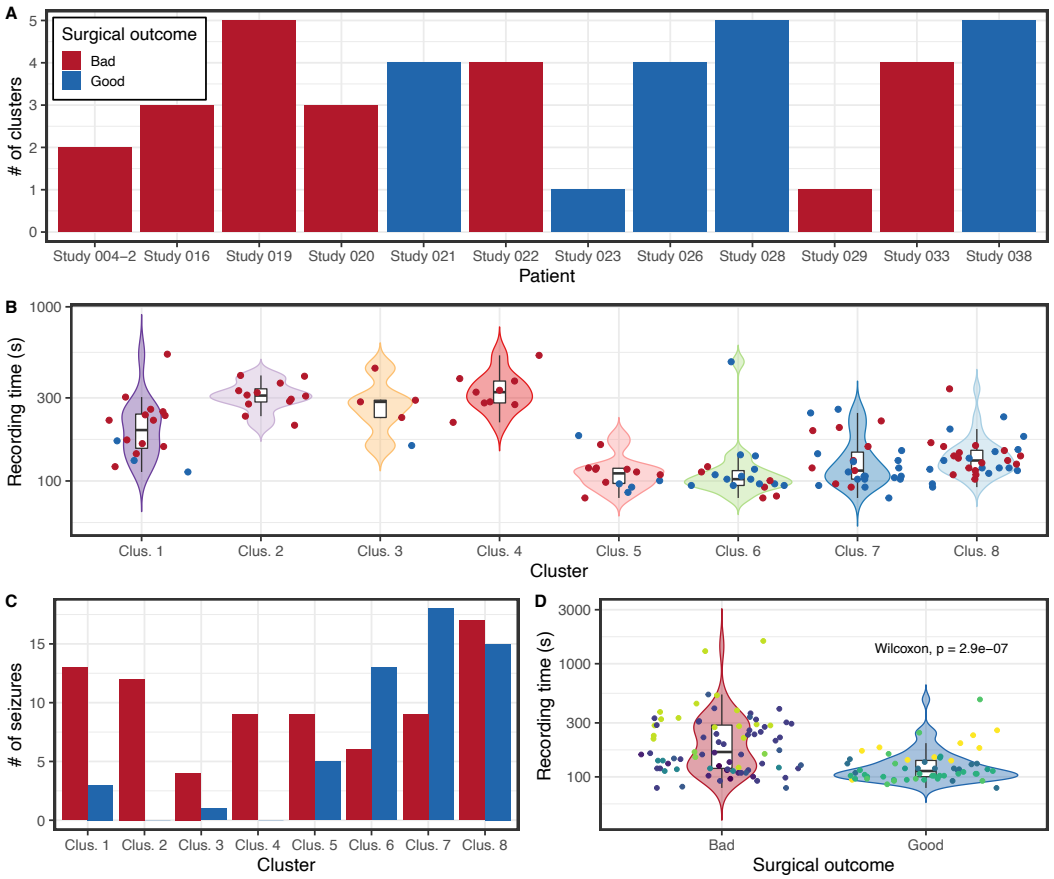

Figure 6.17: Evaluation of the clusters of epileptic seizures. **A** Count of clusters featured per patients for both good and bad outcome cases. At visual inspection, good and bad outcome cases do not feature different numbers of clusters. **B** Violin plots showing the distribution of seizure durations per cluster. The curves suggest that some clusters (#1 - #4) comprehend longer seizures. **C** Count of seizures from both good and bad outcome cases. Some clusters (#1 - #4) are predominantly or exclusively composed by bad outcome cases. **D** Violin plots showing the distribution of seizure durations according to the surgical outcome of the patient. There is a statistically significant difference for the distributions of durations for good and bad surgical outcome case.

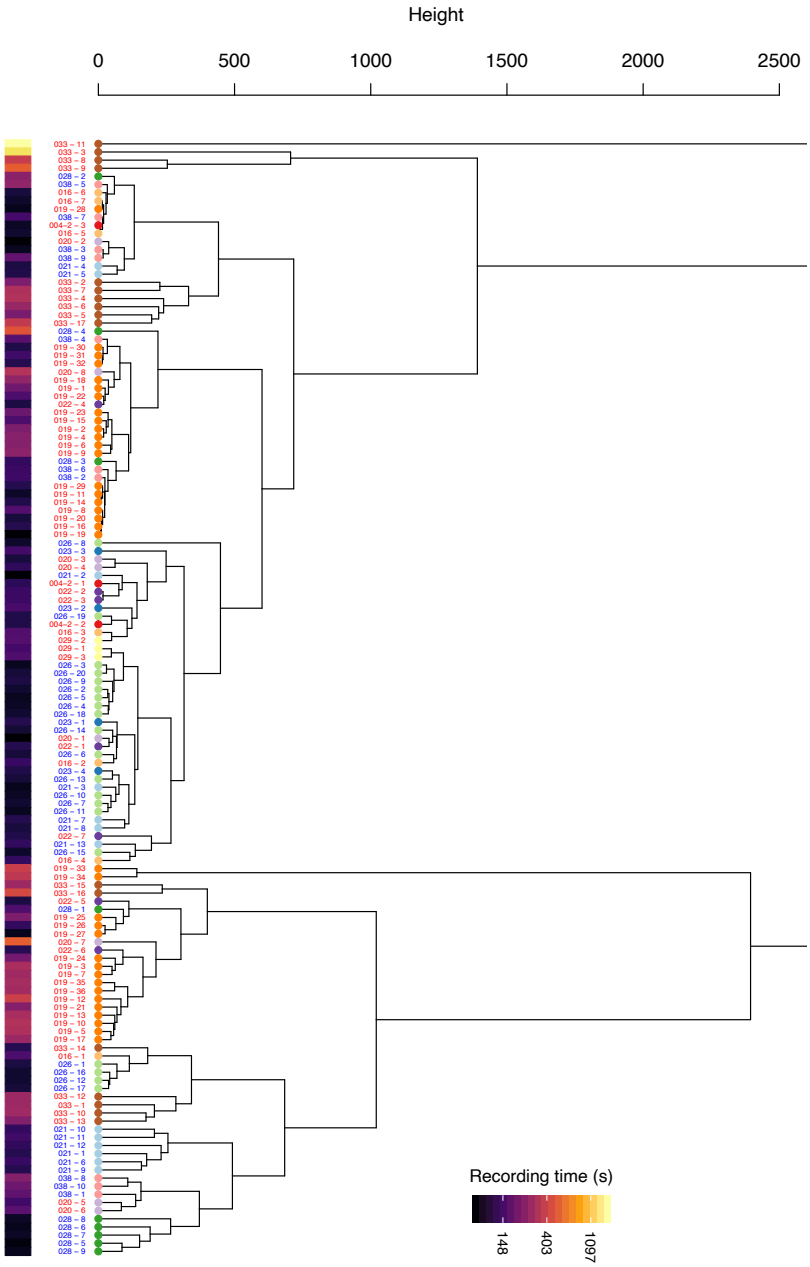

Figure 6.18: Hierarchical cluster, recording time and surgical outcome for all seizures. Cluster obtained from the first components of LL for with dynamic time warping. Patient and seizure numbers were colour-coded according to the patient's surgical outcome in red and blue for bad and good outcomes, respectively. Leaves were coloured according to patient and the bar underneath the dendrogram shows the seizure duration for every recording in the cluster. Patients do have seizures spread across different clusters. The dendrogram obtained with LL shows a different structure to the one obtained with MF measures.

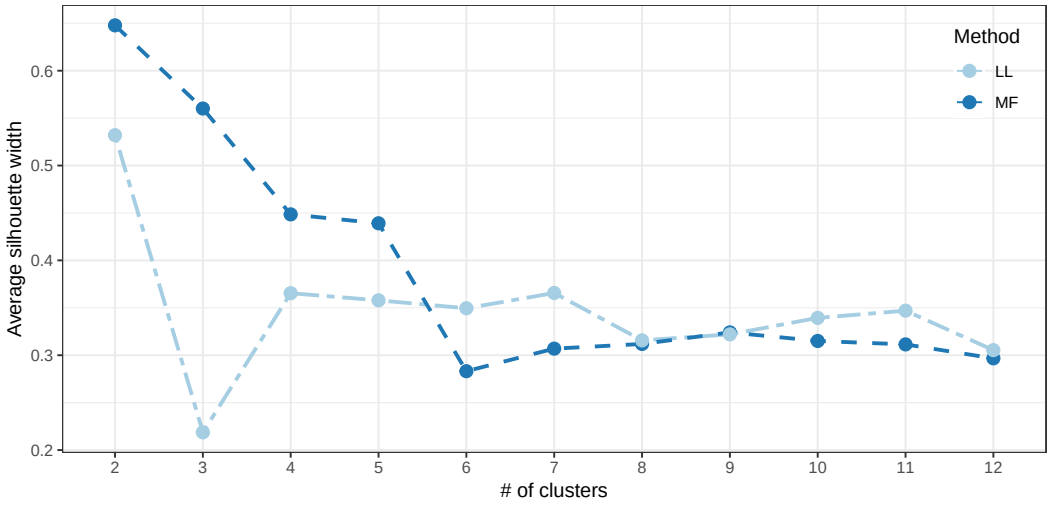

Figure 6.19: Average silhouette for clusters obtained with both MF and LL methods series. MF shows higher quality clustering up to 5 clusters. In subsequent numbers MF measures show similar or slightly worse performance in clustering than LL.

Chapter 7

Scaling properties of local field potential recordings in an animal model of epileptogenesis

7.1 Introduction

Prevention of epilepsy in patients at risk after acquired injury is still not available (Pitkänen and Lukasiuk, 2011; Rizzi et al., 2016; Pitkänen et al., 2016). Anti-epileptogenesis, i.e. a process that opposes epileptogenesis, comprehends prevention, seizure modification and cure (Pitkänen and Engel, 2014).

A putative anti-epileptogenic treatment could be administrated prior to — in order to prevent the disease onset or attenuate the frequency/severity of seizures — or after epilepsy onset (Pitkänen and Engel, 2014; Pitkänen et al., 2016). When applied after the onset of epilepsy, an anti-epileptogenic could reduce the severity of the disease, as well as prevent or slow the progression of the condition. An

anti-epileptic treatment can also modify the dynamics, e.g. from drug-resistant to drug-sensitive (Pitkänen and Engel, 2014).

However, in order to devise such a treatment, biomarkers of epileptogenesis are necessary (Pitkänen and Engel, 2014; Pitkänen et al., 2016). The study of epileptogenesis is still an active area of research (Pitkänen and Engel, 2014). The challenges consist of understanding how epileptogenesis occurs and the changes to the brain that make epileptic seizures happen spontaneously (Pitkänen and Engel, 2014; Pitkänen et al., 2016).

Part of the efforts in acquiring insights into epileptogenesis consist of using animal models. These models of epileptogenesis are generated by inducing status epilepticus (a prolonged seizure) in animals, e.g. rats, and following the evolution until the appearance of spontaneous seizures (Williams et al., 2009; Liu et al., 2016; Jupp et al., 2012; Shekh-Ahmad et al., 2019a,b; Pauletti et al., 2017; Shekh-Ahmad et al., 2018; Hellier et al., 1998).

Status epilepticus causes neuronal damage, inflammation, alterations to underlying neuronal circuitry and neuronal behaviour, and so triggers the epileptogenic process. These models can be produced with both electrical stimulation (Brandt et al., 2016; Pauletti et al., 2017; Rizzi et al., 2019) and chemically, e.g. kainic acid (Hellier et al., 1998; Williams et al., 2009; Liu et al., 2016; Jupp et al., 2012; Shekh-Ahmad et al., 2019a,b, 2018).

These studies can help understanding when neural damage could cause epilepsy and how these processes occur after an insult to the brain. Some studies also point to putative therapeutic interventions to prevent or reduce the impact

of epileptic seizures after neural damage, e.g. stroke and traumatic brain injury (Hellier et al., 1998).

When it comes to the understanding of epileptogenesis, electrophysiological biomarkers might allow a better comprehension of the mechanisms that lead the brain from an insult to a state in which spontaneous seizures occur (Pitkänen et al., 2016). Some previous studies focused on electrophysiological biomarkers for both diagnosis and prognostic evaluations, e.g. pathological high frequency oscillations (HFOs) and markers of seizure threshold/likelihood (Pitkänen et al., 2016).

Another example of such applications is the use of non-linear metrics to evaluate dynamical changes in the brain's behaviour after the initial insult. These techniques analyse long term correlations in the signals and can point to changes in the network of interactions of neurons and neuronal regions — presenting tokens to intrinsic changes that make epileptic seizures more likely. (Rizzi et al., 2016, 2019).

Further research on such markers could provide more insights on how the neuronal damage changes interactions between neurons and neuronal populations in order to generate the increased excitability that could cause spontaneous epileptic seizures.

Chapter 3 shows how multifractal measures can provide additional information on epileptic seizures and their occurrence. Considering the results obtained by Rizzi et al. (2016, 2019), MF measure could also provide additional information on the changes that take place in the brain after an injury/insult.

Hence, this chapter is aimed to evaluate signals from an epileptogenesis experiment and to explore putative information MF measures on changes in the brain

exhibited by electrophysiological signals from an epileptogenesis model in rats.

7.2 Methods

7.2.1 Animal model and data

Sprague-Dawley rats had kainic acid treatment in order to induce status epilepticus and provoke neuronal damage. The procedure is described in more details by Hellier et al. (1998) and Shekh-Ahmad et al. (2018).

The signals presented in this chapter comprises data of fifteen rats that have undergone the procedure. Continuous EEG/EcoG was also recorded from these rats. A subdural electrode was implanted above the right hippocampus and the reference electrode was implanted in the contralateral hemisphere. The apparatus recorded a single channel at 512 Hz in two distinct phases: a baseline stage, in which the data was recorded before the induction of status epilepticus; and a second stage consisting of the epileptogenesis phase, after the neuronal insult and before the subsequent development of spontaneous seizures (first seizure) — in some animals. The series were capped to week-long durations in order to keep consistency across different rats — as the animals develop spontaneous seizures at different paces.

Animals were divided into two groups: the rats that exhibited spontaneous seizures (S), and the rats that did not develop seizures (NS). Some of the rats had been treated to prevent the development of epilepsy (Shekh-Ahmad et al., 2018). The tests were blind and the labels were not available until the end of this analysis.

Table 7.1: Rats IDs classified according to the occurrence of spontaneous epileptic seizures after a neuronal insult. The groups are labelled as S and NS at this point — for animals that show spontaneous seizures and animals that do not yet, respectively.

Rat ID	Group
ID 1	S
ID 2	NS
ID 3	NS
ID 4	NS
ID 5	S
ID 6	S
ID 7	S
ID 8	S
ID 9	NS
ID 10	S
ID 11	S
ID 12	S
ID 13	S
ID 14	NS
ID 15	NS

7.2.2 Multifractal analysis

The multifractal analysis approach presented in chapter 3 was applied to these intracranial EEG obtained from rats as described in previous section.

The Chhabra-Jensen method was applied to time data that was epoch-wise standardised and sigmoid transformed (i.e. the established pipeline introduced in chapter 3). Each epoch was 2 seconds long, as the epochs had 1024 points sampling and sampling frequency at 512 Hz.

7.2.3 Statistics of extremes

Differently from the previous chapters of this thesis, the analysis shown in this chapter consists of long range data, i.e. weeks of recordings — considering all the rats. So a different methodology is necessary when it comes to selecting relevant pieces of data for the study.

Ictal events are rare (Kuhlmann et al., 2018b) — and presumably events related to changes involved in epileptogenesis should follow suit. In order to bypass that issue and remove the potentially uninteresting segments of the time series, the statistics of extremes were performed after the MF analysis of the recordings. Only the MF spectra width bigger than the mean plus three standard deviations were considered when comparing baseline and epileptogenesis phases.

New dynamics and interaction networks caused by the neuronal insult and epileptogenesis could be reflected in the multifractal spectra measures, e.g. these new features could increase the MF spectra width due to an increase in heterogeneity of the signal. Hence the choice for extreme values on MF spectra width.

7.3 Results

The multifractal spectra features exhibited by the rats' electrographic measures show subtle changes between baseline and epileptogenesis phases. Given the amount of measures ($\sim$ 4 mi), it is hard to evaluate the widths and heights for all rats. A possible way to visualise the data consists of employing hexagonal bins to represent the amount of measures in every area of the plot. Figure 7.1 shows the bin count for both baseline and epileptogenesis data.

In both baseline and epileptogenesis phases, it is evident that most of the measures are concentrated around the mean value. Moreover, values that suffer with overfitting are relatively rare with $\log(Count) < 400$. However, at visual inspection most values for width seem to concentrate around the average (slightly below 0.5) - shown by black vertical line. Height measures show some overfitting and result in

values above the topological limit ($height = 1$).

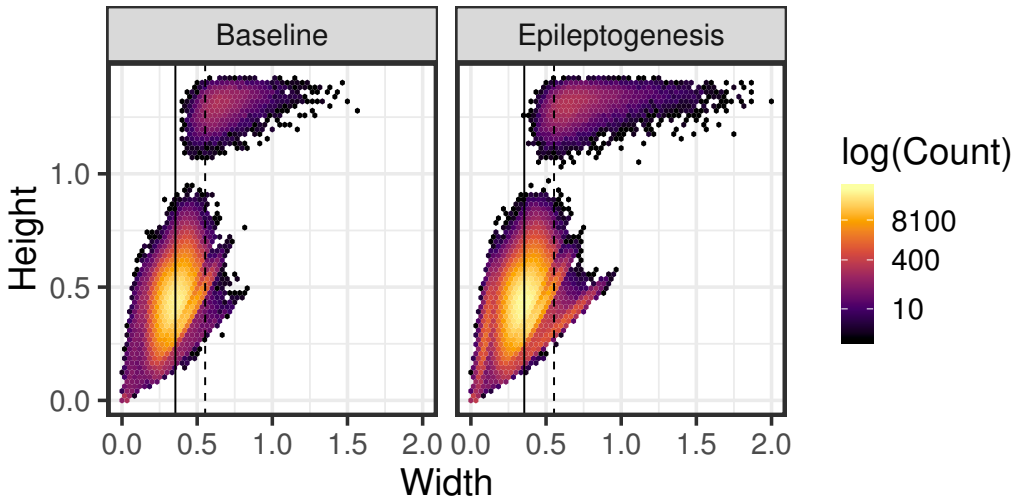

Figure 7.1: Hexagonal bins of width and height for multifractal spectra obtained from all rats for baseline and epileptogenesis stages. The black vertical line shows the mean value for width and the dashed line indicates the value for mean plus three standard deviations. The figure shows the count of points in the different areas and suggests that most of measures are concentrated around the mean width. Moreover, the cluster of the overfitted height measures show relatively few occurrences across the ~ 4 mi measure points.

At visual inspection, figure 7.1 suggests changes on the epileptogenesis phase of the recordings. The widths of the spectra seem to shift towards higher values, suggesting that the brain insult, which the rats have been submitted to, might cause the increase of the multifractal spectra width.

In order to evaluate such putative effects, the percentage of extremes, i.e. widths with values above the mean plus three standard deviations of the distribution, was obtained for every rat in both baseline and epileptogenesis phases. Most rats, except 'ID5', 'ID8', 'ID11', and 'ID12', exhibit similar percentages for both baseline and epileptogenesis — below 1%. The four exceptions mentioned feature notably higher percentages for the epileptogenesis phase, all above 1% and with a peak of $\sim$4% for rat 'ID11'.

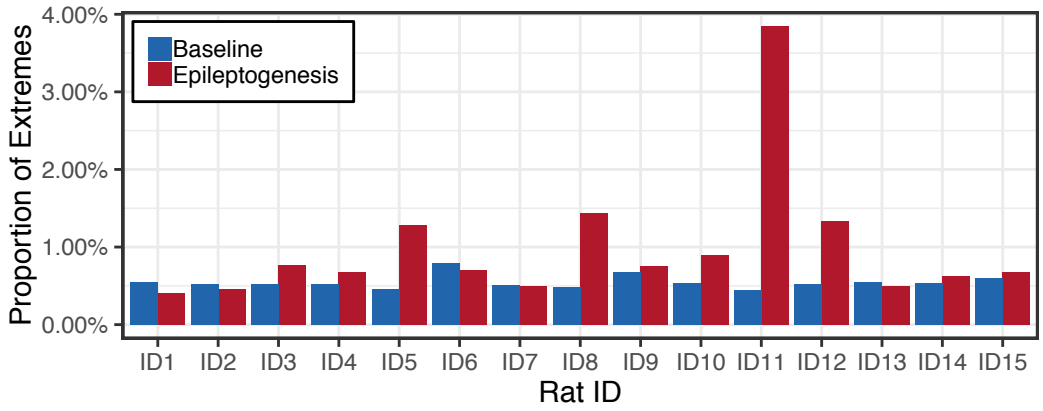

Figure 7.2: Percentage of extremes (larger or equal to mean plus three standard deviations) for each rat in both baseline and epileptogenesis stages. At visual inspection, five rats feature a higher percentage of extremes for the epileptogenesis phase. In particular, four rats show a percentage of extremes above the 1% threshold for the epileptogenesis phase. No rat shows a percentage of extremes above 1% for the baseline phase.

In an alternative view, figure 7.3 shows violin plots of the percentage of extremes for both baseline and epileptogenesis phases. This time rats were divided into two groups concerning the development of spontaneous seizures after the brain insult on the induction of the status epilepticus.

The two groups, at visual inspection, show differences when it comes to variations from baseline to epileptogenesis phase. Both groups exhibit similar percentages of extremes for the baseline phase. Nevertheless, the majority of measures for epileptogenesis phase in S show higher percentages than NS (as indicated by the dashed line) — that features approximately the same percentages for both baseline and epileptogenesis phases.

A hypothesis test does not show statistically significant differences between the two groups, however. A Welch two sample t-test was applied to the differences in percentages of extremes for both groups and returned a non-statistically significant output with $p = 0.117$ and $t = 1.757$.

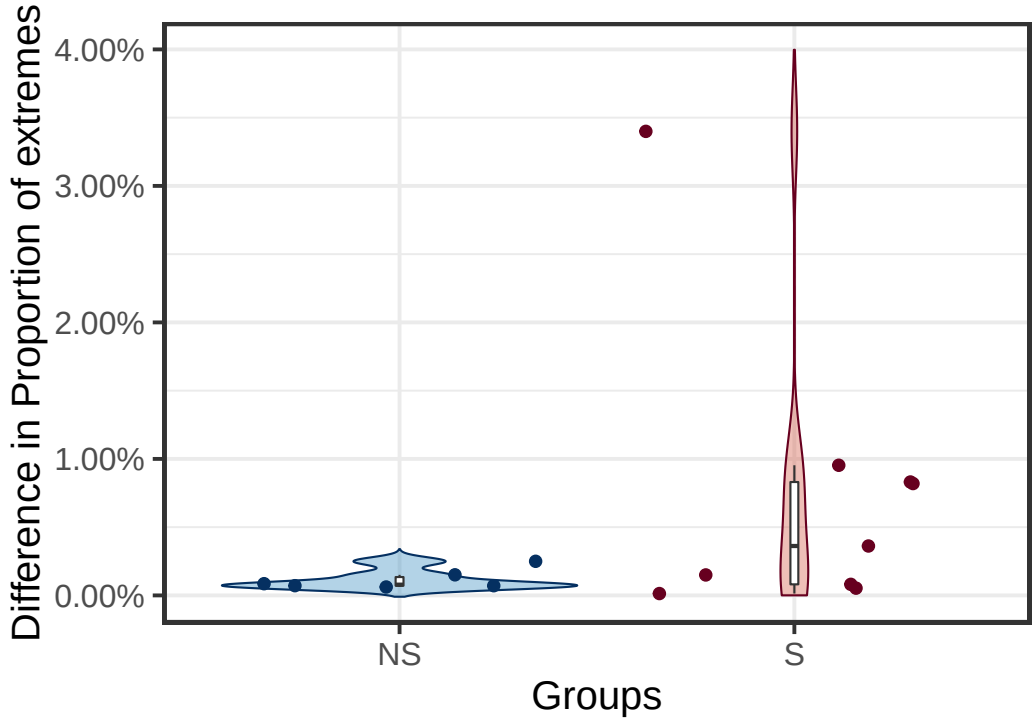

Figure 7.3: Violin plot of the difference in the percentage of extremes for S and NS rats in baseline and epileptogenesis phases. At visual inspection, S seems to feature a higher proportion of epileptogenesis values, which does not show in NS. Such results might indicate that S shows more abnormal activity.

In order to gather a better understanding on the rats of S that show percentages for epileptogenesis phase at similar values to baseline phase, the proportions of extremes were evaluated on a day-to-day basis. The procedure consists of the same steps adopted before, but values are reported for every day of the recording, rather than a global score per rat per phase of monitoring.

Figure 7.4 shows the curves for both epileptogenesis and baseline states in every rat of the experiment. Most of the rats show curves with higher percentages of extremes for the epileptogenesis phase. It is, however, variable and shows increases and drops over the course of days.

In a visualisation similar to figure 7.3, a data point for every day and all rats in

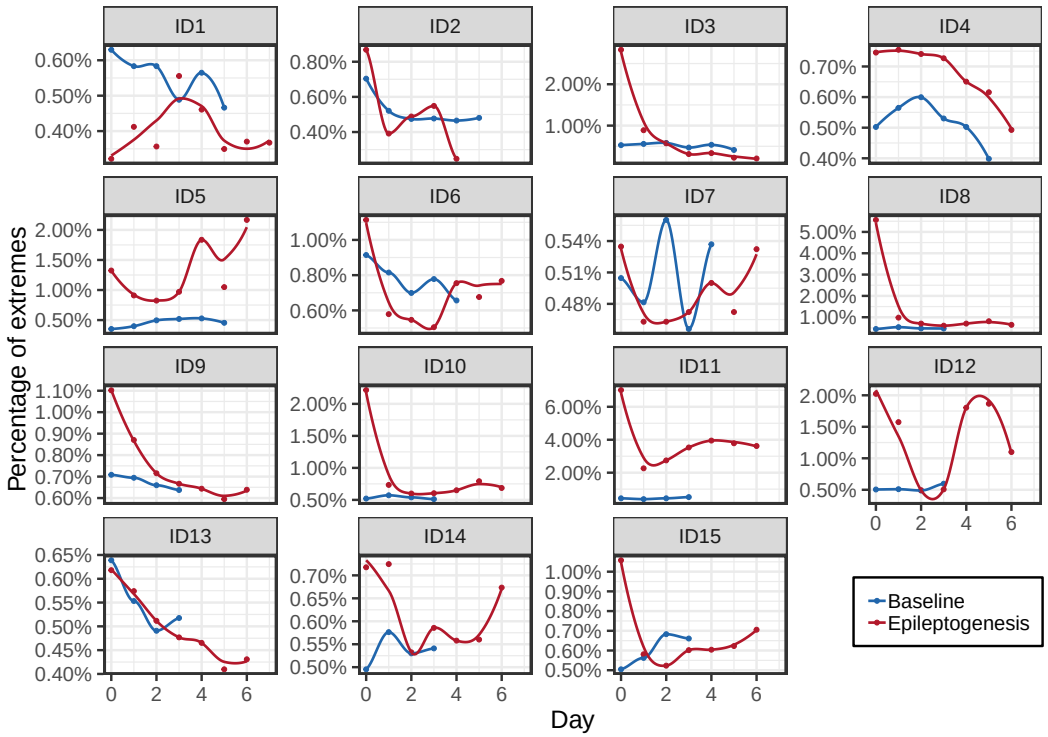

Figure 7.4: Variation of the percentage of extremes (larger or equal to mean plus three standard deviations) for each rat in both baseline and epileptogenesis stages over days. Different rats show distinct patterns. Moreover, the curves show oscillations on the percentage of extremes across days. The solid line represents a LOESS curve fitting of the data points, with formula $y \sim x$.

both S and NS, for both epileptogenesis and baseline phases, was plotted in figure 7.5. The difference between S and NS featured by figure 7.3 is also shown by this version of the plot. This time, however, the difference between S and NS is more pronounced with points related to measures of the epileptogenesis phase for several rats above the mean for the values of the baseline phase.

Further analysis revealed that there are no statistically significant differences between S and NS rats. A Welch two sample t-test was applied to the maximum values of the differences between baseline and epileptogenesis phase measures of the percentage of extremes in S and NS groups and returned a non-statistically significant output with $p = 0.181$ and $t = 1.433$.

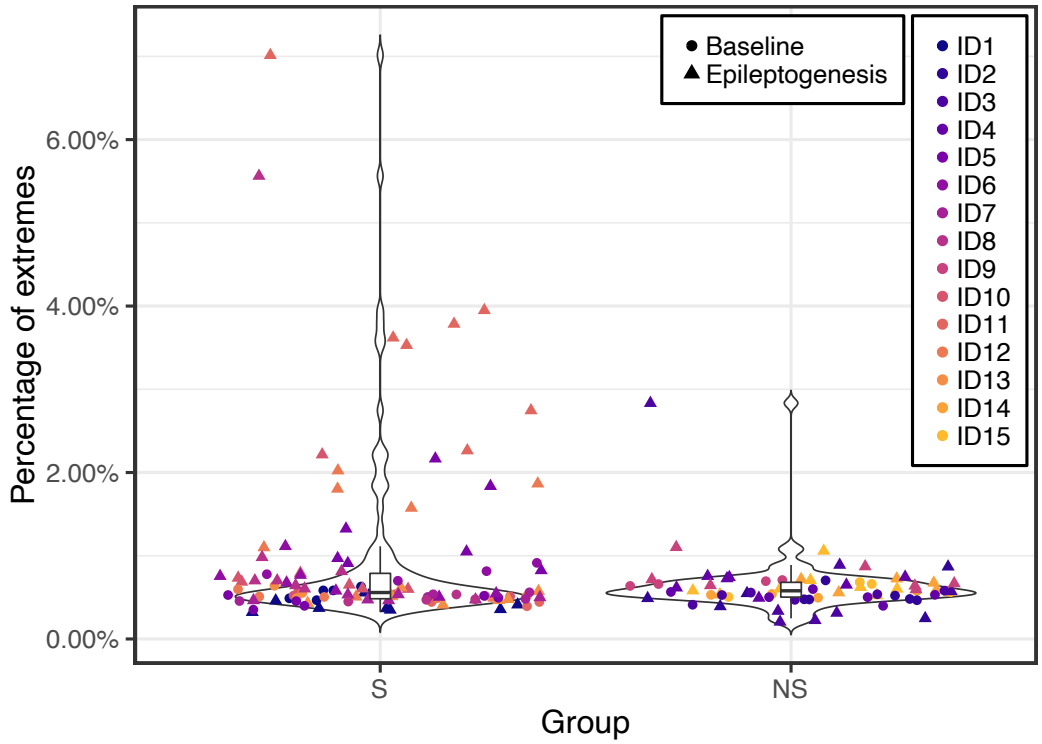

Figure 7.5: Violin plot of percentage of extremes for S and NS rats in both baseline and epileptogenesis phases for daily measures.

After the obtaining these results, it was revealed that S corresponds to the rats that developed spontaneous seizures after the status epilepticus and NS does not show ictal activity after the induced status epilepticus.

7.4 Discussion

There are multiple putative mechanisms underlying epileptogenesis but it has not thus been possible to predict those who will go on to develop epilepsy following a brain insult (Pitkänen and Lukasiuk, 2011; Rizzi et al., 2016; Pitkänen et al., 2016).

This chapter presented a study on brain signals recorded from rats in a long monitoring experiment after induction of status epilepticus with kainic acid. The aim of the experiment was to evaluate and find markers of epileptogenesis that could

be of use in understanding the dynamics of epilepsy and epileptic seizures, as well as putative markers to identify patients that would potentially develop epilepsy after a brain insult such as stroke or traumatic brain injury.

The multifractal approach described in chapter 3 was applied to the signals recorded from fifteen rats of the epileptogenesis animal model described in the previous paragraph. The approach has been applied to epileptic seizures in chapter 3 showing sensitivity to ictal activity. Moreover, if the complexity and nature of the networks are changing with epileptogenesis, multifractal analysis should be sensitive to such changes. This thesis also presents a putative approach employing the same method to the classification of focal epileptic seizures from patients undergoing presurgical evaluation in chapter 6.

The Chhabra-Jensen method (Chhabra and Jensen, 1989) has also been applied to long term monitoring of physiological signals, such as actigraphy signals in patients with fibromyalgia (França et al., 2019). The method was then applied to the longer electrophysiologic signals from animals models of epileptogenesis.

Signals were compared in two distinct phases: the baseline — before the status epilepticus induction — and epileptogenesis - consisting of the phase after the brain insult. The multifractal spectra features showed differences between the baseline and the epileptogenesis phase. In the latter, values seem to shift towards higher values for width of the spectra, suggesting a more heterogeneous recording. Such increase in heterogeneity could be a result from new rhythms caused by the brain insult and be connected to epileptogenesis/ictogenesis.

The Chhabra-Jensen method has been shown to be efficient in the estimation

of dynamical properties of time series, particularly when employed with the standardisation approach mentioned in chapter 3. Nevertheless, when exposed to huge quantities of data in this later experiment, the approach has shown some overfitting with estimated heights of the spectra above the topological limits. The frequency of such issues is relatively small, particularly when compared to the ~ 4 mi measures obtained. In order to avoid such issues, this study focussed on the spectral width, that is not subject to overfitting as much as the spectral height.

One of the difficulties in studying ictal phenomena is the rare occurrence of such episodes (Kuhlmann et al., 2018b). It is also evident from figure 7.1 that most of the data are centred around the mean, suggesting that most of the EEG segments exhibit similar levels of hetereogeneity. It is possible to quantify, however, the occurrence of shifted values as described previously.

The analysis of extremes consists of counting the amount of elements that crosses a certain predefined threshold. In this study, a cut-off limit of the mean value plus three standard deviations was chosen due to the nature of the data. The removal of the other values allows computing information in *a priori* anomalous portions of the data. Probably the simplest way of dealing with extremes consists of calculating the proportion of these in regards to the original raw data.

The percentage of extremes on both baseline and epileptogenesis phases change across rats; however, most of the measures stay below the 1% mark. Four rats, nevertheless, show measures above the 1% cap exhibited by most of the animals. This difference could be a hint that these four animals might show different activity in response to the brain insult during the induction phase. These animals

are candidates to show new rhythms in the EEG — potentially linked to dynamics that generate epileptic seizures (or might generate seizures in future).

This experiment was performed as a blind test and the group of rats that developed spontaneous seizures after the insult in the induction phase was only revealed at the end. Nevertheless, the animals that belonged to the same groups were known beforehand. In order to identify potential statistical differences between these two groups in regard to the percentage of extremes in the multifractal spectra width, a t-test was performed. At visual inspection, S shows higher values for the percentage of extremes — in the epileptogenesis phase. The same cannot be observed for NS. However, a t-test did not show a statistically significant difference between the two groups.

The lack of statistical significance does not support that there are differences between the two groups regarding extremes of multifractal spectra width. However, the pathophysiological aspects of these rats are completely different — one of them presents spontaneous epileptic seizures whereas the second group does not. Possible explanations for such an inconsistency could be a non-continuous epileptogenesis phenomena or the incapacity of multifractal metrics to captures changes in the brain that could lead to the occurrence of spontaneous epileptic seizures. However, multifractal metrics seem to provide information on epileptogenesis for some rats, so this study focussed on the former explanation.

The best way to evaluate the continuity of the epileptogenesis process is to split the recordings in day activities. The curves for variation of percentage of extremes per day suggest a transient condition to the epileptogenesis phase, rather than a con-

tinuous process that starts after the status epilepticus and converges to a state that allows the arise of spontaneous epileptic seizures. However, the evaluation of extremes of multifractal spectra measures in day intervals did not result in significant differences between S and NS.

The results suggest that MF metrics are not sensitive to epileptogenic activity and are not capable of identifying the group of rats that develops spontaneous seizures after an insult, as have previous non-linear metrics reported elsewhere (Rizzi et al., 2016, 2019). However, the same analysis has shown non-monotonic features, i.e. the percentages of extremes can increase osr decrease in subsequent days. Such an outcome elicits questions on the mechanisms of epileptogenesis itself, i.e. is the epileptogenesis a continuous process that starts after the brain insult and finishes after the first spontaneous seizures or slightly after the first spontaneous seizure? Does the epileptogenesis procedure follow any sort of oscillatory mechanism? It has been shown in the literature that epileptic seizures follow periods that span over days (Baud et al., 2018); how these rhythms are established from initial brain insult is a matter for future research.

It is still to be determined in future studies what are the types of oscillations that lead to increased spectral widths of multifractal spectra. Such oscillations might be easier to identify and may be more precise as future markers of epileptogenesis and become predictors of whether a patient is likely to develop epilepsy or not after neuronal damage.

This study is, however, limited by the choice on segments to be evaluated, i.e. the extremes statistics threshold. Activity related to epileptogenesis could also

occur in segments with multifractal spectra widths below the defined threshold. Such activity could even suppress other rhythms, reducing the heterogeneity of the signal and consequently the spectral width. These are issues to be investigated in future work.

Chapter 8

General Conclusions

This thesis discussed scaling properties in electrophysiological signals in epilepsy and implications of these characteristics on current knowledge of epileptogenesis and seizure phenomena in humans and rats.

The document was arranged in seven major chapters, starting with an introduction on relevant information and concepts for the studies presented subsequently. Firstly, a new approach to study epileptic seizures was reported. Secondly, this approach was applied to a series of studies ranging for seizure occurrence (characteristic time and classes) to epileptogenesis in animal models.

8.1 A novel methodology for the study of EEG dynamics

- *Scaling analysis methods, such as multifractal spectrum estimation approaches are affected by changes in the standard deviation of the time series, which can bias seizure analysis applications (Chapter 3).*

Chapter 3 has highlighted several challenges that need to be considered when analysing multifractal properties of EEG signals — namely choice of the appropriate estimation method, estimation parameters, and the influence of the time series variance on signal features.

This chapter suggested some solutions to these problems, such as the use of the Chhabra-Jensen approach combined with an epoch-wise standardisation approach, which has shown potential capabilities as a signal feature for machine learning applications.

The chapter also highlights possible process-specific challenges. In terms of epileptic seizures, future work is required to analyse a larger number of patients in order to draw firmer conclusions on the potential clinical relevance of multifractal analyses. Furthermore, the study of mechanistic generative models of EEG may shed light on why those multifractal changes occur. For example, a generative process of potential interest could feature a modified version of Bak-Tang-Wiesenfeld model (Bak et al. (1987)).

In this chapter, the monofractal and multifractal properties of human EEG recordings have been studied. It has been shown that monofractal estimates are influenced by the standard variation of the time series, thus not capturing features beyond signal variance. For multifractal estimation, it has been shown that the Chhabra-Jensen approach is the most stable, and a method of signal pre-processing to remove the influence caused by the variance of the signal has been developed.

Using the suggested approach, the multifractal estimates do not correlate with traditional EEG measures, thus yielding additional information about the signal and

being a relevant signal feature.

8.2 Insights on sleep

- *Sleep phases impact multifractal estimation in human long term electrographical recordings for epilepsy surgery planning (Chapter 4).*

Chapter 4 focuses on the impact of sleep stages on the metrics developed in chapter 3. This is an important aspect of analysis of intracranial EEG signals, as these are normally long-term recordings and are marked by changes in the sleep-wake state of the patient.

The results shown by the study suggest that multifractal measures and the metrics presented in chapter 3 are sensitive to variations featured by EEG signals during different sleep stages. Such an outcome has two major implications: 1. MF measures and the approach developed could be exploited in the automatic identification of sleep stages. 2. Changes in the wakefulness of a patient on long term monitoring, such as patients with intracranial implantations for epilepsy surgery planning or seizure detection/prediction devices, could affect metrics of evaluation of electrophysiological signals.

When it comes to sleep staging, the outcomes are still very speculative considering the results and the number of individuals with signals analysed. Moreover, the signals used in this study feature some limitations as these were scalp EEG recordings digitised at 100 Hz and two channels. Further research with higher sampling rate signals and more channels might show more robust results.

8.3 Seizures and characteristic time

- *Focal epileptic seizure detection exhibits optimal results at specific sampling frequencies (Chapter 5).*

An open question when it comes to seizure detection/prediction is the choice of parameters for the analysis of EEG signals, i.e. sampling rate and epoch size. Chapter 5 focussed on evaluating how these two parameters affect seizure detection in intracranial EEG from patients undergoing presurgical evaluation with multifractal metrics and the approach presented in chapter 3.

The results show that epileptic seizures do exhibit an optimal time scale, i.e. the combination of sampling rate and epoch length parameters for detection with the evaluated metrics. Such a result suggests that future approaches that aim looking at seizure phenomena, e.g. detection, prediction, characterisation (signal analysis), should look into these properties in order to optimise the outcomes.

Future work should evaluate the existence — and compare the magnitude — of the optimal time scales in different patients. Further research might also evaluate differences between optimal times in different seizures and whether different onsets in the same patients would feature distinct optimal times.

8.4 Seizure classes

- *Focal epileptic seizures can be classified into families using multifractal analysis (Chapter 6).*

To date, there is not an objective method to classify focal epileptic seizures. Classification of such episodes could change the understanding of epilepsy and pro-

vide new insights in potential new therapeutic approaches. Chapter 6 presents a putative approach to devise classes from a population of seizures with unsupervised machine learning techniques.

The proposed approach, when applied to a population of epileptic seizures undergoing planning for epilepsy surgery, suggests that seizures will cluster into groups. Overall, these clusters are not patient dependent — which suggests that there are common dynamics between seizures of different individuals. Moreover, clusters also do not seem to be connected to the surgical outcomes of patients. However, some clusters seem to concentrate longer seizures — that are in general secondarily generalised seizures.

The method is, however, limited by the principal component analysis performed in order to reduce dimensionality. Different patients have distinct implantation schemes, i.e. grids, strips, depth, causing difficulties in the comparison of the seizures.

A possible way to overcome that limitation consists of performing the procedure with the signal obtained from the onset zone, instead of applying principal component analysis to reduce the dimensionality of the problem. Nevertheless, there are also physical and pathological properties of the neural tissue that can make such comparison difficult.

Albeit despite these limitations, the proposed classification scheme offers a tool to class different seizures. The method developed in this study is a first step in the direction of a system of classification of focal epileptic seizures.

8.5 Extremes on epileptogenesis

- *Changes that occur during epileptogenesis can be detected by the multifractal analysis approach. An adequate application of multifractal analysis can identify animals that could develop spontaneous seizures after an initial trauma based on the analysis of local field potential signals. (Chapter 7)*

The mechanisms of epileptogenesis after a brain insult/neural damage are still not fully understood, which hinders the development of new treatments to prevent the evolution of the insult into epilepsy in patients. For example, in individuals after stroke or traumatic brain injury episodes.

Animal models can provide insights on such a process and provide tokens on processes that lead from the insult to the occurrence of epileptic seizures. Chapter 7 focussed on this matter and reports on a study on scaling properties and heterogeneity of electrographic signals obtained from rats after an induced status epilepticus.

The results of this study suggest that there are signs of epileptic activity on scaling properties and heterogeneity of an electrographic signals. The changes are, nonetheless, sporadic and difficult to detect. The changes observed with MF measures are not statistically significant.

Future work should look for strategies in finding such anomalous activity in electrophysiologic signals. Moreover, types of neural activity could follow temporal patterns which — if true — could provide further tokens on how to identify animals with the potential to develop spontaneous epileptic seizures in future.

8.6 Scaling properties in epilepsy

This thesis comprehends studies on scaling properties and heterogeneity in signals containing epileptiform activity in both human and animal data. It has been reported elsewhere that scaling properties are ubiquitous in the brain already and intrinsic to its function. This thesis discussed these concepts and definitions and explored such properties to study epileptiform activity.

The studies reported here were successful in applying multifractal analysis metrics and concepts to answer diverse questions and to analyse signals from different modalities — from identification/detection of ictal segments to finding signs of the development of epilepsy and spontaneous epileptic seizures after a brain insult.

The results obtained in such studies and shown in this thesis indicate that scaling properties are indeed tokens to understand, characterise, detect, and predict epileptiform activity. Hence, multifractal analysis-based metrics can be used in order to study epilepsy and epileptiform activity. The evaluation of the heterogeneity arising from different rhythms in the brain can be a unique biomarker potentially applicable to different analyses, e.g. seizure detection, seizure classification, and diagnosing potential brain injuries that could lead to epileptic seizures.

Future work should focus on these questions and increase the populations studied — to evaluate whether these outcomes hold with a bigger sample. Understanding how the non-classic self-organised criticality concepts, mentioned in chapter 3, would apply to epilepsy and seizure onset is also necessary in order to further develop the multifractal based metrics as biomarkers of epilepsy. Moreover, computational models should be able to help identify the mechanisms that generate

properties shown by multifractal analyses.

The studies in this thesis have taken a new approach to the analysis of EEG data in sleep/wake and epilepsy. It has been shown that in contrast to monofractal analysis, this form analysis can provide information beyond measures of variance. These studies have set out to show how this method can be applied and refined the approach in order to facilitate its use in multiple situations. Further work is required to determine the clinical utility in larger groups and to establish whether this form of analysis can predict clinical outcomes and inform our mechanistic understanding of changes in brain state.

Appendix A

Colophon

This document was set in the Times Roman typeface using LaTeX and BibTeX, composed with a text editor.

All the analysis and figures in this thesis were produced with MathWorks MATLAB, Wolfram Mathematica 12 and R programing language (R Core Team, 2019), as well as the following packages: ggplot2 (Wickham, 2016), RColorBrewer (Neuwirth, 2014), reshape2 (Wickham, 2007), cowplot (Wilke, 2019), cluster (Maechler et al., 2018), R.matlab (Bengtsson, 2018), tidyr (Wickham and Henry, 2019), igraph (Csardi and Nepusz, 2006), factoextra (Kassambara and Mundt, 2017), dendextend (Galili, 2015), dplyr (Wickham et al., 2019), dtw (Tormene et al., 2008), kableExtra (Zhu, 2019), ggnewscale (Campitelli, 2019), glue (Hester, 2019), ggpubr (Kassambara, 2019), lubridate (Grolemund and Wickham, 2011), magrittr (Bache and Wickham, 2014).

Appendix B

Publications associated with this thesis

França, L. G. S., Miranda, J. G. V., Leite, M., Sharma, N. K., Walker, M. C., Lemieux, L., and Wang, Y. (2018). Fractal and multifractal properties of electrographic recordings of human brain activity: toward its use as a signal feature for machine learning in clinical applications. *Frontiers in physiology*, 9.

França, L. G. S., Wang, Y., Gliske, S. V., Stacey, W. C., and Walker, M. C. Characteristic time of focal epileptic seizures in a population of patients. *(In preparation)*

França, L. G. S., Walker, M. C., Wang, Y. A method to classify focal epileptic seizures with dynamic time warping. *(In preparation)*

França, L. G. S., Shekh-Ahmad, T., Wang, Y., and Walker, M. C. Frequency of

occurrence of extreme states in the brain may suggest the development of spontaneous epileptic seizures after a neuronal insult. *(In preparation)*

Appendix C

Conference presentations related to this thesis

França, L. G. S., Walker, M. C., Miranda, J. G. V., Lemieux, L., and Wang, Y. (2017). A multifractal analysis framework for human intracranial electroencephalographic signals. International Conference for Technology and Analysis of Seizures (ICTALS). 20-23 August 2017. Minneapolis, MN, USA.

França, L. G. S., Walker, M. C., Miranda, J. G. V., Lemieux, L., and Wang, Y. (2017). Effect of signal standard deviation and sleep stages in EEG multifractal analysis. Society for Neuroscience (SfN) - Neuroscience 2017. 11-15 November 2017. Washington, DC, USA.

França, L. G. S., Walker, M. C., and Wang, Y. (2018). Do seizures have a characteristic time-scale for detection? Society for Neuroscience (SfN) - Neuroscience 2018. 3-7 November 2018. San Diego, CA, USA.

França, L. G. S., Walker, and Wang, Y. (2019). Can we classify epileptic seizures? International Conference for Technology and Analysis of Seizures (ICTALS). 2-5 September 2019. Exeter, UK.

Bibliography

Agarwal, R. and Gotman, J. (2001). Computer-assisted sleep staging. *IEEE Transactions on Biomedical Engineering*, 48(12):1412–1423.

Aksenova, T. I., Volkovych, V. V., and Villa, A. E. P. (2007). Detection of spectral instability in EEG recordings during the preictal period. *Journal of Neural Engineering*, 4(3):173–178.

American Academy of Sleep Medicine (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*.

Andrade, R., Schellnhuber, H., and Claussen, M. (1998). Analysis of rainfall records: possible relation to self-organized criticality. *Physica A: Statistical Mechanics and its Applications*, 254(3-4):557–568.

Argoul, F., Arnéodo, A., Grasseau, G., Gagne, Y., Hopfinger, E. J., and Frisch, U. (1989). Wavelet analysis of turbulence reveals the multifractal nature of the Richardson cascade. *Nature*, 338(6210):51–53.

Bache, S. M. and Wickham, H. (2014). *magrittr: A Forward-Pipe Operator for R*. R package version 1.5.

Bak, P. and Paczuski, M. (1995). Complexity, contingency, and criticality. *Proceedings of the National Academy of Sciences of the United States of America*, 92(15):6689–96.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/ f noise. *Physical Review Letters*, 59(4):381–384.

Baldassano, S. N., Brinkmann, B. H., Ung, H., Blevins, T., Conrad, E. C., Leyde, K., Cook, M. J., Khambhati, A. N., Wagenaar, J. B., Worrell, G. A., and Litt, B. (2017). Crowdsourcing seizure detection: algorithm development and validation on human implanted device recordings. *Brain*, 140(6):1680–1691.

Bandarabadi, M., Teixeira, C. a., Direito, B., and Dourado, A. (2012). Epileptic seizure prediction based on a bivariate spectral power methodology. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5943–5946. IEEE.

Baud, M. O., Kleen, J. K., Mirro, E. A., Andrechak, J. C., King-Stephens, D., Chang, E. F., and Rao, V. R. (2018). Multi-day rhythms modulate seizure risk in epilepsy. *Nature Communications*, 9(1):88.

Bedeeuzzaman, M., Fathima, T., Khan, Y. U., and Farooq, O. (2013). Seizure prediction using statistical dispersion measures of intracranial EEG. *Biomedical Signal Processing and Control*, 10:338–341.

Beggs, J. M. and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *The*

Journal of neuroscience : the official journal of the Society for Neuroscience, 23(35):11167–77.

Beggs, J. M. and Plenz, D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(22):5216–29.

Bengtsson, H. (2018). *R.matlab: Read and Write MAT Files and Call MATLAB from Within R*. R package version 3.6.2.

Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., van Emde Boas, W., Engel, J., French, J., Glauser, T. a., Mathern, G. W., Moshé, S. L., Nordli, D., Plouin, P., and Scheffer, I. E. (2010). Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia*, 51(4):676–685.

Berg, A. T. and Cross, J. H. (2010). Towards a modern classification of the epilepsies? *The Lancet Neurology*, 9(5):459–461.

Berndt, D. J. and Clifford, J. (1994). Using Dynamic Time Warping to Find Patterns in Time Series. In *AAAIWS'94 Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, pages 359–370, Seattle, WA. AAAI Press.

Bertram, E. H. and Cornett, J. (1993). The ontogeny of seizures in a rat model of

limbic epilepsy: evidence for a kindling process in the development of chronic spontaneous seizures. *Brain Research*, 625(2):295–300.

Bertram, E. H. and Cornett, J. F. (1994). The evolution of a rat model of chronic spontaneous limbic seizures. *Brain Research*, 661(1-2):157–162.

Bianco, S., Ignaccolo, M., Rider, M. S., Ross, M. J., Winsor, P., and Grigolini, P. (2007). Brain, music, and non-Poisson renewal processes. *Physical Review E*, 75(6):061911.

Birjandtalab, J., Pouyan, M. B., and Nourani, M. (2016). Nonlinear dimension reduction for EEG-based epileptic seizure detection. *3rd IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2016*, pages 595–598.

Biswas, A., Cresswell, H. P., and Bing, C. S. (2012). Application of Multifractal and Joint Multifractal Analysis in Examining Soil Spatial Variation: A Review. In Ouadfeul, S.-A., editor, *Fractal Analysis and Chaos in Geosciences*, pages 109–138. InTech.

Bleck, T. P. (1991). Convulsive disorders: Status epilepticus.

Blum, A. L. and Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1-2):245–271.

Brandt, C., Rankovic, V., Töllner, K., Klee, R., Bröer, S., and Löscher, W. (2016). Refinement of a model of acquired epilepsy for identification and validation of biomarkers of epileptogenesis in rats. *Epilepsy and Behavior*, 61:120–131.

Brinkmann, B. H., Bower, M. R., Stengel, K. A., Worrell, G. A., and Stead, M. (2009). Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data. *Journal of Neuroscience Methods*, 180(1):185–192.

Brinkmann, B. H., Wagenaar, J., Abbot, D., Adkins, P., Bosshard, S. C., Chen, M., Tieng, Q. M., He, J., Muñoz-Almaraz, F. J., Botella-Rocamora, P., Pardo, J., Zamora-Martinez, F., Hills, M., Wu, W., Korshunova, I., Cukierski, W., Vite, C., Patterson, E. E., Litt, B., and Worrell, G. A. (2016). Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. *Brain*, 139(6):1713–1722.

Brodie, M. (1990). Status epilepticus in adults. *The Lancet*, 336(8714):551–552.

Camastra, F. and Vinciarelli, A. (2015). Clustering Methods. In *Advanced Information and Knowledge Processing*, number 9781447167341, pages 131–167.

Campitelli, E. (2019). *ggnewscale: Multiple Fill and Color Scales in 'ggplot2'*. R package version 0.3.0.

Cannon, M. J., Percival, D. B., Caccia, D. C., Raymond, G. M., and Bassingthwaighte, J. B. (1997). Evaluating scaled windowed variance methods for estimating the Hurst coefficient of time series. *Physica A: Statistical Mechanics and its Applications*, 241(3-4):606–626.

Carney, P. R., Myers, S., and Geyer, J. D. (2011). Seizure prediction: Methods. *Epilepsy & Behavior*, 22:S94–S101.

Celesia, G. G. (1976). Modern Concepts of Status Epilepticus. *JAMA: The Journal of the American Medical Association*, 235(15):1571–1574.

Chang, N.-f., Chen, T.-C., Chiang, C.-y., and Chen, L.-g. (2012). Channel selection for epilepsy seizure prediction method based on machine learning. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2012:5162–5.

Chaudhary, U. J., Carmichael, D. W., Rodionov, R., Thornton, R. C., Bartlett, P., Vulliemoz, S., Micallef, C., McEvoy, A. W., Diehl, B., Walker, M. C., Duncan, J. S., and Lemieux, L. (2012). Mapping preictal and ictal haemodynamic networks using video- electroencephalography and functional imaging. *Brain*, 135(12):3645–3663.

Cheng-Yi Chiang, Nai-Fu Chang, Tung-Chien Chen, Hong-Hui Chen, and Liang-Gee Chen (2011). Seizure prediction based on classification of EEG synchronization patterns with on-line retraining and post-processing scheme. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 7564–7569. IEEE.

Chhabra, A. and Jensen, R. V. (1989). Direct determination of the f(α) singularity spectrum. *Physical Review Letters*, 62(12):1327–1330.

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nature Physics*, 6(10):744–750.

Chialvo, D. R. (2012). Critical brain dynamics at large scale. *Criticality in Neural Systems*, 2013:1–22.

Ciuciu, P. (2012). Scale-free and multifractal time dynamics of fMRI signals during rest and task. *Frontiers in Physiology*, 3(June):1–18.

Commission on Classification and Terminology of the International League Against Epilepsy (1989). Proposal for Revised Classification of Epilepsies and Epileptic Syndromes. *Epilepsia*, 30(4):389–399.

Consolini, G., Marcucci, M. F., and Candidi, M. (1996). Multifractal structure of auroral electrojet index data. *Physical Review Letters*, 76(21):4082–4085.

Cook, M. J., Karoly, P. J., Freestone, D. R., Himes, D., Leyde, K., Berkovic, S., Brien, T. O., Grayden, D. B., and Boston, R. (2016). Human focal seizures are characterized by populations of fixed duration and interval. pages 359–368.

Cook, M. J., O'Brien, T. J., Berkovic, S. F., Murphy, M., Morokoff, A., Fabinyi, G., D'Souza, W., Yerra, R., Archer, J., Litewka, L., Hosking, S., Lightfoot, P., Ruedebusch, V., Sheffield, W. D., Snyder, D., Leyde, K., and Himes, D. (2013). Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: A first-in-man study. *The Lancet Neurology*, 12(6):563–571.

Cook, M. J., Varsavsky, A., Himes, D., Leyde, K., Berkovic, S., O'Brien, T., and Mareels, I. (2014). The dynamics of the epileptic brain reveal long memory processes. *Frontiers in Neurology*, 5(OCT):1–8.

Costa, I. d. S., Gamundí, A., Miranda, J. G. V., França, L. G. S., De Santana,

C. N., and Montoya, P. (2017). Altered Functional Performance in Patients with Fibromyalgia. *Frontiers in Human Neuroscience*, 11(January):14.

Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, Complex Systems:1695.

Davies, R. B. and Harte, D. S. (1987). Tests for Hurst effect. *Biometrika*, 74(1):95–101.

Davis, A., Marshak, A., Cahalan, R., and Wiscombe, W. (1997). The Landsat Scale Break in Stratocumulus as a Three-Dimensional Radiative Transfer Effect: Implications for Cloud Remote Sensing. *Journal of the Atmospheric Sciences*, 54(2):241–260.

De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., Visentin, D., van den Driessche, G., Lakshminarayanan, B., Meyer, C., Mackinder, F., Bouton, S., Ayoub, K., Chopra, R., King, D., Karthikesalingam, A., Hughes, C. O., Raine, R., Hughes, J., Sim, D. A., Egan, C., Tufail, A., Montgomery, H., Hassabis, D., Rees, G., Back, T., Khaw, P. T., Suleyman, M., Cornebise, J., Keane, P. A., and Ronneberger, O. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9):1342–1350.

de Tisi, J., Bell, G. S., Peacock, J. L., McEvoy, A. W., Harkness, W. F., Sander, J. W., and Duncan, J. S. (2011). The long-term outcome of adult epilepsy

surgery, patterns of seizure remission, and relapse: a cohort study. *The Lancet*, 378(9800):1388–1395.

Dokmanic, I., Parhizkar, R., Ranieri, J., and Vetterli, M. (2015). Euclidean Distance Matrices: Essential theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 32(6):12–30.

Drury, I., Smith, B., Li, D., and Savit, R. (2003). Seizure prediction using scalp electroencephalogram. *Experimental Neurology*, 184(SUPPL. 1).

Eftekhar, A., Juffali, W., El-Imad, J., Constandinou, T. G., and Toumazou, C. (2014). Ngram-Derived Pattern Recognition for the Detection and Prediction of Epileptic Seizures. *PLoS ONE*, 9(6):e96235.

Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., and Apkarian, A. V. (2005). Scale-Free Brain Functional Networks. *Physical Review Letters*, 94(1):018102.

Eke, A., Herman, P., Kocsis, L., and Kozak, L. R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement*, 23(1):R1–R38.

Elger, C. E. (2001). Future trends in epileptology. *Current Opinion in Neurology*, 14(2):185–186.

Esteller, R., Vachtsevanos, G., Echauz, J., Henry, T., Pennell, P., Epstein, C., Bakay, R., Bowen, C., and Litt, B. (1999). Fractal dimension characterizes seizure onset in epileptic patients. In *1999 IEEE International Conference on Acoustics,*

Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258), volume 4, pages 2343–2346 vol.4. IEEE.

Esteller, R., Vachtsevanos, G., Echauz, J., and Lilt, B. (1995). A comparison of fractal dimension algorithms using synthetic and experimental data. In Andresen, B. F. and Strojnik, M., editors, *ISCAS'99. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems VLSI (Cat. No.99CH36349)*, volume 3, pages 199–202. IEEE.

Estrada, E., Nazeran, H., Nava, P., Behbehani, K., Burk, J., and Lucas, E. (2005). EEG feature extraction for classification of sleep stages. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 3, pages 196–199. IEEE.

Falconer, K. (2003). *Fractal Geometry*. John Wiley & Sons, Ltd, Chichester, UK.

Feder, J. (1988). *Fractals*. Springer US, Boston, MA.

Feldwisch-Drentrup, H., Staniek, M., Schulze-Bonhage, A., Timmer, J., Dickten, H., Elger, C. E., Schelter, B., and Lehnertz, K. (2011). Identification of Preseizure States in Epilepsy: A Data-Driven Approach for Multichannel EEG Recordings. *Frontiers in Computational Neuroscience*, 5(July):32.

Fisher, R. S., Boas, W. v. E., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J. (2005). Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–472.

Flexer, A., Gruber, G., and Dorffner, G. (2005). A reliable probabilistic sleep stager based on a single EEG signal. *Artificial Intelligence in Medicine*, 33(3):199–207.

Fraiman, D. and Chialvo, D. R. (2012). What kind of noise is brain noise: anomalous scaling behavior of the resting brain activity fluctuations. *Frontiers in Physiology*, 3(July):1–11.

França, L. G. S., Montoya, P., and Miranda, J. G. V. (2019). On multifractals: A non-linear study of actigraphy data. *Physica A: Statistical Mechanics and its Applications*, 514:612–619.

Freestone, D. R., Karoly, P. J., Peterson, A. D. H., Kuhlmann, L., Lai, A., Goodarzy, F., and Cook, M. J. (2015). Seizure Prediction: Science Fiction or Soon to Become Reality? *Current Neurology and Neuroscience Reports*, 15(11):73.

French, J. A., Williamson, P. D., Thadani, V. M., Darcey, T. M., Mattson, R. H., Spencer, S. S., and Spencer, D. D. (1993). Characteristics of medial temporal lobe epilepsy: I. Results of history and physical examination. *Annals of Neurology*, 34(6):774–780.

Gadhoumi, K., Gotman, J., and Lina, J. M. (2015). Scale Invariance Properties of Intracerebral EEG Improve Seizure Prediction in Mesial Temporal Lobe Epilepsy. *PLoS ONE*, 10(4):e0121182.

Gadhoumi, K., Lina, J. M., and Gotman, J. (2013). Seizure prediction in patients with mesial temporal lobe epilepsy using EEG measures of state similarity. *Clinical Neurophysiology*, 124(9):1745–1754.

Galili, T. (2015). dendextend: an r package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*.

Gastaut, H. (1970). Clinical and Electroencephalographical Classification of Epileptic Seizures. *Epilepsia*, 11(1):102–112.

Gentle, J. E. (2007). *Matrix Algebra*. Springer Texts in Statistics. Springer New York, New York, NY.

Gneiting, T. and Schlather, M. (2004). Stochastic Models That Separate Fractal Dimension and the Hurst Effect. *SIAM Review*, 46(2):269–282.

Goldberg, E. M. and Coulter, D. A. (2013). Mechanisms of epileptogenesis: a convergence on neural circuit dysfunction. *Nature Reviews Neuroscience*, 14(5):337–349.

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23):e215–e220.

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., and Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences*, 99(Supplement 1):2466–2472.

Gómez, C., Mediavilla, Á., Hornero, R., Abásolo, D., and Fernández, A. (2009).

Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients. *Medical Engineering & Physics*, 31(3):306–313.

Gong, P., Nikolaev, A. R., and van Leeuwen, C. (2003). Scale-invariant fluctuations of the dynamical synchronization in human brain electrical activity. *Neuroscience Letters*, 336(1):33–36.

Greenberg, D. A., Aminoff, M. J., and Simon, R. P. (2012). Seizures & Syncope. In *Clinical Neurology*, chapter 12, pages 354–378. Lange Medical Books/McGraw-Hill, 8th editio edition.

Grinenko, O., Li, J., Mosher, J. C., Wang, I. Z., Bulacio, J. C., Gonzalez-Martinez, J., Nair, D., Najm, I., Leahy, R. M., and Chauvel, P. (2018). A fingerprint of the epileptogenic zone in human epilepsies. *Brain*, 141(1):117–131.

Grolemund, G. and Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3):1–25.

Gu, G.-F. and Zhou, W.-X. (2010). Detrending moving average algorithm for multifractals. *Physical Review E*, 82(1):011136.

Güne, S., Polat, K., and Yosunkaya, . (2010). Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. *Expert Systems with Applications*, 37(12):7922–7928.

Guyon, I. and Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)*, 3(3):1157–1182.

Haddad, T., Talbi, L., Lakhssassi, A., Naim, B.-H., and Aouini, S. (2014). Epilepsy seizure prediction using graph theory. In *2014 IEEE 12th International New Circuits and Systems Conference (NEWCAS)*, pages 293–296. IEEE.

Harrison, M. A. F., Frei, M. G., and Osorio, I. (2005a). Accumulated energy revisited. *Clinical Neurophysiology*, 116(3):527–531.

Harrison, M. A. F., Osorio, I., Frei, M. G., Asuri, S., and Lai, Y.-C. (2005b). Correlation dimension and integral do not predict epileptic seizures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(3):033106.

Hellier, J. L., Patrylo, P. R., Buckmaster, P. S., and Dudek, F. (1998). Recurrent spontaneous motor seizures after repeated low-dose systemic treatment with kainate: assessment of a rat model of temporal lobe epilepsy. *Epilepsy Research*, 31(1):73–84.

Hester, J. (2019). *glue: Interpreted String Literals*. R package version 1.3.1.

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, 31(2):277–283.

Hinton, G. and Roweis, S. (2002). Stochastic neighbor embedding. *Advances in Neural Information Processing Systems 15*, pages 833–840.

Hu, K., Ivanov, P. C., Chen, Z., Hilton, M. F., Stanley, H., and Shea, S. A. (2004). Non-random fluctuations and multi-scale dynamics regulation of human activity. *Physica A: Statistical Mechanics and its Applications*, 337(1-2):307–318.

Hu, K., Van Someren, E. J. W., Shea, S. A., and Scheer, F. A. J. L. (2009). Reduction of scale invariance of activity fluctuations with aging and Alzheimer's disease: Involvement of the circadian pacemaker. *Proceedings of the National Academy of Sciences*, 106(8):2490–2494.

Hyunchul Kim and Rosen, J. (2010). Epileptic seizure detection - an AR model based algorithm for implantable device. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, volume 2010, pages 5541–5544. IEEE.

Iasemidis, L. (2003). Epileptic seizure prediction and control. *IEEE Transactions on Biomedical Engineering*, 50(5):549–558.

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

Ihlen, E. A. F. (2012). Introduction to Multifractal Detrended Fluctuation Analysis in Matlab. *Frontiers in Physiology*, 3(June):1–18.

Ihlen, E. A. F. and Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond $1/\alpha$ fluctuation. *Journal of Experimental Psychology: General*, 139(3):436–463.

Ince, R. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., and Schyns, P. G. (2017). A statistical framework for neuroimaging data analysis based on

mutual information estimated via a gaussian copula. *Human Brain Mapping*, 38(3):1541–1573.

Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R., and Stanley, H. E. (1999). Multifractality in human heartbeat dynamics. *Nature*, 399(6735):461–465.

James, C. and Gupta, D. (2009). Seizure prediction for epilepsy using a multi-stage phase synchrony based system. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 44, pages 25–28. IEEE.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. Springer New York, New York, NY.

Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer-Verlag, New York, 2nd edition edition.

Jupp, B., Williams, J., Binns, D., Hicks, R. J., Cardamone, L., Jones, N., Rees, S., and O'Brien, T. J. (2012). Hypometabolism precedes limbic atrophy and spontaneous recurrent seizures in a rat model of TLE. *Epilepsia*, 53(7):1233–1244.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. (2002). Multifractal detrended fluctuation analysis of nonstation-

ary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1-4):87–114.

Karoly, P. J., Ung, H., Grayden, D. B., Kuhlmann, L., Leyde, K., Cook, M. J., and Freestone, D. R. (2017). The circadian profile of epilepsy improves seizure forecasting. *Brain*, 140(8):2169–2182.

Kassambara, A. (2019). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.2.1.

Kassambara, A. and Mundt, F. (2017). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R package version 1.0.5.

Kemp, B., Zwinderman, A., Tuk, B., Kamphuisen, H., and Oberye, J. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9):1185–1194.

Kestener, P. and Arneodo, A. (2003). Three-Dimensional Wavelet-Based Multifractal Method: The Need for Revisiting the Multifractal Description of Turbulence Dissipation Data. *Physical Review Letters*, 91(19):194501.

Koley, B. and Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel EEG signal. *Computers in Biology and Medicine*, 42(12):1186–1195.

Kroese, D. P. and Botev, Z. I. (2015). Spatial Process Simulation. In Schmidt, V., editor, *Lecture Notes in Mathematics*, volume 2120 of *Lecture Notes in Mathematics*, pages 369–404. Springer International Publishing, Cham.

Kuhlmann, L., Karoly, P., Freestone, D. R., Brinkmann, B. H., Temko, A., Barachant, A., Li, F., Titericz, G., Lang, B. W., Lavery, D., Roman, K., Broadhead, D., Dobson, S., Jones, G., Tang, Q., Ivanenko, I., Panichev, O., Proix, T., Náhlík, M., Grunberg, D. B., Reuben, C., Worrell, G., Litt, B., Liley, D. T. J., Grayden, D. B., and Cook, M. J. (2018a). Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain*, 141(9):2619–2630.

Kuhlmann, L., Lehnertz, K., Richardson, M. P., Schelter, B., and Zaveri, H. P. (2018b). Seizure prediction ready for a new era. *Nature Reviews Neurology*, page 1.

Kwan, P., Arzimanoglou, A., Berg, A. T., Brodie, M. J., Allen Hauser, W., Mathern, G., Moshé, S. L., Perucca, E., Wiebe, S., and French, J. (2010). Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia*, 51(6):1069–77.

Lagerlund, T. D., Cascino, G. D., Cicora, K. M., and Sharbrough, F. W. (1996). Long-Term Electroencephalographic Monitoring for Diagnosis and Management of Seizures. *Mayo Clinic Proceedings*, 71(10):1000–1006.

Lajnef, T., Chaibi, S., Ruby, P., Aguera, P.-E., Eichenlaub, J.-B., Samet, M., Kachouri, A., and Jerbi, K. (2015). Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. *Journal of Neuroscience Methods*, 250:94–105.

Le Van Quyen, M., Soss, J., Navarro, V., Robertson, R., Chavez, M., Baulac, M., and Martinerie, J. (2005). Preictal state identification by synchronization changes in long-term intracranial EEG recordings. *Clinical Neurophysiology*, 116(3):559–568.

Lee, J. M., Kim, D. J., Kim, I. Y., Suk Park, K., and Kim, S. I. (2004). Nonlinear-analysis of human sleep EEG using detrended fluctuation analysis. *Medical Engineering and Physics*, 26(9 SPEC.ISS.):773–776.

Li, D., Zhou, W., Drury, I., and Savit, R. (2006). Seizure anticipation, states of consciousness and marginal predictability in temporal lobe epilepsy. *Epilepsy Research*, 68(1):9–18.

Li, X., Polygiannakis, J., Kapiris, P., Peratzakis, A., Eftaxias, K., and Yao, X. (2005). Fractal spectral analysis of pre-epileptic seizures in terms of criticality. *Journal of Neural Engineering*, 2(2):11–16.

Liang, S.-F., Kuo, C.-E., Hu, Y.-H., Pan, Y.-H., and Wang, Y.-H. (2012). Automatic Stage Scoring of Single-Channel Sleep EEG by Using Multiscale Entropy and Autoregressive Models. *IEEE Transactions on Instrumentation and Measurement*, 61(6):1649–1657.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(4):1370–7.

Lipa, P. and Buschbeck, B. (1989). From strong to weak intermittency. *Physics Letters B*, 223(3-4):465–469.

Liu, H., Motoda, H., and Dash, M. (1998). A monotonic measure for optimal feature selection. In *Ecml-98*, pages 101–106.

Liu, H. and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17(4):491–502.

Liu, S. J., Zheng, P., Wright, D. K., Dezsi, G., Braine, E., Nguyen, T., Corcoran, N. M., Johnston, L. A., Hovens, C. M., Mayo, J. N., Hudson, M., Shultz, S. R., Jones, N. C., and O'Brien, T. J. (2016). Sodium selenate retards epileptogenesis in acquired epilepsy models reversing changes in protein phosphatase 2A and hyperphosphorylated tau. *Brain*, 139(7):1919–1938.

Lovejoy, S. and Schertzer, D. (2007). Scaling and multifractal fields in the solid earth and topography. *Nonlinear Processes in Geophysics*, 14(4):465–502.

Lowenstein, D. H. and Alldredge, B. K. (1998). Status Epilepticus. *New England Journal of Medicine*, 338(14):970–976.

Lowenstein, D. H., Bleck, T., and Macdonald, R. L. (1999). It's Time to Revise the Definition of Status Epilepticus. *Epilepsia*, 40(1):120–122.

Ma, Q., Ning, X., Wang, J., and Li, J. (2005). Sleep-stage Characterization by Nonlinear EEG Analysis using Wavelet-based Multifractal Formalism. In *2005 IEEE*

Engineering in Medicine and Biology 27th Annual Conference, pages 4526–4529. IEEE.

Ma, Y., Shi, W., Peng, C. K., and Yang, A. C. (2018). Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. *Sleep Medicine Reviews*, 37:85–93.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2018). *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.7-1 — For new features, see the 'Changelog' file (in the package source).

Mandelbrot, B. B. (1983). *The fractal geometry of nature*. W.H.Freeman & Co Ltd.

Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, 10(4):422–437.

Mandelbrot, B. B. and Wallis, J. R. (1968). Noah, Joseph, and Operational Hydrology. *Water Resources Research*, 4(5):909–918.

Martinerie, J., Adam, C., Quyen, M. L. V., Baulac, M., Clemenceau, S., Renault, B., and Varela, F. (1998). Epileptic seizures can be anticipated by non-linear analysis. *Nature Medicine*, 4(10):1173–1176.

Mazarati, A., Bragin, A., Baldwin, R., Shin, D., Wilson, C., Sankar, R., Naylor, D., Engel, J., and Wasterlain, C. G. (2002). Epileptogenesis After Self-Sustaining Status Epilepticus. *Epilepsia*, 43(SUPPL. 5):74–80.

McDonnell, L. A., Balluff, B., Dijkstra, J., Reinders, M. J. T., Abdelmoula, W. M., Englert, S., Lelieveldt, B. P. F., and Walch, A. (2016). Data-driven identifica-

tion of prognostic tumor subpopulations using spatially mapped t-SNE of mass spectrometry imaging data. *Proceedings of the National Academy of Sciences*, 113(43):12244–12249.

McSharry, P. E., Smith, L. a., and Tarassenko, L. (2003). Prediction of epileptic seizures: are nonlinear methods relevant? *Nature medicine*, 9(3):241–242; author reply 242.

Meinardi, H., Scott, R. A., Reis, R., and On Behalf Of The Ilae Commission on, J. W. A. S. S. (2008). The Treatment Gap in Epilepsy: The Current Situation and Ways Forward. *Epilepsia*, 42(1):136–149.

Mendelson, W. (1987). *Human sleep: research and clinical care*. Springer US.

Meneveau, C. and Sreenivasan, K. R. (1987). Simple multifractal cascade model for fully developed turbulence. *Physical Review Letters*, 59(13):1424–1427.

Meneveau, C. and Sreenivasan, K. R. (1991). The multifractal nature of turbulent energy dissipation. *Journal of Fluid Mechanics*, 224(-1):429.

Merlis, J. K. (1970). Proposal for an International Classification of the Epilepsies. *Epilepsia*, 11(1):114–119.

Miranda, J., Montero, E., Alves, M., Paz González, A., and Vidal Vázquez, E. (2006). Multifractal characterization of saprolite particle-size distributions after topsoil removal. *Geoderma*, 134(3-4):373–385.

Mormann, F., Andrzejak, R. G., Elger, C. E., and Lehnertz, K. (2007). Seizure prediction: the long and winding road. *Brain*, 130(2):314–333.

Mormann, F., Andrzejak, R. G., Kreuz, T., Rieke, C., David, P., Elger, C. E., and Lehnertz, K. (2003). Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Physical Review E*, 67(2):021912.

Mourtazaev, M. S., Kemp, B., Zwinderman, A. H., and Kamphuisen, H. A. C. (1995). Age and Gender Affect Different Characteristics of Slow Waves in the Sleep EEG. *Sleep*, 18(7):557–564.

Mukli, P., Nagy, Z., and Eke, A. (2015). Multifractal formalism by enforcing the universal behavior of scaling functions. *Physica A: Statistical Mechanics and its Applications*, 417:150–167.

Murcio, R., Masucci, A. P., Arcaute, E., and Batty, M. (2015). Multifractal to monofractal evolution of the London street network. *Physical Review E*, 92(6):062130.

Nagy, Z., Mukli, P., Herman, P., and Eke, A. (2017). Decomposing Multifractal Crossovers. *Frontiers in Physiology*, 8(JUL).

Naro, D., Rummel, C., Schindler, K., and Andrzejak, R. G. (2014). Detecting determinism with improved sensitivity in time series: Rank-based nonlinear predictability score. *Physical Review E*, 90(3):032913.

Navarro, V., Martinerie, J., Le Van Quyen, M., Baulac, M., Dubeau, F., and Gotman, J. (2005). Seizure Anticipation: Do Mathematical Measures Correlate with Video-EEG Evaluation? *Epilepsia*, 46(3):385–396.

Navarro, V., Martinerie, J., Le Van Quyen, M., Clemenceau, S., Adam, C., Baulac, M., and Varela, F. (2002). Seizure anticipation in human neocortical partial epilepsy. *Brain : a journal of neurology*, 125(Pt 3):640–655.

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer Palettes*. R package version 1.1-2.

Ngugi, A. K., Bottomley, C., Kleinschmidt, I., Sander, J. W., and Newton, C. R. (2010). Estimation of the burden of active and life-time epilepsy: A meta-analytic approach. *Epilepsia*, 51(5):883–890.

Niknazar, H. and Nasrabadi, A. M. (2015). Epileptic seizure prediction using angle method. In *2015 23rd Iranian Conference on Electrical Engineering*, pages 56–60. IEEE.

Noachtar, S., Binnie, C., Ebersole, J., Mauguière, F., Sakamoto, A., and Westmoreland, B. (1999). A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the EEG findings. The International Federation of Clinical Neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement*, 52:21–41.

Päivinen, N., Lammi, S., Pitkänen, A., Nissinen, J., Penttonen, M., and Grönfors, T. (2005). Epileptic seizure detection: A nonlinear viewpoint. *Computer Methods and Programs in Biomedicine*, 79(2):151–159.

Papo, D. (2014). Functional significance of complex fluctuations in brain activity:

from resting state to cognitive neuroscience. *Frontiers in Systems Neuroscience*, 8(June):112.

Papo, D., Goñi, J., and Buldú, J. M. (2017). Editorial: On the relation of dynamics and structure in brain networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(4):047201.

Pauletti, A., Terrone, G., Shekh-Ahmad, T., Salamone, A., Ravizza, T., Rizzi, M., Pastore, A., Pascente, R., Liang, L. P., Villa, B. R., Balosso, S., Abramov, A. Y., Van Vliet, E. A., Del Giudice, E., Aronica, E., Antoine, D. J., Patel, M., Walker, M. C., and Vezzani, A. (2017). Targeting oxidative stress improves disease outcomes in a rat model of acquired epilepsy. *Brain*, 140(7):1885–1899.

Paz Ferreiro, J., Miranda, J., and Vidal Vázquez, E. (2010). Multifractal Analysis of Soil Porosity Based on Mercury Injection and Nitrogen Adsorption. *Vadose Zone Journal*, 9(2):325.

Paz-Ferreiro, J., Vázquez, E. V., and Miranda, J. G. (2010). Assessing soil particle-size distribution on experimental plots with similar texture under different management systems using multifractal parameters. *Geoderma*, 160(1):47–56.

Pechlivanidis, I. G. and Arheimer, B. (2015). Large-scale hydrological modelling by using modified PUB recommendations: the India-HYPE case. *Hydrology and Earth System Sciences*, 19(11):4559–4579.

Peng, C., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger,

A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2):1685–1689.

Peng, C., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87.

Penry, K. (1981). Proposal for Revised Clinical and Electroencephalographic Classification of Epileptic Seizures. *Epilepsia*, 22(4):489–501.

Pereda, E., Gamundi, A., Rial, R., and González, J. (1998). Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience Letters*, 250(2):91–94.

Petitjean, F., Ketterlin, A., and Gançarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3):678–693.

Pitkänen, A. and Engel, J. (2014). Past and Present Definitions of Epileptogenesis and Its Biomarkers. *Neurotherapeutics*, 11(2):231–241.

Pitkänen, A., Kharatishvili, I., Karhunen, H., Lukasiuk, K., Immonen, R., Nairismägi, J., Gröhn, O., and Nissinen, J. (2007). Epileptogenesis in Experimental Models. *Epilepsia*, 48(s2):13–20.

Pitkänen, A., Löscher, W., Vezzani, A., Becker, A. J., Simonato, M., Lukasiuk, K., Gröhn, O., Bankstahl, J. P., Friedman, A., Aronica, E., Gorter, J. A., Ravizza, T.,

Sisodiya, S. M., Kokaia, M., and Beck, H. (2016). Advances in the development of biomarkers for epilepsy. *The Lancet Neurology*, 15(8):843–856.

Pitkänen, A. and Lukasiuk, K. (2011). Mechanisms of epileptogenesis and potential treatment targets. *The Lancet Neurology*, 10(2):173–186.

Platzer, A. (2013). Visualization of SNPs with t-SNE. *PLoS ONE*, 8(2):e56883.

Polychronaki, G. E., Ktonas, P. Y., Gatzonis, S., Siatouni, A., Asvestas, P. A., Tsekou, H., Sakas, D., and Nikita, K. S. (2010). Comparison of fractal dimension estimation algorithms for epileptic seizure onset detection. *Journal of Neural Engineering*, 7(4):046007.

Protopopescu, V. a., Hively, L. M., and Gailey, P. C. (2001). Epileptic Event Forewarning From Scalp EEG. *Journal of Clinical Neurophysiology*, 18(3):223–245.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Racz, F. S., Mukli, P., Nagy, Z., and Eke, A. (2018). Multifractal dynamics of resting-state functional connectivity in the prefrontal cortex. *Physiological Measurement*, 39(2):024003.

Radha, M., Garcia-Molina, G., Poel, M., and Tononi, G. (2014). Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1876–1880. IEEE.

Rizzi, M., Brandt, C., Weissberg, I., Milikovsky, D. Z., Pauletti, A., Terrone, G., Salamone, A., Frigerio, F., Löscher, W., Friedman, A., and Vezzani, A. (2019). Changes of dimension of EEG/ECoG nonlinear dynamics predict epileptogenesis and therapy outcomes. *Neurobiology of Disease*, 124(December 2018):373–378.

Rizzi, M., Weissberg, I., Milikovsky, D. Z., and Friedman, A. (2016). Following a potential epileptogenic insult, prolonged high rates of nonlinear dynamical regimes of intermittency type is the hallmark of epileptogenesis. *Scientific Reports*, 6(July):1–12.

Rogawski, M. A. and Cavazos, J. E. (2015). Mechanisms of action of antiepileptic drugs. In Wyllie, E., Cascino, G. D., Gidal, B. E., Goodkin, H. P., Loddenkemper, T., and Sirven, J., editors, *Wyllie's Treatment of Epilepsy: Principles and Practice*, page 1128. Lippincott Williams & Wilkins, 6th edition.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C):53–65.

Saeys, Y., Inza, I., and Larranaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517.

Sakoe, H. and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49.

Sander, J. W. (2003). The epidemiology of epilepsy revisited. *Current opinion in neurology*, 16(2):165–70.

Schroeder, G. M., Diehl, B., Chowdhury, F. A., Duncan, J. S., and Tisi, J. D. (2019). Slow changes in seizure pathways in individual patients with focal epilepsy. pages 1–37.

Sharafat, A. R. and Nesaei, S. (2015). Real-time mining of epileptic seizure precursors via nonlinear mapping and dissimilarity features. *IET Signal Processing*, 9(3):193–200.

She, Z.-S. and Leveque, E. (1994). Universal scaling laws in fully developed turbulence. *Physical Review Letters*, 72(3):336–339.

Shekh-Ahmad, T., Eckel, R., Dayalan Naidu, S., Higgins, M., Yamamoto, M., Dinkova-Kostova, A. T., Kovac, S., Abramov, A. Y., and Walker, M. C. (2018). KEAP1 inhibition is neuroprotective and suppresses the development of epilepsy. *Brain*, 141(5):1390–1403.

Shekh-Ahmad, T., Kovac, S., Abramov, A., and Walker, M. (2019a). Reactive oxygen species in status epilepticus. *Epilepsy & Behavior*, (xxxx):106410.

Shekh-Ahmad, T., Lieb, A., Kovac, S., Gola, L., Wigley, W. C., Abramov, A. Y., and Walker, M. C. (2019b). Combination antioxidant therapy prevents epileptogenesis and modifies chronic epilepsy. *Redox Biology*, 26(July):101278.

Shevchenko, G. (2015). Fractional Brownian motion in a nutshell. *International Journal of Modern Physics: Conference Series*, 36(2014):1560002.

Shishikura, M. (1998). The Hausdorff Dimension of the Boundary of the Mandelbrot Set and Julia Sets. *The Annals of Mathematics*, 147(2):225.

Shufang Li, Weidong Zhou, Qi Yuan, and Yinxia Liu (2013). Seizure Prediction Using Spike Rate of Intracranial EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(6):880–886.

Sinha, N., Dauwels, J., Kaiser, M., Cash, S. S., Brandon Westover, M., Wang, Y., and Taylor, P. N. (2017). Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling. *Brain*, 140(2):319–332.

Smith, J. R. and Karacan, I. (1971). EEG sleep stage scoring by an automatic hybrid system. *Electroencephalography and Clinical Neurophysiology*, 31(3):231–237.

Snyder, D. E., Echauz, J., Grimes, D. B., and Litt, B. (2008). The statistics of a practical seizure warning system. *Journal of Neural Engineering*, 5(4):392–401.

Sornette, A. and Sornette, D. (1989). Self-Organized Criticality and earthquakes. *EPL (Europhysics Letters)*, 9(3):197.

Spasic, S., Kalauzi, A., Grbic, G., Martac, L., and Culic, M. (2005). Fractal analysis of rat brain activity after injury. *Medical & Biological Engineering & Computing*, 43(3):345–348.

Sreenivasan, K. R. and Antonia, R. A. (1997). The phenomenology of small-scale turbulence. *Annual Review of Fluid Mechanics*, 29(1):435–472.

Stanley, H., Amaral, L., Goldberger, A., Havlin, S., Ivanov, P., and Peng, C.-K. (1999). Statistical physics and physiology: Monofractal and multifractal ap-

proaches. *Physica A: Statistical Mechanics and its Applications*, 270(1-2):309–324.

Suckling, J., Wink, A. M., Bernard, F. A., Barnes, A., and Bullmore, E. (2008). Endogenous multifractal brain dynamics are modulated by age, cholinergic blockade and cognitive performance. *Journal of Neuroscience Methods*, 174(2):292–300.

Takahashi, H., Takahashi, S., Kanzaki, R., and Kawai, K. (2012). State-dependent precursors of seizures in correlation-based functional networks of electrocorticograms of patients with temporal lobe epilepsy. *Neurological Sciences*, 33(6):1355–1364.

Taskesen, E. and Reinders, M. J. (2016). 2D representation of transcriptomes by t-SNE exposes relatedness between human tissues. *PLoS ONE*, 11(2):1–6.

Thijs, R. D., Surges, R., O'Brien, T. J., and Sander, J. W. (2019). Epilepsy in adults. *The Lancet*, 6736(18).

Tormene, P., Giorgino, T., Quaglini, S., and Stefanelli, M. (2008). Matching incomplete time series with dynamic time warping: An algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine*, 45(1):11–34.

Treiman, D. M., Meyers, P. D., Walton, N. Y., Collins, J. F., Colling, C., Rowan, A. J., Handforth, A., Faught, E., Calabrese, V. P., Uthman, B. M., Ramsay, R. E., Mamdani, M. B., Yagnik, P., Jones, J. C., Barry, E., Boggs, J. G., and Kanner,

A. M. (1998). A Comparison of Four Treatments for Generalized Convulsive Status Epilepticus. *New England Journal of Medicine*, 339(12):792–798.

Trinka, E., Cock, H., Hesdorffer, D., Rossetti, A. O., Scheffer, I. E., Shinnar, S., Shorvon, S., and Lowenstein, D. H. (2015). A definition and classification of status epilepticus - Report of the ILAE Task Force on Classification of Status Epilepticus. *Epilepsia*, 56(10):1515–1523.

van den Berg, J. C., editor (1999). *Wavelets in Physics*. Cambridge University Press, Cambridge.

Van Der Maaten, L. J. P. and Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*.

van Drongelen, W., Nayak, S., Frim, D. M., Kohrman, M. H., Towle, V. L., Lee, H. C., McGee, A. B., Chico, M. S., and Hecox, K. E. (2003). Seizure anticipation in pediatric epilepsy: use of kolmogorov entropy. *Pediatric Neurology*, 29(3):207–213.

Varatharajah, Y., Berry, B., Cimbalnik, J., Kremen, V., Van Gompel, J., Stead, M., Brinkmann, B., Iyer, R., and Worrell, G. (2018). Integrating artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy. *Journal of Neural Engineering*, 15(4):046035.

Vázquez, E. V., Ferreiro, J. P., Miranda, J. G. V., and González, A. P. (2008). Mul-

tifractal Analysis of Pore Size Distributions as Affected by Simulated Rainfall. *Vadose Zone Journal*, 7(2):500.

Venema, V., Bachner, S., Rust, H. W., and Simmer, C. (2006). Statistical characteristics of surrogate data based on geophysical measurements. *Nonlinear Processes in Geophysics*, 13(4):449–466.

Vidal-Vázquez, E., Camargo, O., Vieira, S., Miranda, J., Menk, J., Siqueira, G., Mirás-Avalos, J., and Paz González, A. (2013). Multifractal Analysis of Soil Properties along Two Perpendicular Transects. *Vadose Zone Journal*, 12(3).

Wagenaar, J. B., Brinkmann, B. H., Ives, Z., Worrell, G. A., and Litt, B. (2013). A multimodal platform for cloud-based collaborative research. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1386–1389. IEEE.

Walker, M. C., White, H. S., and Sander, J. W. A. S. (2002). Disease modification in partial epilepsy. *Brain : a journal of neurology*, 125(Pt 9):1937–1950.

Wang, L., Wang, C., Fu, F., Yu, X., Guo, H., Xu, C., Jing, X., Zhang, H., and Dong, X. (2011). Temporal lobe seizure prediction based on a complex Gaussian wavelet. *Clinical Neurophysiology*, 122(4):656–663.

Wang, Y., Goodfellow, M., Taylor, P. N., and Baier, G. (2014). Dynamic Mechanisms of Neocortical Focal Seizure Onset. *PLoS Computational Biology*, 10(8):e1003787.

Wang, Y., Trevelyan, A. J., Valentin, A., Alarcon, G., Taylor, P. N., and Kaiser, M. (2017). Mechanisms underlying different onset patterns of focal seizures. *PLOS Computational Biology*, 13(5):e1005475.

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12):1–20.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H., Franois, R., Henry, L., and Mller, K. (2019). *dplyr: A Grammar of Data Manipulation*. R package version 0.8.0.1.

Wickham, H. and Henry, L. (2019). *tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions*. R package version 0.8.3.

Wiebe, S. and Jette, N. (2012). Pharmacoresistance and the role of surgery in difficult to treat epilepsy. *Nature Reviews Neurology*, 8(12):669–677.

Wilke, C. O. (2019). *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'*. R package version 1.0.0.

Williams, P. A., Hellier, J. L., White, A. M., Staley, K. J., and Dudek, F. E. (2007). Development of Spontaneous Seizures after Experimental Status Epilepticus: Implications for Understanding Epileptogenesis. *Epilepsia*, 48(s5):157–163.

Williams, P. A., White, A. M., Clark, S., Ferraro, D. J., Swiercz, W., Staley, K. J., and Dudek, F. E. (2009). Development of spontaneous recurrent seizures after kainate-induced status epilepticus. *Journal of Neuroscience*, 29(7):2103–2112.

Williamson, J. R., Bliss, D. W., Browne, D. W., and Narayanan, J. T. (2012). Seizure prediction using EEG spatiotemporal correlation structure. *Epilepsy & Behavior*, 25(2):230–238.

Wolfe, H. E. (2012). *Introduction to non-Euclidean geometry*. Dover Publications Inc., New York.

Wong, M. T., Chen, J., Narayanan, S., Lin, W., Anicete, R., Kiaang, H. T. K., De Lafaille, M. A. C., Poidinger, M., and Newell, E. W. (2015). Mapping the Diversity of Follicular Helper T Cells in Human Blood and Tonsils Using High-Dimensional Mass Cytometry Analysis. *Cell Reports*, 11(11):1822–1833.

Xu, H.-C., Gu, G.-F., and Zhou, W.-X. (2017). Direct determination approach for the multifractal detrending moving average analysis. *Physical Review E*, 96(5):052201.

Yuan, Q., Zhou, W., Liu, Y., and Wang, J. (2012). Epileptic seizure detection with linear and nonlinear features. *Epilepsy & Behavior*, 24(4):415–421.

Zandi, A. S., Dumont, G. a., Javidan, M., and Tafreshi, R. (2011). Epileptic seizure prediction using variational mixture of Gaussians. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 7549–7552. IEEE.

Zandi, A. S., Tafreshi, R., Javidan, M., and Dumont, G. a. (2013). Predicting Epileptic Seizures in Scalp EEG Based on a Variational Bayesian Gaussian Mixture

Model of Zero-Crossing Intervals. *IEEE Transactions on Biomedical Engineering*, 60(5):1401–1413.

Zappasodi, F., Olejarczyk, E., Marzetti, L., Assenza, G., Pizzella, V., and Tecchio, F. (2014). Fractal dimension of EEG activity senses neuronal impairment in acute stroke. *PLoS ONE*, 9(6):e100199.

Zeleke, T. B. and Si, B. C. (2006). Characterizing scale-dependent spatial relationships between soil properties using multifractal techniques. *Geoderma*, 134(3-4):440–452.

Zhang, Y., Zhou, W., Yuan, Q., and Wu, Q. (2014). A low computation cost method for seizure prediction. *Epilepsy Research*, 108(8):1357–1366.

Zhang, Y., Zhou, W., and Yuan, S. (2015). Multifractal Analysis and Relevance Vector Machine-Based Automatic Seizure Detection in Intracranial EEG. *International Journal of Neural Systems*, 25(06):1550020.

Zheng, Y., Wang, G., Li, K., Bao, G., and Wang, J. (2014). Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition. *Clinical Neurophysiology*, 125(6):1104–1111.

Zhou, W.-X. (2008). Multifractal detrended cross-correlation analysis for two nonstationary signals. *Physical Review E*, 77(6):066211.

Zhu, H. (2019). *kableExtra: Construct Complex Table with 'kable' and Pipe Syntax*. R package version 1.1.0.

Zorick, T. and Mandelkern, M. A. (2013). Multifractal Detrended Fluctuation Analysis of Human EEG: Preliminary Investigation and Comparison with the Wavelet Transform Modulus Maxima Technique. *PLoS ONE*, 8(7):e68360.