

**Nabeel Durrani**

✉ nabeel.durrani@uqconnect.edu.au

**Prof Amin Abbosh**

Acting Head of School  
School of Information Technology and Electrical  
Engineering  
The University of Queensland  
St Lucia QLD 4072

Dear Professor Abbosh,

In accordance with the requirements of the Bachelors of Engineering (Honours) and Science in the School of Information Technology and Electrical Engineering, I submit the following thesis entitled *Phase Coherence Classifier for Multivariate Non-Linear Time Series Based On Ordinal Partition Transition Network Topology*. The thesis was undertaken under the supervision of Associate Professor Marcus Gallagher. I declare that the work submitted in the thesis is my own, except as acknowledged in the text, and that it has not previously been submitted for a degree at the University of Queensland or any other institution.

Yours sincerely,

**Nabeel Durrani**





THE UNIVERSITY OF QUEENSLAND  
A U S T R A L I A

# Phase Coherence Classifier for Multivariate Non-Linear Time Series Based on Ordinal Partition Transition Network Topology

by  
NABEEL DURRANI  
*under the supervision of*  
MARCUS GALLAGHER.

School of Information Technology and Electrical Engineering,  
The University of Queensland.

*Submitted for the*  
Bachelors of Engineering (Honours) and Science  
*in the division of*  
Electrical & Biomedical Engineering and Mathematics.

November 3, 2019



## ABSTRACT

*Phase coherence* is a measure of *phase synchronisation*, a phenomenon known to occur in non-linear systems whereby the phases but not necessarily the amplitudes of multiple signals are time-synchronized. A novel multivariate phase coherence classifier based on *persistent homology* from *Topological Data Analysis* and *ordinal partition transition networks* from non-linear *symbolic dynamics* was developed then evaluated using a simulated (non-linear) Rössler system and using experimental electroencephalogram (EEG) data collected during seizure events. It was possible to evaluate the classifier, to test whether it measured phase coherence effectively, using

- Rössler system data because the system may be calibrated to produce multivariate time series in both phase coherent and non-phase coherent regimes – and
- seizure data, which is known to exhibit phase coherence.

In fact, a Amplitude Adjusted Fourier Transform surrogate data method was used to represent the null hypothesis that the classifier was measuring a linear or stochastic property of the data, as opposed to the non-linear property of phase coherence. Since the surrogate data procedure used was stochastic, there existed the potential for the null hypothesis to be rejected purely by chance – that is, by coincidental false positives produced by the surrogate data generation algorithm. Hence the typical 5% significance level required for rejecting the null hypothesis was strengthened to a lower value using the Šidák correction method.

Certain key results were obtained:

- The classifier was evaluated using a phase coherent Rössler system time window: Using  $m = 99$  surrogate data realizations and a 5% significance level Šidák corrected to  $\alpha$ , a one tailed  $z$ -test on the classifier values yielded  $p = 5.7 \times 10^{-41} < \alpha$ .
- A similar analysis was applied using  $m = 19$  with 40 seizure epochs (2 s) randomly selected from each of 3 patients: Permutation entropy selected the five most phase coherent channels per sliding window to evaluate the classifier on. The  $z$ -test null hypothesis was rejected for 22%, 16%, and 14% of all epochs for increasingly stronger Šidák corrections on the level of an epoch, a patient, and all patients respectively. These values were significantly higher than the 5% value that would be expected to occur purely by chance.

The following conclusions were made:

- Phase coherent and non-phase coherent classifier values may be considered to be drawn from distinct distributions.
- The results provide further confidence that phase coherence as opposed to some other property of the data is being measured by the mean of the classifier.
- The standard deviation of the classifier value is a result of linear properties of the data, and may not be employed to measure phase coherence.

The method may be further evaluated through

- comparing it systematically to other measures of phase coherence, and
- applying it to the problem of seizure detection, or the detection of other pathologies that may be identified through measures of phase coherence.

# CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Project Context and Background . . . . .	1
1.1.1	Cybernetics and Systems Theory . . . . .	1
1.1.2	Linear Systems, Non-Linear Systems, and Chaos . . . . .	1
1.1.3	Chaos Theory in the Biomedical World . . . . .	1
1.1.4	Synchronisation: How Spontaneous Order Arises in Non-Linear Systems . . . . .	2
1.2	Project Introduction . . . . .	2
1.2.1	Description . . . . .	2
1.2.2	Motivation, Importance, and Relevance . . . . .	4
<b>2</b>	<b>Technical Background</b>	<b>5</b>
2.1	Non-Linear Dynamics, Chaos Theory, and Bifurcation Theory . . . . .	5
2.1.1	The Butterfly Effect, The Lorenz System, and The Rössler System . . . . .	5
2.1.2	Chaos Theory and Bifurcation Theory . . . . .	5
2.1.3	Phase Portraits . . . . .	7
2.1.4	Nullclines and Symbolic Dynamics . . . . .	8
2.1.5	Takens' Theorem and Takens' Time Delay Embedding . . . . .	8
2.2	Information Theory and Permutation Entropy . . . . .	10
2.2.1	Information and Chaos . . . . .	10
2.2.2	Computing Permutation Entropy . . . . .	10
2.3	Epileptic Seizures . . . . .	11
2.3.1	Electroencephalography . . . . .	11
2.3.2	Differential Montage vs Referential Montage . . . . .	12
2.3.3	Characteristics of Scalp EEG Signals . . . . .	12
2.3.4	Epileptic Seizures . . . . .	13
2.3.5	Characteristics of Seizure Signals . . . . .	13
2.3.6	Characteristics of Neonatal Seizure Signals . . . . .	14
2.3.7	Seizure detection . . . . .	15
2.3.8	Channel selection . . . . .	15
2.4	Synchronisation and Phase Coherence . . . . .	16
2.4.1	What Is Synchronisation . . . . .	16
2.4.2	What Synchronisation Is Not . . . . .	16
2.4.3	On Phase Coherence and How Phase Synchronisation Is Measured . . . . .	18
2.5	Topological Data Analysis . . . . .	18
2.5.1	Vietoris-Rips Simplicial Complex . . . . .	19
2.5.2	Computing Persistent Homology using the Vietoris-Rips Complex . . . . .	19
2.5.3	Clique-Rank Filtration . . . . .	19
<b>3</b>	<b>Literature Review</b>	<b>25</b>
3.1	Sliding windows and persistence: An Application of Topological Methods to Signal Analysis . . . . .	25
3.2	Persistent Homology of Complex Networks for Dynamic State Detection . . . . .	25
3.3	Constructing Ordinal Partition Transition Networks from Multivariate Time Series . . . . .	26

<b>4 Project Overview</b>	<b>29</b>
4.1 Methodology Introduction . . . . .	29
4.1.1 Project Objectives . . . . .	29
4.1.2 Methodology Rationale and Motivation . . . . .	29
4.1.3 Data Sets . . . . .	29
4.2 Project Scope . . . . .	30
<b>5 Computing the Classifier</b>	<b>31</b>
5.1 Time Series Symbolization . . . . .	31
5.2 Periodicity in Symbolized Data . . . . .	31
5.3 Ordinal Partition Transition Networks . . . . .	31
5.4 Computing Persistent Homology . . . . .	31
5.5 Why Should the Classifier Measure Phase Coherence? . . . . .	33
5.5.1 A Heuristic Argument . . . . .	33
5.5.2 The Robustness of the Algorithm . . . . .	33
<b>6 Evaluating the Classifier</b>	<b>35</b>
6.1 Introduction . . . . .	35
6.1.1 Overview . . . . .	35
6.1.2 Computation of Surrogate Data . . . . .	35
6.1.3 Computation of Statistical Significance . . . . .	37
6.1.4 Evaluation of Statistical Significance . . . . .	37
6.2 Rössler System Data . . . . .	38
6.3 Seizure EEG Data . . . . .	44
<b>7 Conclusions and Reflections</b>	<b>51</b>
7.1 Conclusions . . . . .	51
7.2 Critical Appraisal of Project . . . . .	51
7.3 Future Work . . . . .	51
7.3.1 Further Evaluation of the Method . . . . .	51
7.3.2 Development of a Mathematical Framework for the Method . . . . .	52
<b>References</b>	<b>53</b>

# LIST OF FIGURES

1	Illustrates the concept of phase synchronisation. Three pairs of variables $x_1$ (solid) and $x_2$ (dotted) are plotted for (1) the uncoupled case (a, b), (2) the phase synchronized case (c, d), and for (3) the phase and amplitude synchronized case (e, f). For each case, they are plotted with respect to time (left) and with respect to each other (right). . . . .	3
2	Illustrates how scalp electroencephalogram (EEG) data is collected through electrodes (left) yielding a multivariate time series representation (right). . . . .	3
3	Illustrates the Lorenz system for particular parameter values by (a) plotting its variables with respect to each other — by plotting its projections onto the $xy$ -plane and $xz$ -plane, and (b) plotting each of its variables with respect to time. . . . .	6
4	Illustrates the Rössler system for particular parameter values by (a) plotting its variables with respect to each other — by plotting its projections onto the $xy$ -plane and $xz$ -plane, and (b) plotting each of its variables with respect to time. . . . .	6
5	Plots the solution of the logistic map for $r = 3.5$ . . . . .	7
6	Plots the phase portrait of a periodic dynamical system (left) of two variables $x$ and $y$ and the corresponding time series for one of its variables $x$ . . . . .	8
7	Illustrates how the nullcline of the system presented in Section 2.1.4 partitions the system's phase portrait into four regions corresponding to the $2^2$ possible values of $(\text{sign } \dot{x}, \text{sign } \dot{y})$ . . . . .	9
8	Illustrates the Takens' embedding applied to the $x(t)$ Lorenz system variable data using parameters $d = 3$ and $\tau = 0.1$ , as described in Section 2.1.5. Delayed copies of $x(t)$ , where the delay is indicated by line colour, are sampled as illustrated in (B). The resulting $d = 3$ time series are given in (A) and are used to produce the time delay embedding illustrated in (C). . . . .	9
9	Illustrates how a time series is symbolised for the purposes of calculating order 3 permutation entropy, as described in Section 2.2.2. . . . .	11
10	Displays a recording of EEG data. . . . .	12
11	The time series and phase portraits of a noise-driven dynamical system model of healthy seizure signals consisting of a fixed point (red dot) and a limit cycle (red circle). The black circle represents a <i>basin boundary</i> . When the system is perturbed by low amplitude noise, the system is trapped in either the fixed point (left) or the limit cycle (middle). On the other hand, when the system is perturbed by high amplitude noise, the system jumps erratically between the fixed point and limit cycle (right), and is hence considered a <i>multistable</i> system. This latter multistable system fits experimental data most closely, and is hence considered by Breakspear to be the model of healthy cortical activity. . . . .	14
12	Illustrates the electrode placement for the International 10-20 system. . . . .	15
13	Depicts two independent uncoupled pendulums (left) and their phases & periods (right). The oscillators may be coupled if they share a common mechanical support, so that their phases coincide or coincide with a constant phase offset. . . . .	17

14	Depicts two pendulums connected by a rigid beam. The pendulums are <i>strongly</i> as opposed to <i>weakly</i> coupled. Hence the pendulums <i>do not</i> exhibit synchronisation. . . . .	17
15	Displays a Topological Data Analysis pipeline . Given a point cloud (left), balls of radius $\mathbf{r}$ are constructed around each point, and based on the radius of these balls, a simplicial complex is constructed (centre: small $\mathbf{r}$ value, right: large $\mathbf{r}$ value). . . . .	20
16	Illustrates the construction of the Vietoris-Rips complex as described in Section 2.5.1. . . . .	20
17	Illustrates a persistence barcode. . . . .	21
18	Illustrates how a persistence barcode maps to a persistence diagram. The horizontal axis are the hole birth times, the vertical axis are the hole death times. The diagonal line represents when the birth and death times are equal, and as such, points close to this line are considered noise while points far from this line are considered to be significant. . . . .	21
19	Illustrates the clique-rank filtration as described in Section 2.5.3. A graph (top left) and its associated shortest path distance matrix (top middle) is used to compute a clique rank filtration (bottom row, using indices $i = 0$ through to $i = 5$ , cliques are shaded in blue). The corresponding persistence diagram (top right) contains two points, each corresponding to a cycle in the graph. . . . .	23
20	Illustrates cliques of various sizes. . . . .	23
21	Illustrates the graphs that may be generated in subfigure (C) of the Figure 22 pipeline, as described in Section 3.2. . . . .	27
22	Illustrates the pipeline employed by Myers et al. as described in Section 3.2.	27
23	Plots a multivariate dynamical system. . . . .	27
24	Illustrates the classifier computation methodology. The top row describes from left to right the classifier computation pipeline, while the bottom row provides supplementary explanatory illustrations. . . . .	32
25	A phase coherent Rössler system as described in Section 6.2. . . . .	40
26	A realization of surrogate data for the Figure 25 phase coherent Rössler system, as described in Section 6.2. . . . .	40
27	Plots in the time domain the $\mathbf{x}$ -component of the Figure 25 phase coherent Rössler system and the $\mathbf{x}$ -component of the Figure 25 surrogate data. . . . .	40
28	The persistence of the Figure 25 phase coherent Rössler system, as described in Section 6.2. . . . .	41
29	The persistence of the Figure 26 surrogate data, as described in Section 6.2. . . . .	41
30	A non-phase coherent Rössler system as described in Section 6.2. . . . .	42
31	The persistence of the Figure 30 non-phase coherent Rössler system, as described in Section 6.2. . . . .	42
32	The classifier means and standard deviations of the Section 6 Rössler system for various parameter values. The dotted line indicates the boundary between phase coherent parameter values ( $a < 0.206$ ) and non-phase phase coherent parameter values ( $a > 0.206$ ). . . . .	43
33	The classifier medians, maximums and minimums of the Section 6 Rössler system for various parameter values. The dotted line indicates the boundary between phase coherent parameter values ( $a < 0.206$ ) and non-phase phase coherent parameter values ( $a > 0.206$ ). . . . .	43
34	The classifier values with respect to time, for various values of $N$ , where $n = 4$ , as described in Section 6.3. . . . .	46

35	The classifier values with respect to time, for $\mathbf{N} = 5$ , where $n = 4$ , as described in Section 6.3. . . . .	46
36	The classifier values with respect to time, for various values of $\mathbf{N}$ , where $n = 3$ , as described in Section 6.3. . . . .	47
37	The classifier values with respect to time, for $\mathbf{N} = 3$ , where $n = 9$ , as described in Section 6.3. . . . .	47
38	The classifier values with respect to time, for various values of $\mathbf{N}$ , where $n = 2$ , as described in Section 6.3. . . . .	48
39	The classifier values with respect to time, for $\mathbf{N} = 9$ , where $n = 2$ , as described in Section 6.3. . . . .	48
40	An example on of a seizure epoch from a single EEG channel together with a realization of the corresponding surrogate data. The patient from which the depicted data was collected differs from the patient of Figure 41 . . . .	49
41	An example on of a seizure epoch from a single EEG channel together with a realization of the corresponding surrogate data. The patient from which the depicted data was collected differs from the patient of Figure 40 . . . .	49

# 1 INTRODUCTION

The minimal background required to describe the relationship of the project to various disciplines is presented below. Then the project itself, and its social relevance, is introduced and situated within the broad context of these disciplines.

## 1.1 Project Context and Background

### 1.1.1 Cybernetics and Systems Theory

Complex systems are biological, social, electronic, mechanical, & chemical systems and systems more generally. The field of cybernetics is the study of such complex systems using the tools of control theory, information theory, and other fields of applied mathematics [1]. It was founded by Norbert Wiener in 1948 with the publication of his seminal book *Cybernetics: Or Control and Communication in the Animal and the Machine* [1]. Although the ideas from the book were revolutionary at the time, within Wiener's lifetime they soon made their way into the mainstream engineering curriculum [1] under the umbrella term *systems theory*.

### 1.1.2 Linear Systems, Non-Linear Systems, and Chaos

Systems theory to the typical engineer focuses on using *Fourier analysis* to analyse *linear systems*, systems that may be understood in a reductionist manner: Linear systems are systems whose whole can be understood as a *superposition* of its parts. Although this superposition principle suffices for simple electromechanical systems, it often fails for more complicated systems, such as physiological systems. In this case the theory of *non-linear systems* is required.

Non-linear systems are characterized by the so called *Butterfly Effect*: The potential for small changes in the initial state of a system to have a large effect on the system. This is in contrast to linear systems, for which small changes in the initial state necessarily have commensurate small effect on the system.

Some non-linear systems give rise to the phenomenon of *chaos*, a phenomenon whereby the system exhibits complicated, ostensibly random and unpredictable behaviour [2]. Chaotic systems, however, are not stochastic [2]. Instead, they exhibit unpredictable behaviour *deterministically* [2].

### 1.1.3 Chaos Theory in the Biomedical World

As non-linear science and chaos theory matured, its findings were subsumed into systems theory [2]. In particular, non-linear methods were particularly beneficial in biomedical science where it came to be understood that reductionist methods are often inadequate in the study of complex systems [2]. That is, a physiological system, for example, cannot always be understood as the sum of its parts and a holistic approach must instead be taken [2]. Consider the following passage from Gleick [2]:

One perplexing feature of fibrillation is that many of the heart's individual components can be working normally. Often the heart's pacemaking nodes continue to send out regular electrical ticks. Individual muscle cells respond properly. Each cell receives its stimulus, contracts, passes the stimulus on, and relaxes to wait for the next stimulus. In autopsy the muscle tissue may

reveal no damage at all. That is one reason chaos experts believed that a new, global approach was necessary: the parts of a fibrillating heart seem to be working, yet the whole goes fatally awry. Fibrillation is a disorder of a complex system, just as mental disorders—whether or not they have chemical roots—are disorders of a complex system.

#### 1.1.4 Synchronisation: How Spontaneous Order Arises in Non-Linear Systems

Synchronisation is a characteristic of some non-linear systems, including biological systems [3]. It is the beneficial process of organization underlying the behaviour of schools of fish that allows them to coordinate so as to evade predators, but also the deleterious process underlying human epileptic seizures [3]. Consider the following passage from Strogatz [3]:

Until just a few years ago, the study of synchrony was a splintered affair, with biologists, physicists, mathematicians, astronomers, engineers, and sociologists labouring in their separate fields, pursuing seemingly independent lines of inquiry. Yet little by little, a science of sync has begun coalescing out of insights from these and other disciplines. This new science centres on the study of "coupled oscillators." Groups of fireflies, planets, or pacemaker cells are all collections of oscillators—entities that cycle automatically, that repeat themselves over and over again at more or less regular time intervals. Fireflies flash; planets orbit; pacemaker cells fire. Two or more oscillators are said to be coupled if some physical or chemical process allows them to influence one another. Fireflies communicate with light. Planets tug on one another with gravity. Heart cells pass electrical currents back and forth. As these examples suggest, nature uses every available channel to allow its oscillators to talk to one another. And the result of those conversations is often synchrony, in which all the oscillators begin to move as one.

## 1.2 Project Introduction

### 1.2.1 Description

*Phase coherence* is a measure of *phase synchronisation*, a phenomenon known to occur in non-linear systems whereby the phases but not necessarily the amplitudes of multiple signals are time-synchronized (Figure 1). A novel multivariate phase coherence classifier based on *persistent homology* from *Topological Data Analysis* and *ordinal partition transition networks* from non-linear *symbolic dynamics* was developed then evaluated using a simulated (non-linear) Rössler system and using experimental electroencephalogram (EEG) data collected during seizure events (Figure 2). It was possible to evaluate the classifier using

- Rössler system data because the system may be calibrated to produce multivariate time series in both phase coherent and non-phase coherent regimes – and
- seizure data, which is known to exhibit phase coherence [4].

The terms *phase coherence*, *persistent homology*, *Topological Data Analysis*, *ordinal partition transition network*, and *symbolic dynamics* will be described and expounded on in Section 2. Further, a more complete description of the project and its scope is deferred until Section 4.

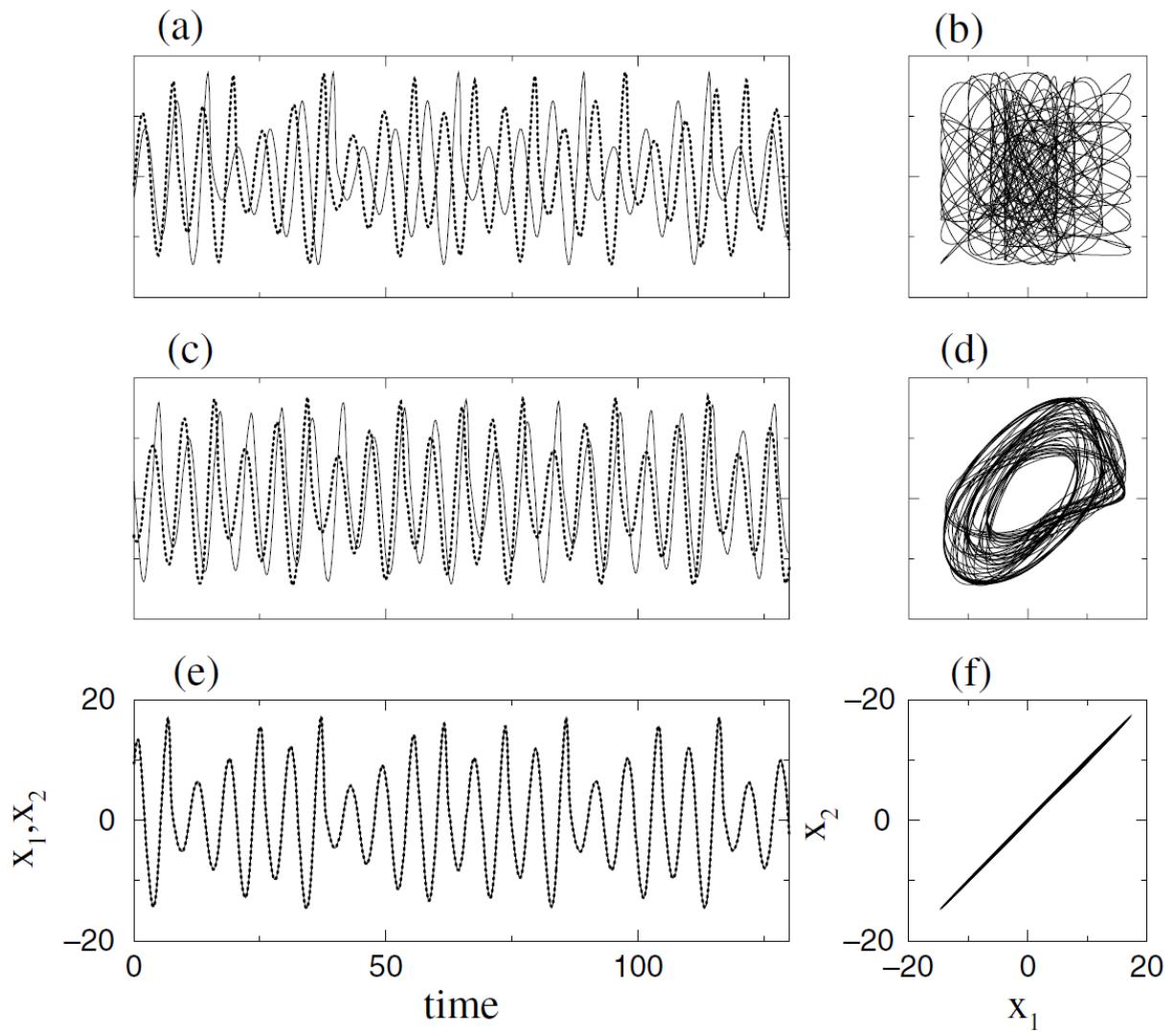


Figure 1: Illustrates the concept of phase synchronisation. Three pairs of variables  $x_1$  (solid) and  $x_2$  (dotted) are plotted for (1) the uncoupled case (a, b), (2) the phase synchronized case (c, d), and for (3) the phase and amplitude synchronized case (e, f). For each case, they are plotted with respect to time (left) and with respect to each other (right).

Source: [5]

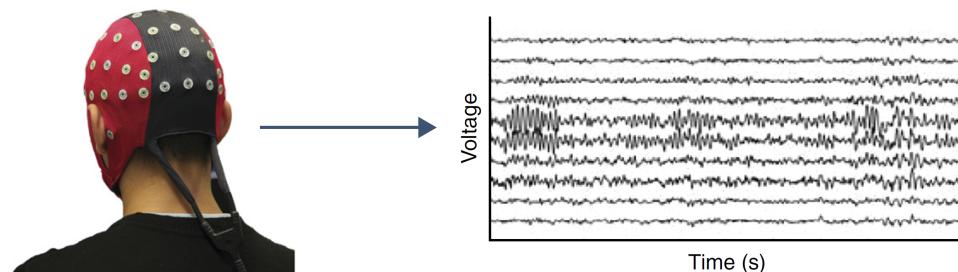


Figure 2: Illustrates how scalp electroencephalogram (EEG) data is collected through electrodes (left) yielding a multivariate time series representation (right).

Source: Adapted from [4]

### 1.2.2 Motivation, Importance, and Relevance

The reliable and robust measurement of phase coherence is beneficial in biomedical diagnosis, among other disciplines. It may be employed to classify

- seizure from non-seizure time intervals for EEG data [4], as alluded to above,
- vigilance states (i.e. eyes closed vs eyes open) using EEG data [4], and
- respiratory sinus arrhythmia using measures of heart rate and respiration [6]

among other states or pathologies.

## 2 TECHNICAL BACKGROUND

### 2.1 Non-Linear Dynamics, Chaos Theory, and Bifurcation Theory

#### 2.1.1 The Butterfly Effect, The Lorenz System, and The Rössler System

In 1961 while working on the problem of weather prediction using electronic computers, Edward Lorenz stumbled on the *Butterfly Effect* — that is, extreme sensitivity to initial conditions [2]. Rather than simulating his non-linear model of the weather from its initial time point, he took the shortcut of simulating it from an intermediate time point, manually providing his model with the initial conditions corresponding to the intermediate time point [2]. Defying all expectations, the resulting simulation did not reproduce the behaviour of his original simulation from the initial time point [2].

Lorenz was able to determine that his shortcut had failed to reproduce the expected behaviour, despite the fact that the simulation was deterministic as opposed to stochastic, as a result of finite precision in entering the initial conditions [2]. To investigate this phenomenon of extreme sensitivity to initial conditions he devised a simplified model, known as the *Lorenz system* (Figure 3), which also exhibited the Butterfly Effect under certain conditions [2]. The equations of the Lorenz system are as follows [7]:

$$\dot{x} = \sigma(y - x) \quad \dot{y} = rx - y - xz \quad \dot{z} = xy - bz$$

where  $\sigma, r, b > 0$  are parameters.

This model was further simplified by Otto Rössler in 1976 (Figure 4): The Rössler system shares many of the properties of the Lorenz system, but is more amenable to analysis [7]. The Rössler system equations are given by

$$\dot{x} = -y - z \quad \dot{y} = x + ay \quad \dot{z} = b + z(x - c)$$

where  $a, b, c$  are parameters [7, 8].

#### 2.1.2 Chaos Theory and Bifurcation Theory

An extremely brief overview of the phenomenon of chaos and bifurcations is provided by examining the *logistic map*  $x_{n+1} = rx_n(1-x_n)$  following the highly influential review article by Robert May in 1976 [2, 7]. The behaviour observed as the parameter  $r$  is incremented is as follows [7]:

- For  $0 < r < 3$ , the solution for  $x_n$  with respect to  $n$  converges to a single value.
- For  $r > 3$ , the solution for  $x_n$  with respect to  $n$  is periodic and oscillates between fixed values in the steady state. For example, for the case of  $r = 3.5$ ,  $x_n$  oscillates between four different values (Figure 5).
- At certain threshold parameter values ( $r_1 = 3, r_2 \approx 3.449, r_3 \approx 3.54, \dots$ ) a qualitative changes known as *bifurcations* are observed in the system: the number of points  $x_n$  oscillates between doubles. For example, for  $r_1 < r = 3.4 < r_2$ , the map oscillates between two values while for  $r_2 < r = 3.5 < r_3$  the map oscillates between four values (Figure 5).
- The differences between the threshold values  $r_1, r_2, \dots$  gradually decreases and approaches a limit: In fact  $r_5 \approx 3.568$  and  $r_\infty \approx 3.569$ .

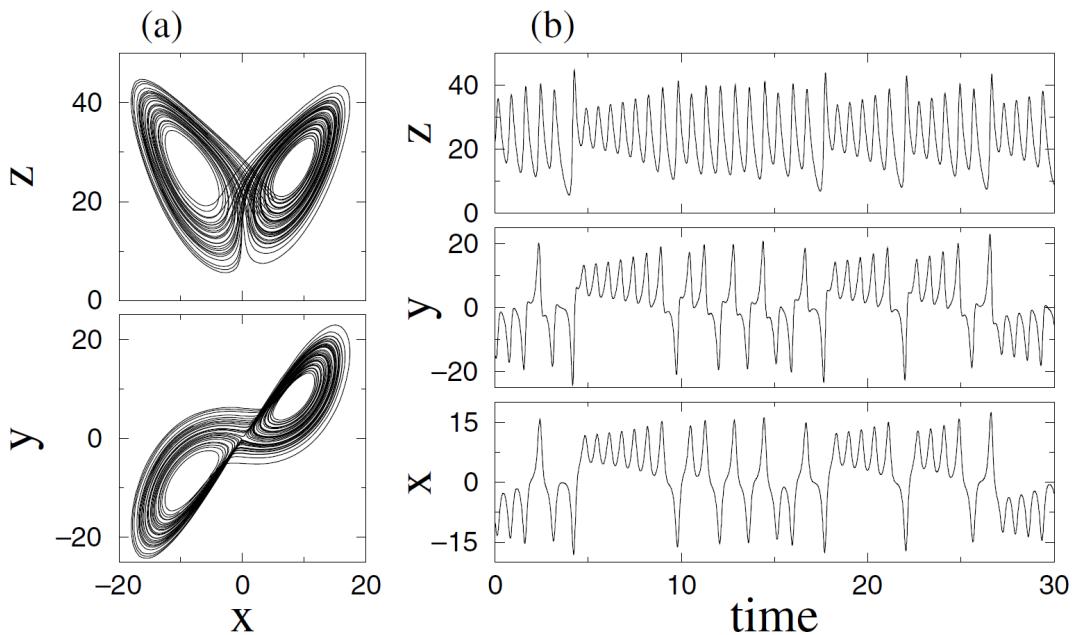


Figure 3: Illustrates the Lorenz system for particular parameter values by (a) plotting its variables with respect to each other – by plotting its projections onto the  $xy$ -plane and  $xz$ -plane, and (b) plotting each of its variables with respect to time.

Source: [5]

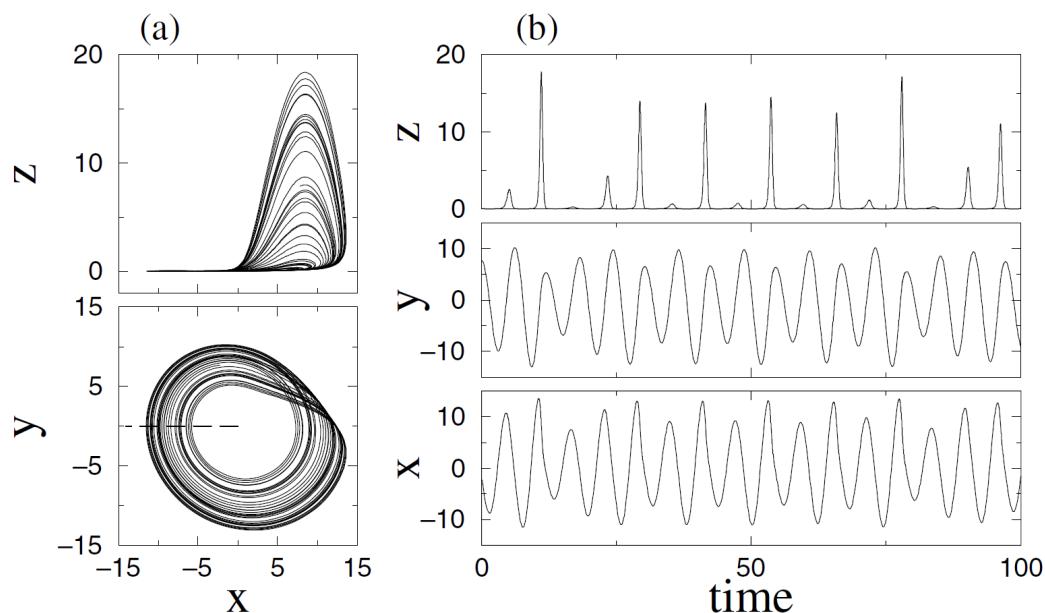


Figure 4: Illustrates the Rössler system for particular parameter values by (a) plotting its variables with respect to each other – by plotting its projections onto the  $xy$ -plane and  $xz$ -plane, and (b) plotting each of its variables with respect to time.

Source: [5]

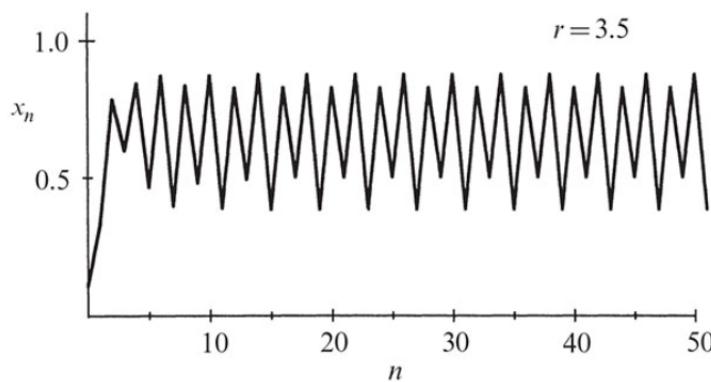


Figure 5: Plots the solution of the logistic map for  $r = 3.5$ .

Source: [7]

- For  $r > r_\infty$  the map oscillates between *infinitely* many values and is said to be chaotic.

Hence a deterministic process is able to produce effectively unpredictable behaviour. It may be challenging to determine whether unpredictable time series data has been generated chaotically or stochastically [9].

Although the logistic map is a discrete iterated map, the phenomenon of bifurcations and chaos are also observed in continuous systems, such as the Lorenz and Rössler system. Certain analogies may be made:

- Just as for the case of  $0 < r < 3$  the logistic map converged to a single value, each variable in a continuous system may converge to a single value given the appropriate initial conditions. That is, the system may converge to a *fixed point*,
- just as for the case of  $r > 3$  the logistic map exhibited periodic behaviour, each variable in a continuous system may exhibit periodic behaviour with respect to time and with respect to other system variables (Figure 3, Figure 4). That is, the system may converge to a *limit cycle*,
- just as qualitative changes known as bifurcations may occur in the logistic map for certain threshold parameter values, bifurcations may occur in continuous systems for certain threshold parameter values, and
- just as the logistic map may exhibit chaos for certain parameter values, continuous system may exhibit chaos for certain parameter values.

### 2.1.3 Phase Portraits

In the nineteenth century, Henri Poincaré pioneered a qualitative, geometric method of analysing dynamical systems based on the system's *phase portrait*, also known as its *state space* [7]. The phase portrait is a broad term referring to the plot of a system's variables, or their derivatives, with respect to each other rather than with respect to time. Although representation of a dynamical system has already been presented in Figure 3 and Figure 4, Figure 6 expounds on the concept.

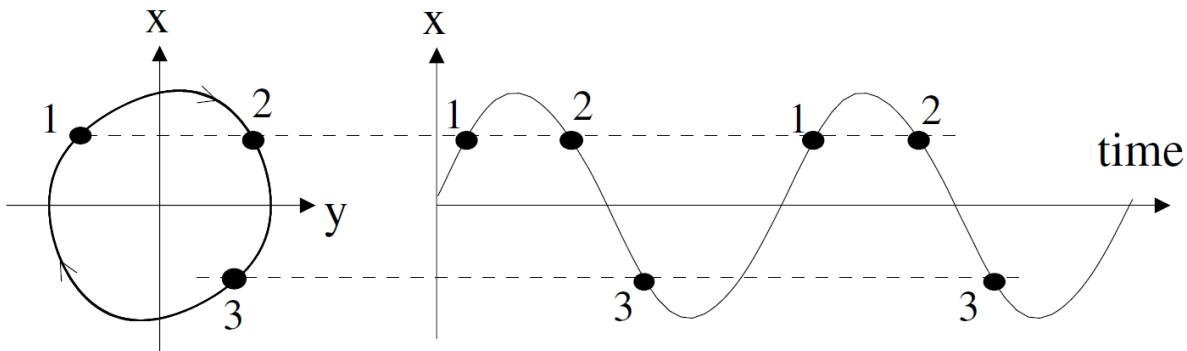


Figure 6: Plots the phase portrait of a periodic dynamical system (left) of two variables  $x$  and  $y$  and the corresponding time series for one of its variables  $x$ .

Source: [5]

### 2.1.4 Nullclines and Symbolic Dynamics

The *nullclines* of a system are the curves  $\dot{x} = 0$  or  $\dot{y} = 0$ . The nullclines of the system  $\dot{x} = x + e^{-y}$  and  $\dot{y} = -y$  for example are given by Figure 7 [7]. They partition the phase portrait into four possible regions, as indicated by the Figure.

The field of *symbolic dynamics* involves modelling time series using a space of discrete symbols [10, 11]. One possible way to assign symbols to time series values is to partition the phase portrait by nullclines as is done in Figure 7, assigning to each partition a symbol.

### 2.1.5 Takens' Theorem and Takens' Time Delay Embedding

Given a system of differential equations, its phase portrait may be faithfully constructed by solving the system for — say  $x(t)$ ,  $y(t)$ , and  $z(t)$  — through numerical integration with a sufficiently fine time step [7]. *Takens' Theorem* gives conditions under which samples of  $x$ ,  $y$  and  $z$  may be used to construct a phase portrait that faithfully preserves the qualitative properties of the original system [6]. The construction is referred to as a *Time Delay Embedding*, *Takens' Embedding*, or *State Space Reconstruction* and is given by

$$X(t_i) = \begin{pmatrix} x(t_i) \\ x(t_i + \tau) \\ \vdots \\ x(t_i + (d-1)\tau) \end{pmatrix}$$

for the variable  $x(t)$  where  $\tau$  and  $d$  are parameters, the *delay time* and *embedding dimension* respectively, that must be selected carefully [6]. It is possible to extend the embedding to incorporate all variables rather than just  $x(t)$ , as is described above [6].

Takens' theorem

- may be considered the non-linear analogue of Nyquist's theorem from linear system analysis, and more crucially
- may be employed, in the absence of a differential equation model, to construct a phase portrait from experimental data thought to be generated from a non-linear process.

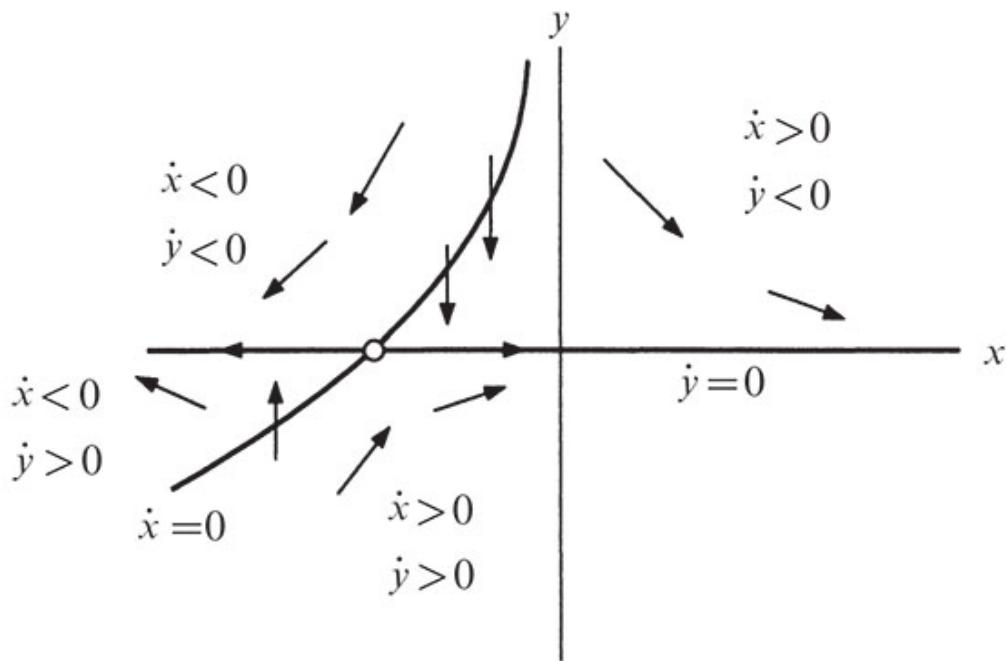


Figure 7: Illustrates how the nullcline of the system presented in Section 2.1.4 partitions the system's phase portrait into four regions corresponding to the  $2^2$  possible values of  $(\text{sign } \dot{x}, \text{sign } \dot{y})$ .

Source: [7]

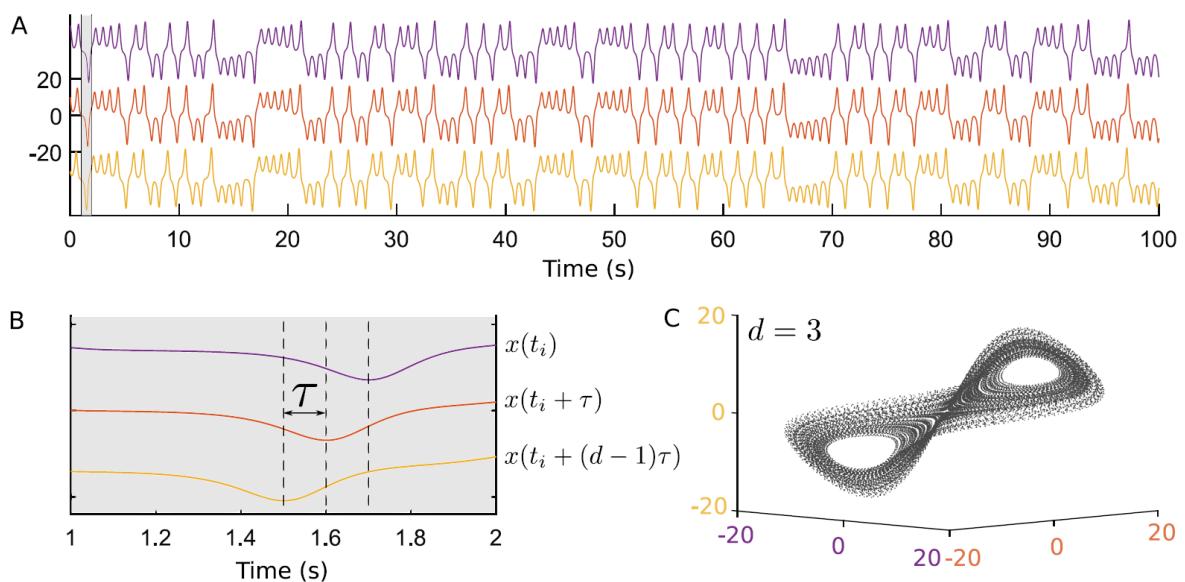


Figure 8: Illustrates the Takens' embedding applied to the  $\mathbf{x}(t)$  Lorenz system variable data using parameters  $d = 3$  and  $\tau = 0.1$ , as described in Section 2.1.5. Delayed copies of  $\mathbf{x}(t)$ , where the delay is indicated by line colour, are sampled as illustrated in (B). The resulting  $d = 3$  time series are given in (A) and are used to produce the time delay embedding illustrated in (C).

Source: [6]

## 2.2 Information Theory and Permutation Entropy

### 2.2.1 Information and Chaos

Claude Shannon's information theory was developed for the field of electronic communication systems, but has a history closely tied to the field of non-linear dynamics, chaos, and symbolic dynamics — in fact, to quote Gleick's description [2] of Robert Shaw's highly influential 1981 work *Strange attractors, chaotic behaviour, and information flow* [12]:

Chaos was the creation of information. One could imagine water flowing past an obstruction. As every hydrodynamicist and white-water canoeist knows, if the water flows fast enough, it produces whorls downstream. At some speed, the whorls stay in place. At some higher speed, they move. An experimenter could choose a variety of methods for extracting data from such a system, with velocity probes and so forth, but why not try something simple: pick a point directly downstream from the obstruction and, at uniform time intervals, ask whether the whorl is to the right or the left. If the whorls are static, the data stream will look like this: left-left-left-left-left-left-left-left-left-left-left-left-left-left-left-left-left-. After a while, the observer starts to feel that new bits of data are failing to offer new information about the system. Or the whorls might be moving back and forth periodically: left-right-left-right-left-right-left-right-left-right-left-right-left-right-left-right-left-right-. Again, though at first the system seems one degree more interesting, it quickly ceases to offer any surprises. As the system becomes chaotic, however, strictly by virtue of its unpredictability, it generates a steady stream of information. Each new observation is a new bit. This is a problem for the experimenter trying to characterize the system completely. "He could never leave the room," as Shaw said. "The flow would be a continuous source of information."

When talking of the information, as opposed to redundancy, of a data source, Shaw of course is referring to the mathematical quantity *Shannon entropy*.

### 2.2.2 Computing Permutation Entropy

Although Shannon entropy may be applied directly to a non-linear signal, an alternative *permutation entropy* based on symbolic dynamics and the time delay embedding is more robust to noise [11, 13]. This robustness owes to the fact that it is based on inequalities rather than on signal amplitude values — that is, it is based on *ordinal patterns*, the ordering of amplitudes in a time window relative to each other [11].

The computation of permutation entropy is described below [13]:

- A signal is divided into windows using a sliding window, and an *order*  $n$  of the permutation entropy calculation is chosen.
- As described in Figure 9 for  $n = 3$ , each window is assigned one of  $n!$  ordinal patterns. That is, for each window,  $n$  equally spaced samples are selected, and based on the relative amplitudes of these samples, the window is assigned a motif. Note that here a time delay embedding as described in Section 2.1.5 is effectively constructed.
- Given a section of the symbolised time series, the frequency that each of the  $n!$  permutations  $\pi_i$  occurs is collected to yield a discrete distribution  $p(\pi)$

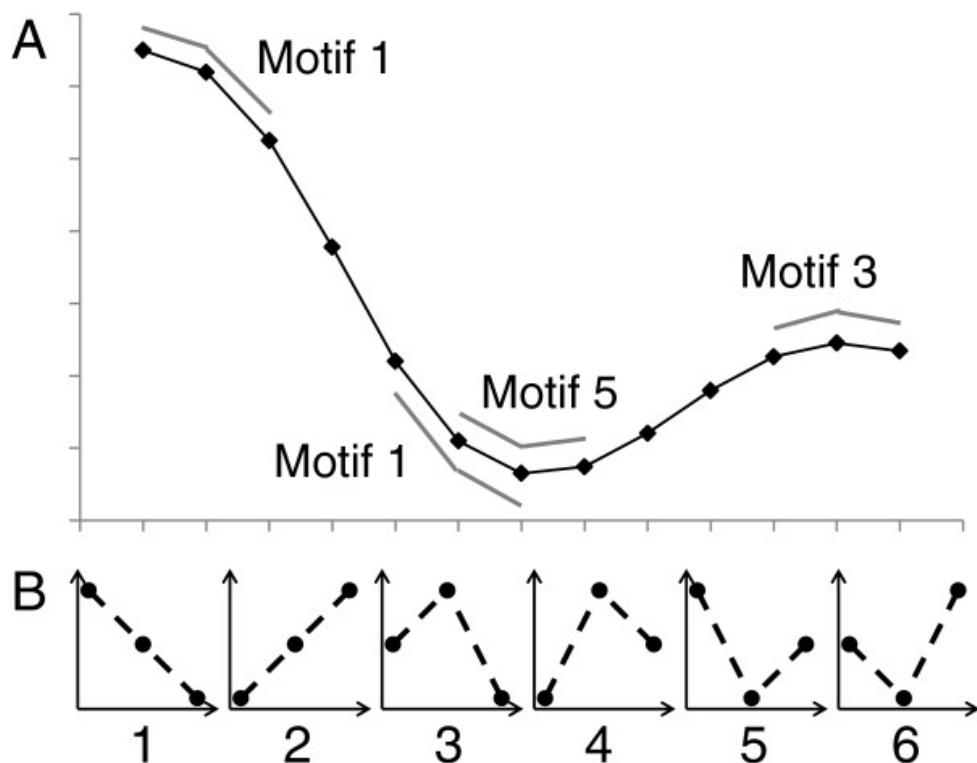


Figure 9: Illustrates how a time series is symbolised for the purposes of calculating order 3 permutation entropy, as described in Section 2.2.2.

Source: [14]

- The Shannon entropy of  $p(\pi)$  is computed to yield the permutation entropy of the section of the symbolised time series:

$$H(m) = - \sum_{i=1}^{n!} p(\pi_i) \ln(p(\pi_i))$$

## 2.3 Epileptic Seizures

### 2.3.1 Electroencephalography

*Electroencephalography (EEG)* is method commonly employed by physicians whereby an *electroencephalogram* is used to measure patients' brain electrical activity, which results from neuron ionic currents [15]. Typically

- multiple electrodes are placed along the scalp, with each electrode recording the amplitude of brain signals to an *EEG channel* (Figure 10), and
- *scalp EEG data* is collected – that is, EEG electrodes are placed non-invasively on the scalp [16].

However, in cases where greater accuracy and reduced noise is required [16], *intracranial EEG (iEEG) data* may instead be collected invasively by placing electrodes directly on the brain [15, 17]. Intracranial EEG is typically only performed in exceptional cases, for example, in cases of medically intractable epilepsy [17].

Often EEG data is divided into time windows referred to as *epochs*. The term *seizure epoch*, for example, would refer to a time window of an understood length during which a seizure event is known to have occurred.

### 2.3.2 Differential Montage vs Referential Montage

EEG data may be recorded in a *referential montage* or a *differential montage*, the most common of which is a *bipolar montage* [17]. A referential montage displays the brain signal amplitudes with respect to the electric potential of a reference electrode [17]. On the other hand, differential montages display the electric potential difference between pairs of electrodes [17].

### 2.3.3 Characteristics of Scalp EEG Signals

## Scalp EEG signals

- have a frequency between 0.5 Hz and 100 Hz and an amplitude of approximately 100  $\mu$ V [18],
  - possess properties characteristic of *non-linear* and *non-stationary* dynamical systems [19]:
    - Non-linear signals do not satisfy superposition and scaling properties and hence are not amenable to the usual electrical engineering frequency domain techniques typically, and instead must employ techniques from chaos and bifurcation theory [19].
    - Non-stationary signals are signals that may be modelled as a realization of a stochastic process whose *distribution at each time step changes* [20].

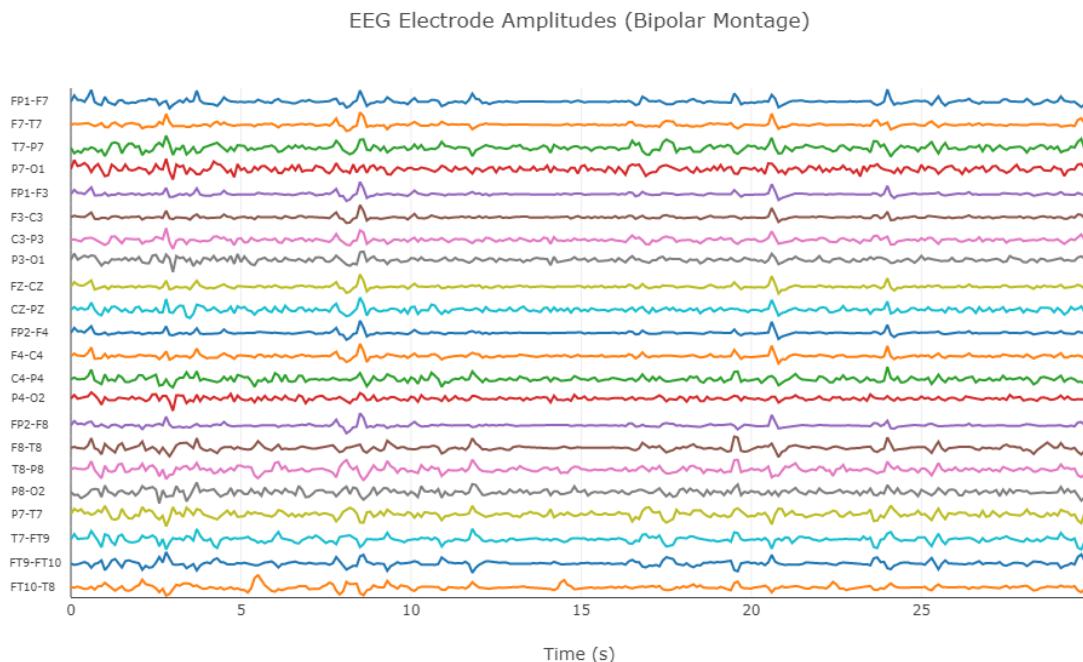


Figure 10: Displays a recording of EEG data.

- Stationary signals, on the other hand, can be generated by sampling from a single distribution across all time steps [20].

### 2.3.4 Epileptic Seizures

Epilepsy is a relatively common neurological disorder that afflicts 0.6% to 0.8% of the world population and is characterised by seizures [20]. An epileptic seizure may be

- *focal*, in which case the seizure is localized to certain regions of the brain, and hence to certain EEG channels if EEG data is collected from an epileptic patient [17];
- *general*, in which case the activity of the whole brain region is impaired by the seizure, and in which case the seizure is not localized to a particular region of the brain [17],
- *clinical*, in which case the epileptic patient exhibits overt behavioural symptoms [21], or
- *subclinical*, in which case the epileptic patient *does not* exhibit overt behavioural symptoms [21].

Neonatal seizures, as opposed to adult seizures, are effectively subclinical and focal [21]. Hence neonates require continuous scalp EEG monitoring [21].

Epileptic patients are said to be in

- the *ictal state* during a seizure,
- the *preictal state* immediately before a seizure,
- the *postictal state* immediately after a seizure, and
- the *interictal state* at all other times [22].

Clearly, an epileptic patient is in the interictal state the vast majority of the time [22].

### 2.3.5 Characteristics of Seizure Signals

In his widely cited 2017 work, Breakspear [4] summarizes the history of chaos theory vis-à-vis seizure signals:

Chaotic dynamics arise from unstable non-linear processes, yielding complex attractors with fractal geometry. In the 1980s, new computational algorithms allowed the quantification of these properties in empirical data. The broad appeal of chaos—the emergence of complex dynamics from simple rules—fueled the application of these algorithms to data from diverse systems. As in many fields, large-scale neurophysiological data, acquired during rest, sleep, cognition and seizures, were subsequently observed to possess the classic hallmarks of chaotic dynamics. The algorithms for detecting chaos certainly yielded valid results for the theoretical dynamical systems in which they were developed—lengthy time series data obtained by integrating non-linear systems. Unfortunately, substantial limitations were soon identified in the application of these algorithms to noisy, non-stationary and often relatively brief empirical data. In particular, it was shown that filtered (linear) noise could also yield the numeric values that had been previously associated with chaos. Recognition of these limitations led to a reappraisal of prior findings and to

more circumspect conclusions regarding the role of simple chaotic dynamics in large-scale neural systems. In retrospect, the initial application of measures of chaotic dynamics to time series data may have reflected a confirmation bias – an implicit objective to show that empirical data are consistent with an appealing theoretical framework. Alternative hypotheses – that the measures of chaos might be generated by filtered linear noise with no further temporal structure, for example – were not systematically tested.

His research indicates that

- healthy EEG signals are not chaotic but rather jump unpredictably between 10 Hz oscillations and noise. In the language of non-linear theory, healthy EEG signals may be modelled by a dynamical system with a limit cycle and a fixed point in its phase portrait, as described in further detail by Figure 11. Further,
- seizure EEG signals exhibit sustained non-linear behaviour.

### 2.3.6 Characteristics of Neonatal Seizure Signals

We focus henceforth on neonatal seizures, which

- are typically 1.5 min to 2 min in duration and are localized in the central and temporal (i.e. near the temples) regions of the brain,
- are characterised in their *ictal state* by high energy, determinism, a narrow frequency band, pseudo-periodicity, and low Shannon entropy (a measure of uncertainty), and
- are characterised in their *interictal state* by low energy, (Gaussian) randomness, and high Shannon entropy [21].

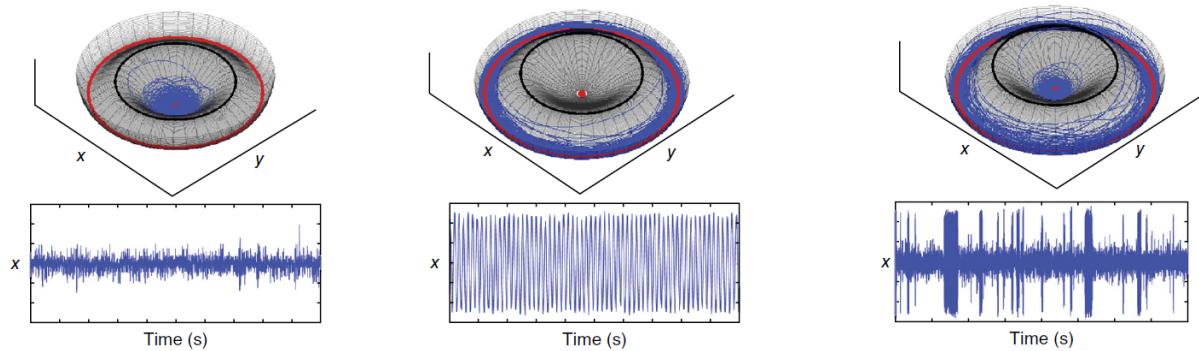


Figure 11: The time series and phase portraits of a noise-driven dynamical system model of healthy seizure signals consisting of a fixed point (red dot) and a limit cycle (red circle). The black circle represents a *basin boundary*. When the system is perturbed by low amplitude noise, the system is trapped in either the fixed point (left) or the limit cycle (middle). On the other hand, when the system is perturbed by high amplitude noise, the system jumps erratically between the fixed point and limit cycle (right), and is hence considered a *multistable* system. This latter multistable system fits experimental data most closely, and is hence considered by Breakspear to be the model of healthy cortical activity.

Source: [4]

### 2.3.7 Seizure detection

In principle, seizures may transition between the interictal and ictal states

- abruptly, in which case the seizure is *unable* to be detected, or
- gradually, in which case the change in dynamics of the non-linear system model that generates the EEG signal is *able* to be detected [20].

In the latter case, according to Ghadyali [17], automated seizure detection systems may be employed and classified into the following two categories.

- *Online seizure detection*:
  - *Objective*: To detect the seizure prior to its occurrence
  - *Motivation*: To allow the epileptic patient to secure themselves to prevent injury or to administer medication prior to the seizure event
  - *Remark*: There is a lower tolerance for false positives compared to post-hoc seizure detection. This is because false positives burden patient care providers.
- *Post-hoc seizure detection*:
  - *Objective*: To detect the seizure prior after it has occurred
  - *Motivation*: Physicians must manually review EEG recordings and note the time, duration, spatial localization (or lack thereof) of seizure events. This process is time consuming, so an automated system is highly desirable.
  - *Remark*: There is a higher tolerance for false positives compared to online seizure detection

### 2.3.8 Channel selection

The International 10-20 system is a standardised placement and labelling system for scalp EEG electrodes (Figure 12). In the case of focal seizures, which is the norm in neonates, the seizure is localized to certain channels so that many channels are redundant [23, 24]. Hence a channel selection method is desirable in order to reduce the dimensionality of the

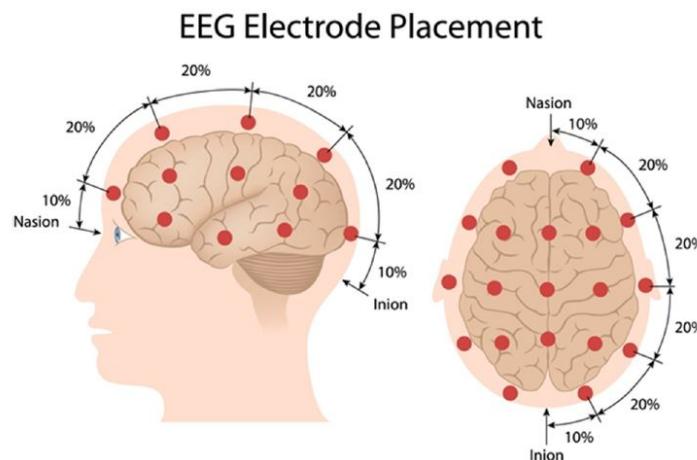


Figure 12: Illustrates the electrode placement for the International 10-20 system.

Source: [16]

problem and to reduce overfitting of the model [23]. Channel selection methods follow the following pipeline [16]:

1. Multichannel signal acquisition and pre-processing
2. Channel selection
3. Feature extraction
4. EEG signal classification

## 2.4 Synchronisation and Phase Coherence

The concept of synchronisation introduced in Section 1.1.4 is expounded upon here

### 2.4.1 What Is Synchronisation

Synchronisation is

- "adjustment of rhythms of oscillating objects due to their weak interaction" [5] , or
- the coupling of oscillators through weak interaction.

The meaning of the term *weak interaction* here is described in Section 2.4.2.

Referring to Figure 13, to quote Pikovsky et al. [5]

Suppose that the two clocks are fixed on a common support, and let this be a not absolutely rigid beam. This beam can bend, or it may vibrate slightly, moving from left to right, this does not matter much. What is really important is only that the motion of each pendulum is transmitted through the supporting structure to the other pendulum and, as a result, both clocks "feel" each other: they interact through the vibration of the common support ...it may alter the rhythms of both clocks! ...Examples of self-sustained oscillatory systems are electronic circuits used for the generation of radio-frequency power ...Belousov-Zhabotinsky and other oscillatory chemical reactions, pacemakers (sino-atrial nodes) of human hearts or artificial pacemakers that are used in cardiac pathologies, and many other natural and artificial systems.

Synchronization may result in the following [5]:

- The relationship of frequencies between oscillators in what is known as *locking*
- A relationship, such as equality with a constant offset, in what is known as *phase locking*

### 2.4.2 What Synchronisation Is Not

Synchronisation is not the following [5]:

- A state — it is a process, and as such synchronisation is not merely synchronous variation between oscillators
- Synchronisation is the result of two independent systems, with their own independent behaviour. If the coupling between two oscillators is too strong, as depicted by Figure 14, then the oscillators are said to be *strongly coupled* as opposed to *weakly coupled*, because both systems may be modelled by a single dynamical system rather than two.

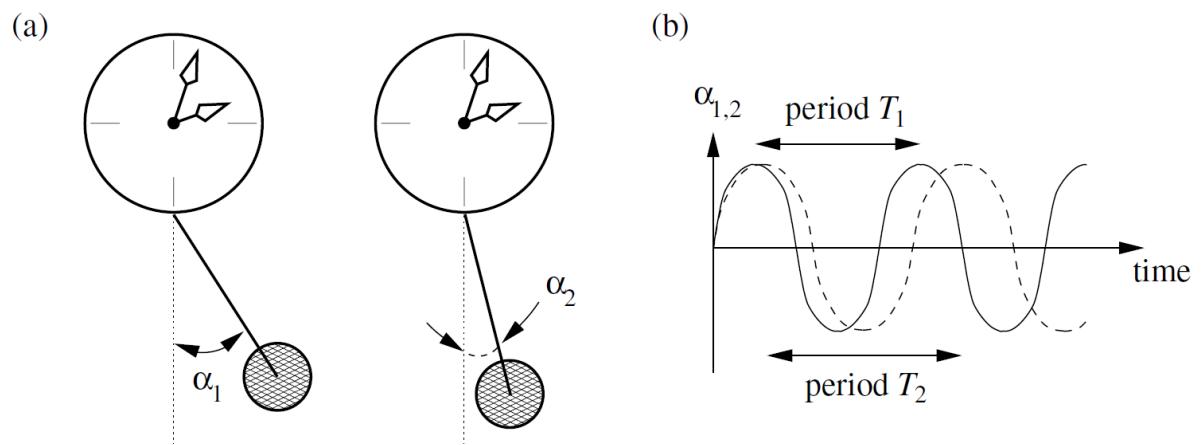


Figure 13: Depicts two independent uncoupled pendulums (left) and their phases & periods (right). The oscillators may be coupled if they share a common mechanical support, so that their phases coincide or coincide with a constant phase offset.

Source: [5]

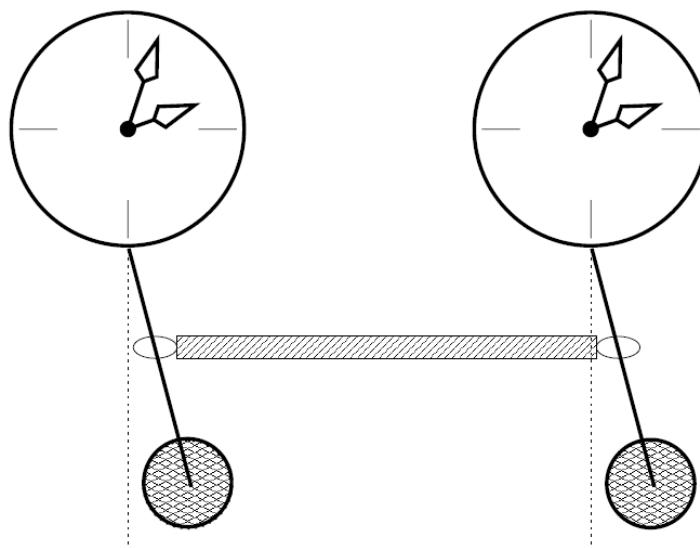


Figure 14: Depicts two pendulums connected by a rigid beam. The pendulums are *strongly* as opposed to *weakly* coupled. Hence the pendulums *do not* exhibit synchronisation.

Source: [5]

### 2.4.3 On Phase Coherence and How Phase Synchronisation Is Measured

Given two oscillators with phases  $\alpha_1$  and  $\alpha_2$ , one definition of phase synchronization is as the locking of phases [6]:

$$|m\alpha_1 - n\alpha_2| < \text{constant} \quad m, n \in \mathbb{N} \quad \text{Oscillator amplitudes may be independent}$$

Further, one definition of the phase coherence — a measure of the degree of phase synchronisation between signals corresponding to the indices 1 and 2 — is

$$\Pi(\omega) = \frac{1}{L} \left| \sum_{n=1}^L e^{i(\alpha_1(\omega, t_n) - \alpha_2(\omega, t_n))} \right| \quad 0 \leq \Pi(\omega) \leq 1$$

where the following should be noted [6]:

- $\alpha_1(\omega, t_n) - \alpha_2(\omega, t_n)$  is the phase difference of frequency  $\omega$  Fourier components
- The phases may be defined using the *Hilbert Transform*
- $\Pi(\omega) = 1$  indicates perfect phase locking
- $\Pi(\omega) = 0$  indicates perfect independence
- Phase coherence is merely an indication of phase synchronization. Signals that are perfectly phase coherent are not necessarily synchronized: The relationship between signal phases may be the result of a coincidence as opposed to a synchronization process.

## 2.5 Topological Data Analysis

Topology is the study of geometry without reference to a distance function (i.e. a metric such as Euclidean distance): To a topologist a doughnut and a coffee cup are one and the same because they each possess exactly one hole so that one may be elastically deformed to form the other. A subfield of topology, *algebraic topology*, is responsible for a technique called *homology*, or hole-counting [25, 26]. To describe how homology may be computed, terms germane to another subfield of topology known as *computational topology* must be defined [25]:

- A *0-simplex* is a point, a *1-simplex* is a line, and a *2-simplex* is a triangle. An *n-simplex* is said to be the *n*-dimensional generalization of these geometric objects [25, 26].
- A *simplicial complex* is a collection of simplices, possibly of mixed dimensions, defined in relation to each other such that certain properties are satisfied [25, 26].
- *Persistent homology* refers to a method of constructing a simplicial complex from point cloud data and computing homology from this complex **in a manner that is robust to noise** [25, 26].

The prototypical pipeline for computing persistent homology, which employs the Vietoris-Rips complex, is given by Figure 15.

### 2.5.1 Vietoris-Rips Simplicial Complex

Referring to Figure 15, the Vietoris-Rips simplicial complex is constructed through the following process [26]:

1. Vertices  $\{v_1, v_2, \dots, v_n\}$  in  $\mathbb{R}^n$  are selected by selecting points from a data set.
2. For a given  $r$  balls of radius  $r$  are constructed around each vertex.
3. As illustrated in Figure 16, the simplex  $[v_1, v_2, \dots, v_k]$  is added to the simplicial complex if and only if the diameter of the set  $\{x_1, x_2, \dots, x_k\}$  is less than  $r$ .

### 2.5.2 Computing Persistent Homology using the Vietoris-Rips Complex

Referring to Section 2.5.1 and Figure 15, note that the homology of the complex depends on the parameter  $r$  [25, 26].

- If  $r$  is too small many holes may be observed in the complex that are merely artefacts of noise – and alternately,
- if  $r$  is too large potentially all components may end up being connected (Figure 15, right) yielding little information from the data.

To select an appropriate  $r$  value a *persistence barcode* must be defined. A persistence barcode is a diagram that contains  $r$  as the independent variable on the horizontal axis with the vertical axis corresponding to holes (or the  $n$ -dimensional generalization thereof) present in the simplicial complex (Figure 17) [25, 26]. The initial points of each bar are referred to as hole *birth* times and the final points of each bar are referred to as hole *death* times [25, 26]. Referring to the Figure 17, small bars correspond to features that are the result of noise while long bars correspond to features that represent the data [29].

An equivalent representation of the persistence diagram is a *persistence barcode*, which is described in Figure 18 [25, 30]. *Persistent homology* refers to the process of only considering bars with sufficiently large length.

### 2.5.3 Clique-Rank Filtration

In describing the Vietoris-Rips complex in Section 2.5.1, the construction process was described in terms of constructing the complex directly from data points. Briefly and without describing the details, it may alternately be described as constructing a graph, a *proximity graph*, then finding the *clique complex*, also known as the *flag complex* of this graph [28].

The computation of the *clique rank filtration* based on the clique complex is described below [31] (Figure 19):

- The pairwise distance matrix of shortest paths is computed (e.g. using Dijkstra's algorithm)
- To index  $i = 0$  associate to each node  $N$  in the graph all nodes a shortest path distance of  $i = 0$  from  $N$  – that is, associate *all* nodes in the graph to the index  $i = 0$ . Similarly, to index  $i = 1$  associate to each node  $N$  in the graph all nodes a shortest path distance of  $i = 1$  from  $N$  – that is, associate *all* edges in the graph to the index  $i = 1$ .

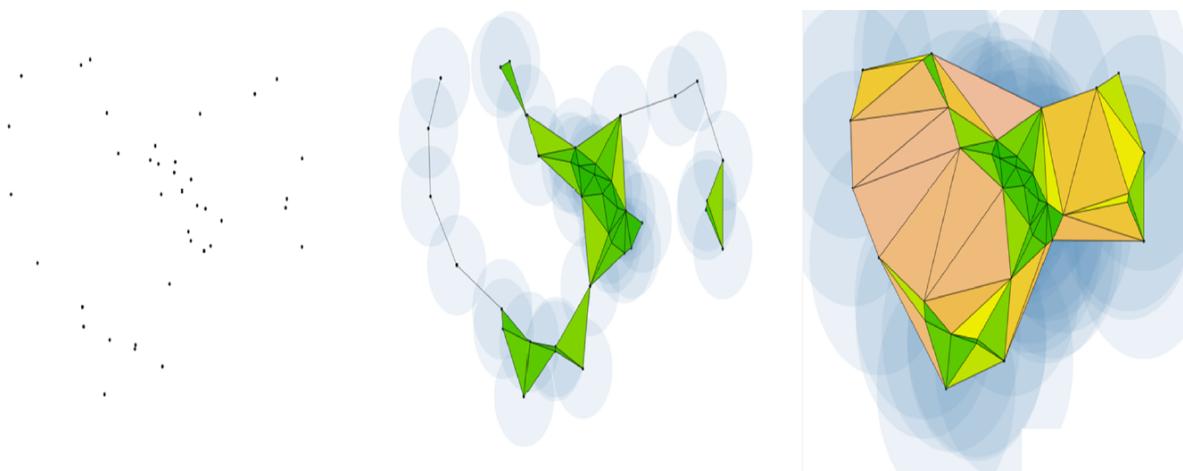


Figure 15: Displays a Topological Data Analysis pipeline . Given a point cloud (left), balls of radius  $r$  are constructed around each point, and based on the radius of these balls, a simplicial complex is constructed (centre: small  $r$  value, right: large  $r$  value).

Source: [27]

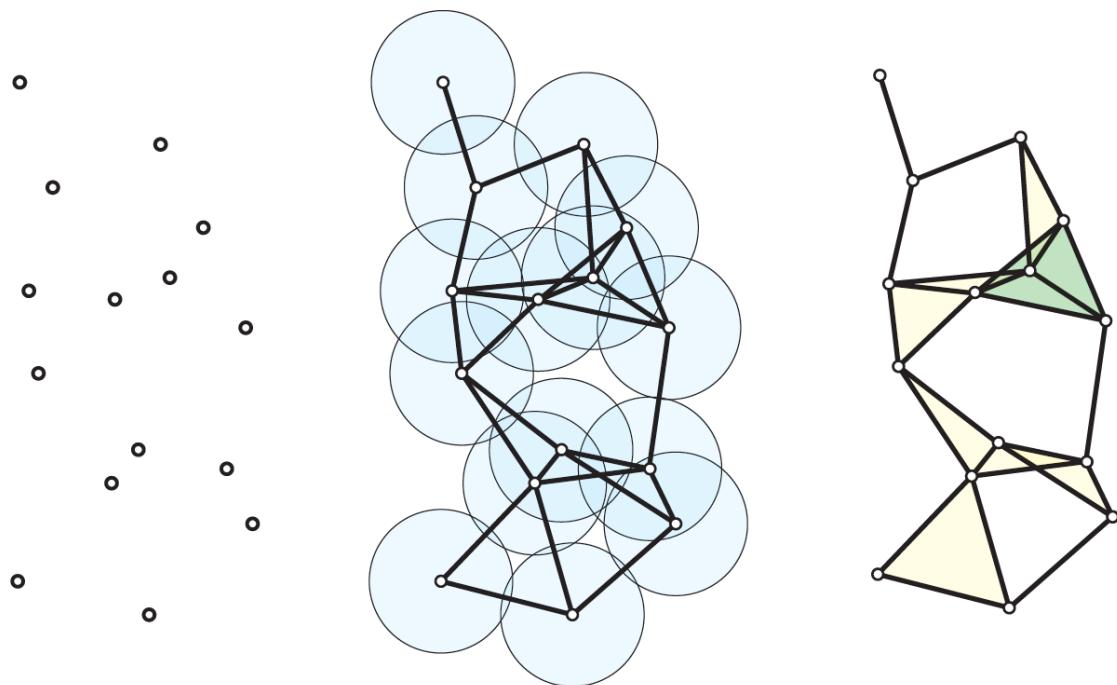


Figure 16: Illustrates the construction of the Vietoris-Rips complex as described in Section 2.5.1.

Source: Adapted from [28]

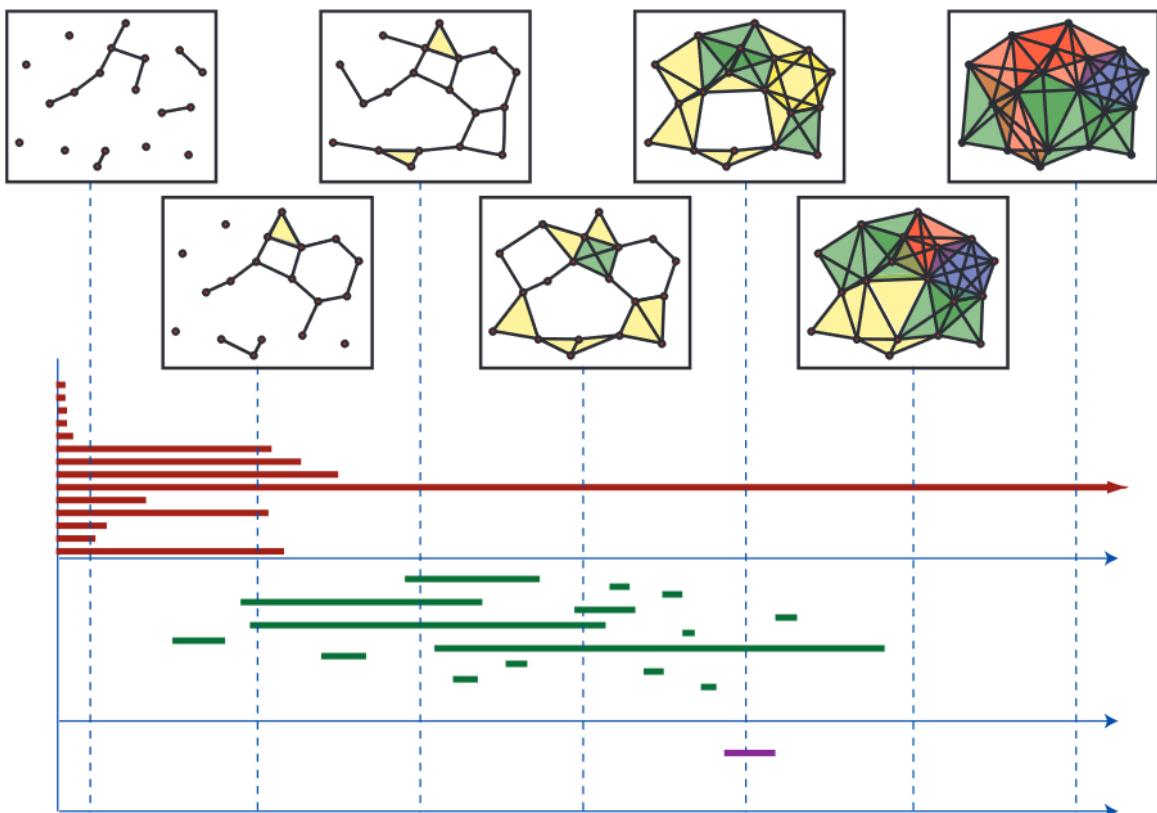


Figure 17: Illustrates a persistence barcode.

Source: [29]

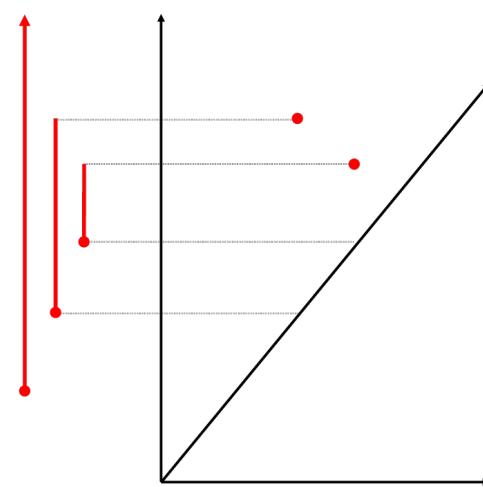


Figure 18: Illustrates how a persistence barcode maps to a persistence diagram. The horizontal axis are the hole birth times, the vertical axis are the hole death times. The diagonal line represents when the birth and death times are equal, and as such, points close to this line are considered noise while points far from this line are considered to be significant.

Source: [30]

- To index  $i = m$  associate to each node  $N$  in the graph all nodes a shortest path distance of  $i = m$  from  $N$ .

Before the remainder of the procedure is described, the following terminology and clarifications are introduced:

- **The index  $i$  takes the role of the variable  $r$  in the barcode diagram for the Vietoris-Rips filtration, as described in Section 2.5.2.**
- A *clique* of size  $n$  is a graph or subgraph in which all  $n$  nodes are mutually connected by edges (Figure 20). A clique may be considered a simplex of dimension  $n - 1$ .
- The graph may be considered a simplicial complex, with nodes as 0-simplices and edges as 1-simplices. For example, for indices  $0 \leq i \leq 1$  the simplicial complex of the graph consists of all nodes and edges in the graph.
- For indices  $0 \leq i \leq 1$ , as described above, the corresponding simplicial complex is the simplicial complex consisting of all cliques of size  $0 \leq n \leq 1$ . This is because both nodes and edges are cliques.

The remainder of the procedure is described below in broad strokes — a rigorous mathematical exposition using computational topology terminology is given by Myers [31] (Figure 19).

- A succession of graphs, considered as simplicial complexes, are considered for each index. For  $i = 0$ , the corresponding simplicial complex consists only of nodes, and for  $i = 1$ , the corresponding simplicial complex consists of both nodes and edges. Note that the simplicial complex for  $i = 0$  is a subset of the simplicial complex for  $i = 1$  and more generally, simplicial complexes of a lower index are subsets of simplicial complexes of a higher index.
- For each index  $i = m$ , edges are added to the simplicial complex based on shortest paths: for each node  $N$ , all nodes associated with node  $N$  at index  $i = m$  as described above are joined to node  $N$ . As  $i$  increases, an increasing number of cliques appear in the graph, and these are added to the simplicial complex as higher dimensional simplices.
- All birth times in the persistence diagram of the clique rank filtration occur at index  $i = 1$ . This is because, for the clique rank filtration, holes correspond to graph cycles and cycle consist of edges, which appear at  $i = 1$ .
- Death times in the persistence diagram of the clique rank filtration occur when a cycle is covered by a clique.

Although the problem of finding cliques in a graph is computationally intensive, in practice persistent homology and the persistence diagram may be computed directly from the shortest path distance matrix [31].

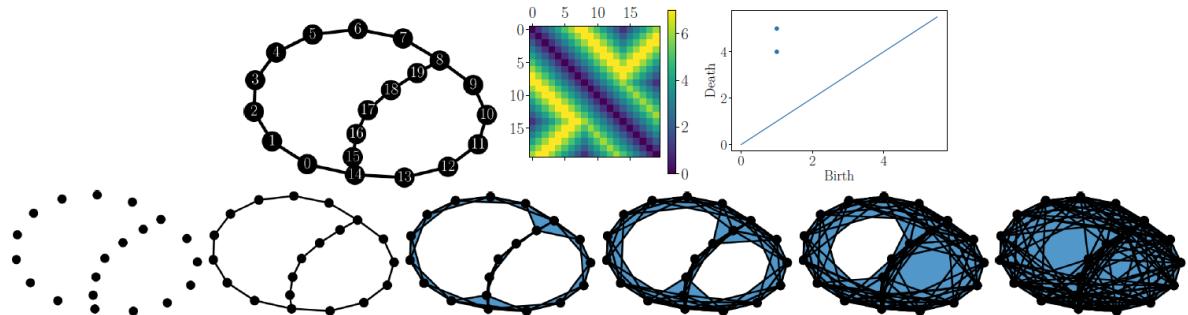


Figure 19: Illustrates the clique-rank filtration as described in Section 2.5.3. A graph (top left) and its associated shortest path distance matrix (top middle) is used to compute a clique rank filtration (bottom row, using indices  $i = 0$  through to  $i = 5$ , cliques are shaded in blue). The corresponding persistence diagram (top right) contains two points, each corresponding to a cycle in the graph.

Source: [31]

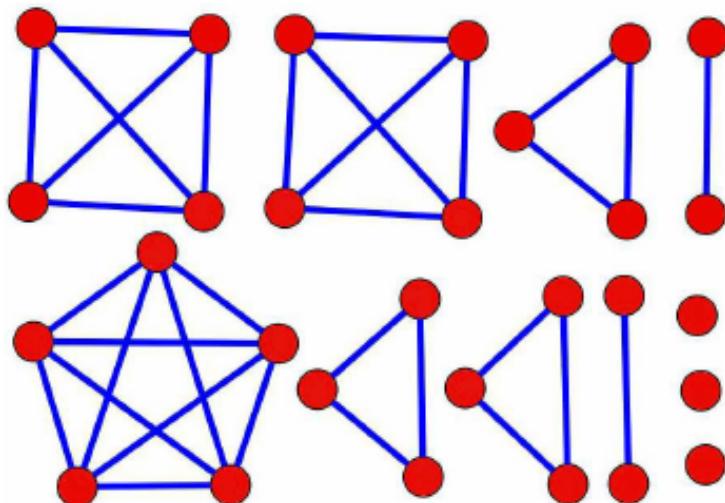


Figure 20: Illustrates cliques of various sizes.

Source: [32]



### 3 LITERATURE REVIEW

The phase coherence classifier is based on the following works

- Sliding windows and persistence: An application of Topological Methods to Signal Analysis (2015) [33]
- Persistent Homology of Complex Networks for Dynamic State Detection (2019) [31]
- Constructing Ordinal Partition Transition Networks from Multivariate Time Series (2017) [8]

They are each described, insofar as they relate to the project, in the following sections.

#### 3.1 Sliding windows and persistence: An Application of Topological Methods to Signal Analysis

Perea and Harer's widely cited work [33] is a milestone in the fertile field of time series analysis subfield of topological data analysis. It provides the mathematical framework able to place a particular pipeline popular in the subfield described by Krim et al. as follows [26]:

- A time delay embedding is applied to a signal to produce point cloud data, before
- the point cloud is preprocessed and cleaned. Then
- A Vietoris-Rips filtration is applied directly to the point-cloud data to produce a persistence diagram.

Points on this persistence diagram correspond to holes in the point cloud, which correspond to periods in the original time series.

In this sense, the author of this thesis makes the observation that the analysis described by the pipeline may be considered the non-linear analogue of Fourier Analysis:

- Points in the frequency domain, which represent periods of the time series just as points in the persistence diagram, correspond to loops in the phase portrait, which in turn correspond to periods in the time series.
- As mentioned in Section 2.1.5, Takens' theorem may be considered an analogue of Nyquist's sampling theorem.

#### 3.2 Persistent Homology of Complex Networks for Dynamic State Detection

Myers et al. [31] investigate the properties of univariate time series by constructing graphs, and investigating the properties of the graphs through various scalar measures. In general, their methodology is such that chaotic time series produce disordered graphs while periodic time series produce ordered graphs with cycles corresponding to time series periods (Figure 21).

Referring to Figure 22, two methods are presented:

- A method similar to that described in Section 3.1 wherein a time delay embedding is used to ultimately produce a persistence diagram, from which point summaries may be obtained.

- A method based on the order  $\mathbf{n}$  ordinal rank symbolisation of the univariate time series, as described by Section 2.2.2 and Figure 9. Essentially, the method entails assigning a node to each of  $\mathbf{n}!$  symbols, and connecting nodes based on time successions (e.g. if symbol 1 follows symbol 2).

Only the latter method is relevant to the project.

### 3.3 Constructing Ordinal Partition Transition Networks from Multivariate Time Series

Zhang et al. [8] build on the computational biology work of authors such as Wilds et al. [34].

Referring to Figure 23, Wilds et al. describe a method by which a multivariate time series is able to be symbolized:

- Consider a multivariate dynamical system with variables  $\mathbf{x}_1(\mathbf{t})$ ,  $\mathbf{x}_2(\mathbf{t})$ , and  $\mathbf{x}_3(\mathbf{t})$ .
- For each sample  $\mathbf{t}_j$  in time, the sign of the derivative (e.g. based on the first order differences) is considered:  $\Delta_j = (\text{sign } \mathbf{x}'_1(\mathbf{t}_j), \text{sign } \mathbf{x}'_2(\mathbf{t}_j), \text{sign } \mathbf{x}'_3(\mathbf{t}_j))$ . Note that  $\Delta_j$  may take on  $2^3$  possible values. Each of these values is assigned a symbol  $\pi_i$  for  $0 \leq i \leq 8$  (e.g.  $\pi_1 = (+++)$ ).
- Hence the time series is able to be symbolized, with each sample at  $\mathbf{t}_j$  taking on one of  $2^3$  values.

In fact, for the special case of the Figure 23 system, Wilds et al. provide the following symbolization:

$$(+ - +) \rightarrow (- - +) \rightarrow (- - -) \rightarrow \dots \rightarrow (+ + +) \rightarrow (- + +) \rightarrow (- - +) \rightarrow \dots$$

In fact, for reasons described in detail in Section 2.1.4, each of the symbols corresponds to a region of the phase portrait partitioned by nullclines. This is illustrated for the case of the Rössler system in subfigure (C2) from Figure 24.

Zhang et al. extend the method, and the extension most relevant to the project is described below [8]:

- Each coordinate of  $\Delta_j = (\text{sign } \mathbf{x}'_1(\mathbf{t}_j), \text{sign } \mathbf{x}'_2(\mathbf{t}_j), \text{sign } \mathbf{x}'_3(\mathbf{t}_n))$  is substituted with an order  $\mathbf{n}$  ordinal rank symbolisation, as described by Section 2.2.2 and Figure 9.
- For the special case of  $\mathbf{n} = 2$ , the symbolization is as described by Wilds et al. [34] assuming  $\Delta_n$  was computed using first order differences. Hence, as expected, an  $\mathbf{N} = 3$  dimensional time series is able to be symbolized with each sample at  $\mathbf{t}_j$  taking on one of  $(2)^3$  values.
- In the general case, an  $\mathbf{N}$  dimensional time series is able to be symbolized with each sample at  $\mathbf{t}_j$  taking on one of  $(\mathbf{n}!)^{\mathbf{N}}$  values.

Note the following [8]:

- For  $\mathbf{n} > 2$  the interpretation described above of symbolization as a process of partitioning the state space no longer holds, since the symbols no longer may be interpreted as the sign of a derivative.
- Zhang et al. are able to use their methodology to measure phase coherence.

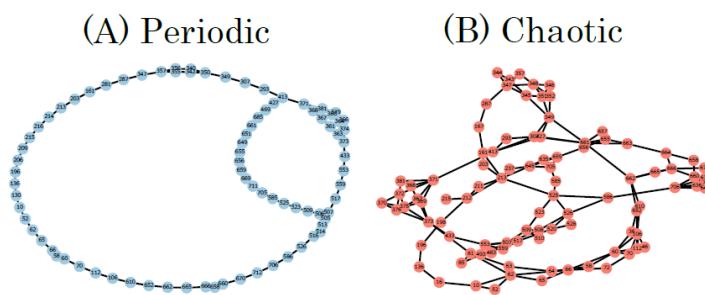


Figure 21: Illustrates the graphs that may be generated in subfigure (C) of the Figure 22 pipeline, as described in Section 3.2.

Source: [31]

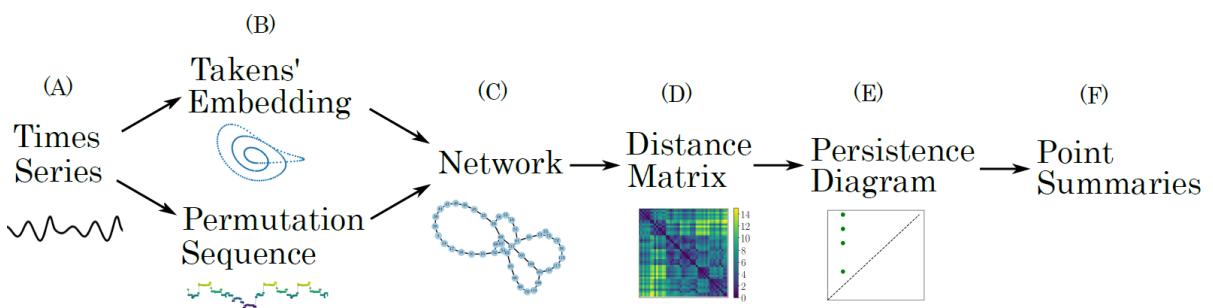


Figure 22: Illustrates the pipeline employed by Myers et al. as described in Section 3.2.

Source: [31]

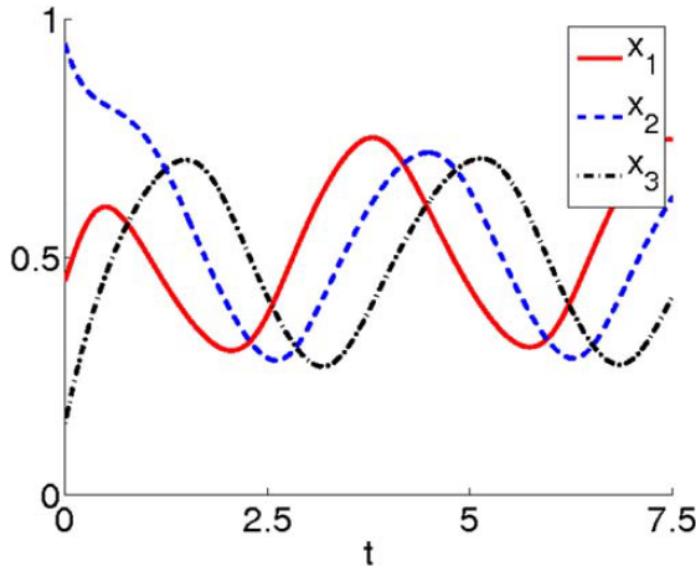


Figure 23: Plots a multivariate dynamical system.

Source: Adapted from [34]



# 4 PROJECT OVERVIEW

The methodology for computing the classifier is introduced in Section 4.1.1 before the scope and objectives of the project are described. A detailed account of the methodology for computing the classifier, however, is deferred till Section 5.

## 4.1 Methodology Introduction

### 4.1.1 Project Objectives

The objective of the project is to implement and evaluate a method of measuring phase coherence. In fact, a hybrid of the methods of Myers et al. [31] and Zhang et al. [8] – respectively described in Section 3.2 and Section 3.3 – is used. Specifically, the Figure 22 pipeline of Myers et al. is used, albeit with three modifications:

- The point summaries of subfigure (f) are not produced or considered.
- A multivariate time series rather than the univariate time series of subfigure (a) are used.
- The permutation sequence method of subfigure (b) is used, albeit using the multivariate symbolization method of Zhang et al. [8] substituted for the univariate symbolization method of Myers et al. [31].

### 4.1.2 Methodology Rationale and Motivation

The relevance and importance of the project was described in Section 1.2.2. Here the choice of methodology is motivated:

- EEG data is inherently noisy due to artefacts (e.g. eye blink artefacts, ECG artefacts, etc) [35]. Hence the methodology must be robust to noise.
- As alluded to in Section 2.2.2, methods based on symbolization (e.g. permutation entropy) tend to be more robust to noise.
- As alluded to in Section 2.5, methods based on Topological Data Analysis tend to be robust to noise.
- Referring to Figure 22, Myers et al. [31] found that the time delay embedding method was not effective when applied to EEG data. They reasoned that this was a result of the noise inherent in EEG data. Hence a method based on symbolization – as opposed to the class of time delay embedding methods described in Section 3.1 – was used.

### 4.1.3 Data Sets

The classifier is evaluated using simulated Rössler system data, as described in detail in Section 6. Further, the expert annotated neonatal seizure EEG dataset of Stevenson et al. [36] released in 2019 is used to evaluate the classifier.

## 4.2 Project Scope

The scope of the project is to evaluate whether the classifier effectively measures phase coherence – that is, to evaluate whether the classifier measures phase coherence, as opposed to some other property of multivariate datasets. Section 6 describes how this is done.

# 5 COMPUTING THE CLASSIFIER

The computation of the phase coherence classifier is described below. A detailed explanation of the concepts found in this section were described in Section 2 and Section 3. The methodology was introduced in Section 4, but is expounded on in the following sections, each of which refer to Figure 24.

## 5.1 Time Series Symbolization

Given multivariate time series data (A1), each of the  $N$  constituent series (A2) are symbolized (B1) using a sliding window into  $n!$  symbols based on ordinal rank as illustrated for  $n = 3$  (B2). At each point in time in the symbolized time series (B1) there are  $(n!)^N$  possible states, each of which can be given a symbol to produce a univariate symbolized time series as illustrated for  $n = 2$ ,  $N = 3$ , and  $(n!)^N = 8$  (C1).

## 5.2 Periodicity in Symbolized Data

A Rössler non-linear time series ( $N = 3$ ) symbolized with  $n = 2$  partitions the state space into  $(n!)^N = 8$  symbols (C2). In the phase coherent case, interdependence between variables and periodic time series patterns generate state space orbits that oscillate around an annulus (C2), corresponding to a periodic symbolic sequence (C1). The relationship between phase coherence and periodic symbolic sequences holds in general for  $n > 2$ , despite the fact that the state space cannot be interpreted as being partitioned by symbols in this case.

## 5.3 Ordinal Partition Transition Networks

An ordinal partition transition network is a graph representation (D1) of a symbolic sequence (C1) with nodes as symbols, edges as time successions, and cycles as state space orbits (C2). Where in Fourier analysis periodic time patterns correspond to frequency domain points, in our analysis they correspond to cycles in the graph of the symbolized data.

## 5.4 Computing Persistent Homology

The number of graph cycles and the quantity of nodes in each cycle is estimated through persistent homology: A clique rank filtration (D2) is performed by successively adding edges to the graph in an ordered determined by a shortest path distance matrix (E1). A persistence diagram (F1) is produced by recording when holes are created, and when they are covered by cliques, which are shaded below in blue (D2). The differences between hole death and birth times in the persistence diagram (F1) points are the computed classifier.

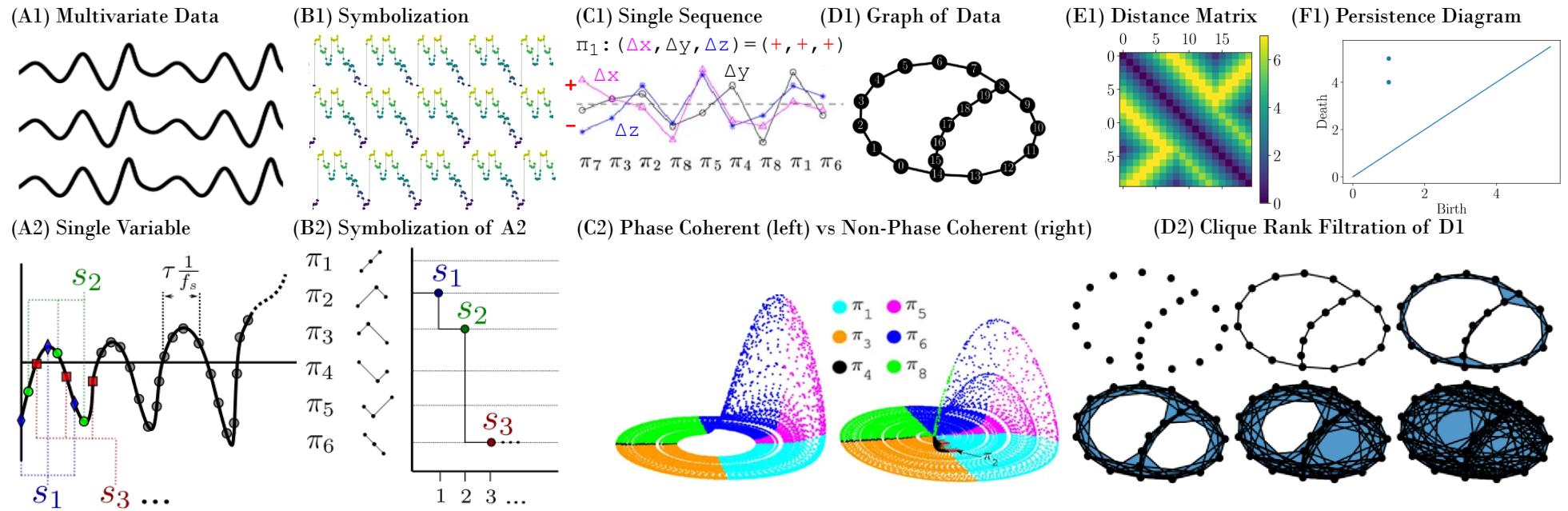


Figure 24: Illustrates the classifier computation methodology. The top row describes from left to right the classifier computation pipeline, while the bottom row provides supplementary explanatory illustrations.

Source: Adapted from [8] and [31]

## 5.5 Why Should the Classifier Measure Phase Coherence?

### 5.5.1 A Heuristic Argument

Section 3.3 described the multivariate symbolization procedure of Zhang et al. [8]. At the end of this section it was noted that for  $n > 2$  the interpretation described of symbolization as a process of partitioning the state space no longer holds. **In the following heuristic argument the mentioned interpretation is nonetheless applied, and that the behaviour whereby loops in the phase portrait correspond to a periodic succession of symbols still holds.** Currently a mathematical framework to justify this does not exist, and hence this assumption must be considered an example of reasoning by analogy. The heuristic argument is now given below.

When a subset of variables from a multivariate time series has channels that are phase synchronized, each variable from the subset should exhibit a periodic pattern, so that the phase portrait of that subset exhibits loops. These loops produce cycles in the graph that is produced by the methodology, which produces points on the persistence diagram. **Conversely, points on the persistence diagram correspond to cycles in the graph, which correspond to the simultaneously occurring periodic patterns in a subset of variables in the time series.**

### 5.5.2 The Robustness of the Algorithm

Errors in the symbolization of course occur due to noise, and the suboptimal selection of parameters (e.g. if the number of variables selected to produce a graph are not mutually phase coherent). Ideally the algorithm is robust to these errors, but an evaluation of this robustness is beyond the scope of this thesis. In fact, devising a

- theoretical framework for the algorithm within which the robustness of the algorithm can be understood, and a
- method to evaluate the robustness of the algorithm

is considered by the author to pose significant challenge.



# 6 EVALUATING THE CLASSIFIER

The methodology for evaluating the classifier and the results obtained are given below.

## 6.1 Introduction

### 6.1.1 Overview

Breakspear and Terry's [9] statistical methods were applied to phase coherent and non-phase coherent Rössler systems in addition to seizure epochs, which are known to be phase coherent [9], from artefact (e.g. eye blinks and heartbeats) corrected bandpass filtered downsampled EEG recordings. They write:

[Artefacts] such as ...sampling error [and] filtering can lead to the false identification of non-linearity ...In order to ensure that [classifier values] are not artefacts caused by these properties of the data ...surrogate data ...was employed. [Surrogate data computation involves] preserving all the original linear properties but destroying any non-linear structure present by phase randomising the Fourier components. This allows testing of the null hypothesis that the time series are exclusively linear, with purely linear coherence.

### 6.1.2 Computation of Surrogate Data

Breakspear and Terry summarize the computation of surrogate data according to their method as follows [9]:

Stochastic processes and low-dimensional non-linear systems both produce irregular, broad spectrum signals. Stochastic signals can be fully described by linear measures, such as the spectral density function. Non-linear signals have additional structure which is not captured by the spectral density function. This structure derives from the geometry of the non-linear systems' phase space orbits ...

Surrogate data was constructed from the experimental data by a combination of the amplitude-adjusted and multivariate surrogate data techniques. Given a set of **K** concurrent series of data each of length **L**, the algorithm produces a set of **K** 'surrogate data' series each of length **L**, preserving the amplitude distribution, spectral density and cross-spectral density functions of the original data. However, non-linear structure contained within and between the time series are destroyed. Thus the surrogate algorithm allows testing of the null hypothesis that the time series are produced by a cross-correlated stochastic system. The amplitude-adjusted step controls for any static non-linear distortion of the signal due the measurement process.

In their appendices, they provide a full description of their method, which was implemented using the **Python 3** code on the following page.

```

class Surrogate:
# http://www.pik-potsdam.de/~donges/pyunicorn/_modules/
# pyunicorn/timeseries/surrogates.html
def __call__(self, data):
    # Amplitude adjusted (fourier transform)
    # phase randomized surrogate, as described in
    # Breakspear (2002)
    gaussian = random.randn(data.shape[0], data.shape[1])
    gaussian.sort(axis=1)
    # Rescale data to Gaussian distribution
    rescaledData = np.zeros(data.shape)
    for i in range(data.shape[0]):
        rescaledData[i, :] = \
            gaussian[i, self._get_ranks(data)[i, :]]
    # Phase randomize rescaled data
    phaseRandomizedData = \
        self._phase_shuffled_surrogate(rescaledData)
    # Rescale back to amplitude distribution of original data
    sortedOriginal = data.copy()
    sortedOriginal.sort(axis=1)
    for i in range(data.shape[0]):
        rescaledData[i, :] = sortedOriginal[
            i,
            self._get_ranks(phaseRandomizedData)[i, :]]
    return rescaledData
def _get_ranks(self, data):
    return data.argsort(axis=1).argsort(axis=1)
def _phase_shuffled_surrogate(self, data):
    (N, nTime) = data.shape
    # Calculate FFT of original_data time series
    surrogates = np.fft.rfft(data, axis=1)
    lenPhase = surrogates.shape[1]
    # Generate random phases uniformly distributed
    # in the interval [0, 2*Pi]
    phases = np.random.uniform(
        low=0, high=2 * np.pi,
        size=(N, lenPhase))
    # Add random phases uniformly distributed
    # in the interval [0, 2*Pi]
    surrogates *= np.exp(1j * phases)
    # Calculate IFFT and take the real part, the
    # remaining imaginary part is due to numerical errors.
    return np.ascontiguousarray(np.real(np.fft.irfft(
        surrogates, n=nTime, axis=1)))

```

### 6.1.3 Computation of Statistical Significance

Following Breakspear and Terry [9], the surrogate data described above is used to compute statistical significance using a one tailed  $z$ -test, where

$$z = \frac{\mu_{\text{surrogate}} - \mu_{\text{experimental}}}{\sigma_{\text{surrogate}}}$$

where  $\mu$  and  $\sigma$  designate mean and standard deviation respectively. Specifically, as will be elaborated on, for given experimental data

- The classifier is computed and used to compute to calculate  $\mu_{\text{experimental}}$ .
- Multiple realizations of the surrogate data, which are generated through a stochastic process, are produced.
- The classifier is computed for each surrogate data realization, and the resulting surrogate classifier values are used to compute  $\mu_{\text{surrogate}}$  and  $\sigma_{\text{surrogate}}$ .
- The  $z$ -value is computed, which can then be used to produce a  $p$ -value.

Breakspear and Terry [9] comment that

The  $p$ -values, were ... obtained via a one-tailed  $z$ -test. These represent the probability that the experimental measures would be observed by chance alone if the null hypothesis was true. A one-tailed test was chosen since we are interested in rejecting the null hypothesis only when the experimental prediction is better than that calculated from the surrogate data.

### 6.1.4 Evaluation of Statistical Significance

Following Breakspear and Terry [9], note that the surrogate data procedure used is stochastic, so there exists the potential for the null hypothesis to be rejected purely by chance — that is, by coincidental false positives produced by the surrogate data generation algorithm. Hence the typical condition of  $p < \alpha = 0.05$  for rejecting the null hypothesis must be strengthened [9]. Breakspear and Terry use Šidák correction for this purpose:

$$p_{\text{corrected}} = 1 - (1 - \alpha)^{1/n}$$

where  $n$  is the number of repeated observations. For example if  $\alpha = 0.05$  and  $n_1 = 19$  sets of surrogate data values are paired with  $n_2 = 1$  experimental values, then  $n = n_1 + n_2 = 20$  and the condition for rejecting the null hypothesis is

$$p < 1 - (1 - 0.05)^{1/20} < \alpha = 0.05.$$

## 6.2 Rössler System Data

The parameters used by Zhang et al. [8] are used to simulate the Rössler system using a fourth order Runge-Kutta numerical integration with step size 0.01:

$$\dot{x} = -y - z \quad \dot{y} = x + ay \quad \dot{z} = 0.4 + z(x - 8.5)$$

Following Zhang et al. [8]

- To discard the data points corresponding to the initial transient behaviour and retain only the steady state values, the first 16000 of a total 20000 steps are discarded; used a similar processing step.
- The data was simulated to produce phase coherent ( $a = 0.165$ , Figure 25) and non-phase coherent ( $a = 0.26$ , Figure 30) Rössler systems.

For the case of the phase coherent system, examples of surrogate data realizations are given by Figure 26 and Figure 27.

Recall from Section 2.5.3 that all birth times produced in persistence diagrams produced by a clique rank filtration. This means, since the classifier indeed employs a clique rank filtration, that the total persistence (the difference between birth times and death times) is able to be expressed on a single axis. Hence it is possible, **using a 120 data point non-overlapping sliding window** with our classifier, to produce plots of persistence with respect to time. **Noting that an order  $n = 5$  symbolization was used to compute all classifier values for the Rössler system**, examples of such plots are given by Figure 28, Figure 29, and Figure 31 – for which the following observations may be made:

- The quantity of points and their variance is greater for diagrams produced by surrogate data than for diagrams produced by experimental data
- The quantity of points, their variance is greater for diagrams produced by non-phase coherent data than for diagrams produced by phase coherent data.

Such diagrams were produced across a range of  $a$  parameter values, and the mean & standard deviation (Figure 32) and median, maximum, & minimum (Figure 33) were computed for the data points across all time values. Figure 32 and Figure 33 summarize these statistics, and the following observations may be made for them. The dotted line indicates the boundary – as determined by Zhang et al. [8] – between phase coherent parameter values ( $a < 0.206$ ) and non-phase coherent parameter values ( $a > 0.206$ ). After this boundary value

- the mean decreases and variance increases, and
- the median decreases and the minimum decreases.

To confirm that phase coherent and non-phase coherent classifier values may be considered to be drawn from distinct distributions an exact permutation test was performed using the `permutation_test` function from the `mlxtend` library for **Python 3**. From the means from Figure 32 were divided two datasets were extracted: Ten means prior to  $a = 0.206$  and ten means after. Further data points were not used for reasons to ensure that the exact permutation test was computationally tractable. A  $p$ -value of  $1.29 \times 10^{-4}$  was obtained, and – applying the same procedure Figure 32 standard deviations –  $p = 2.16 \times 10^{-5}$  was obtained for the standard deviations. **Hence it may be concluded that phase coherent and non-phase coherent classifier values may be considered to be drawn from distinct distributions.**

To confirm that phase coherence, as opposed to some other property of the data, was being measured, the surrogate data procedure described in Section 6.1.2 was applied:

- $m = 99$  realizations of the seizure data were used, so that given a base significance threshold of  $\alpha = 0.05$ , the Šidák corrected value is  $p_{\text{corrected}} = 5.128014 \times 10^{-4}$
- For a phase coherent Rössler system ( $a = 0.165$ ),  $p = 5.659 \times 10^{-41} < p_{\text{corrected}}$  was obtained, firmly rejecting the null hypothesis.
- Similarly, for a non-phase coherent Rössler system ( $a = 0.26$ ), the null hypothesis was rejected but to a lesser degree:  $p = 2.584 \times 10^{-24} < p_{\text{corrected}}$ .

Recall that the null hypothesis is not that phase coherence is *not* being measured, but rather that "the time series are produced by a cross-correlated stochastic system" and that "the time series [is] exclusively linear, with purely linear coherence" [9]. **Hence the results provide further confidence that phase coherence as opposed to some other property of the data is being measured by the mean classifier value.** Note that

- the fact that a lesser  $p$  value was obtained for the non-coherent provides further confidence that phase coherence is being measured,
- although the conclusions here drawn may not answer the binary question of whether phase coherence is being measured or not, this is a limitation of surrogate data for phase coherence, and other possible surrogate data methods can only reproduce similar results [6].

The same procedure was repeated, albeit substituting classifier standard deviation for classifier means, but the null hypothesis was rejected  $p \approx 1$ . **Hence it may be concluded that the standard deviation of the classifier value is a result of linear properties of the data, and may not be employed to measure phase coherence.**

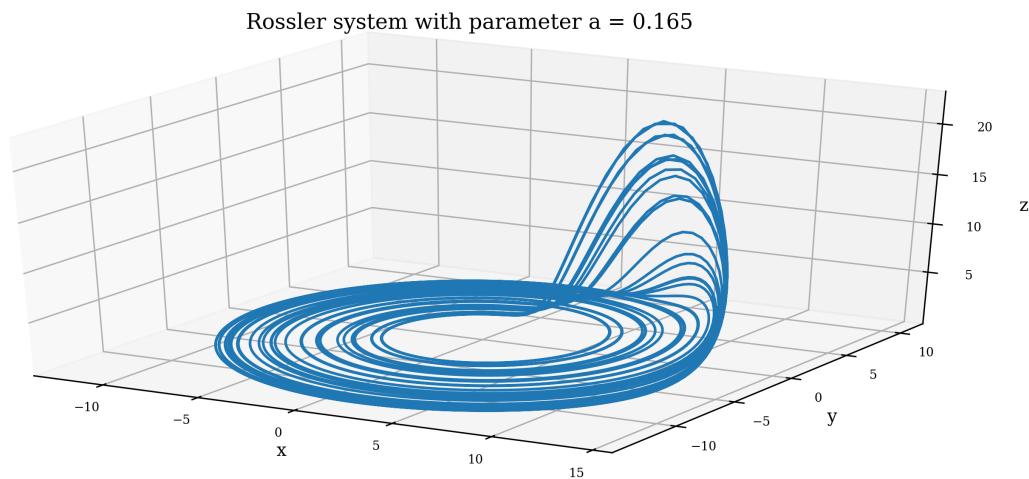


Figure 25: A phase coherent Rössler system as described in Section 6.2.

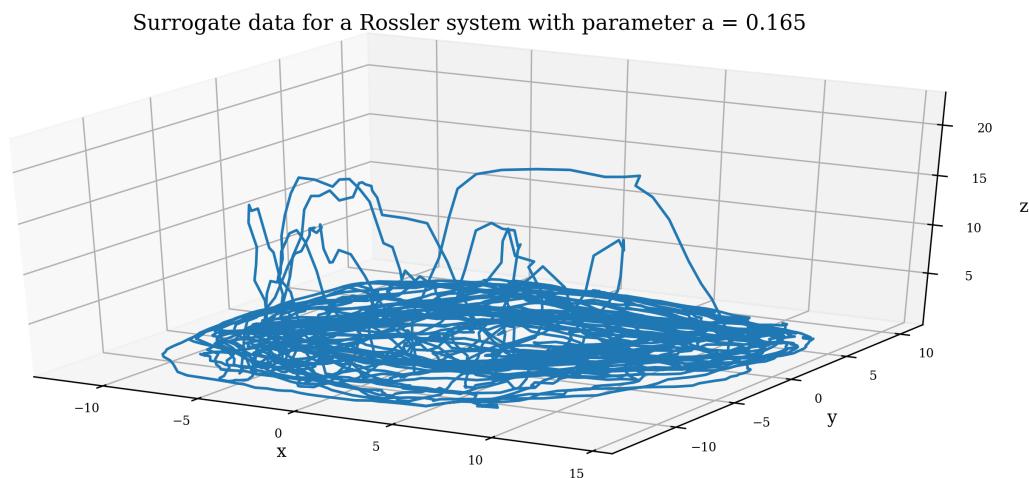


Figure 26: A realization of surrogate data for the Figure 25 phase coherent Rössler system, as described in Section 6.2.

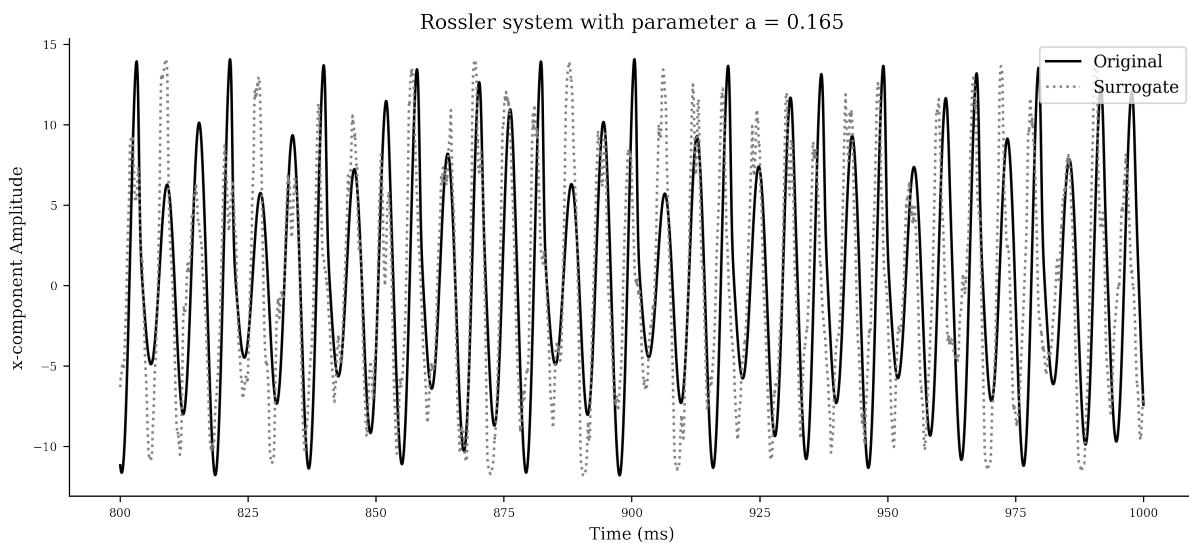


Figure 27: Plots in the time domain the **x**-component of the Figure 25 phase coherent Rössler system and the **x**-component of the Figure 25 surrogate data.

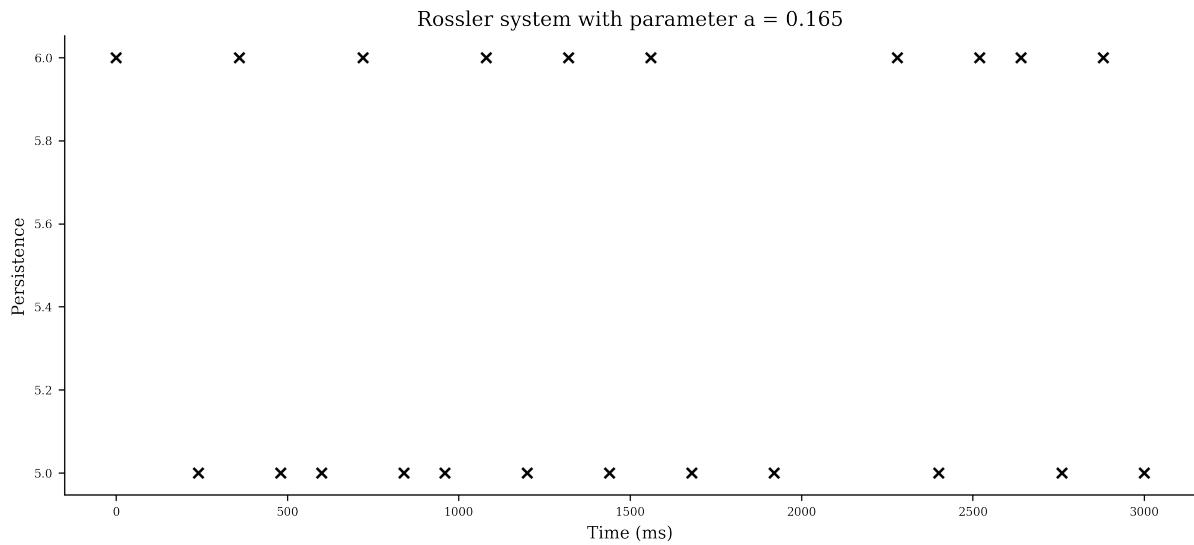


Figure 28: The persistence of the Figure 25 phase coherent Rössler system, as described in Section 6.2.

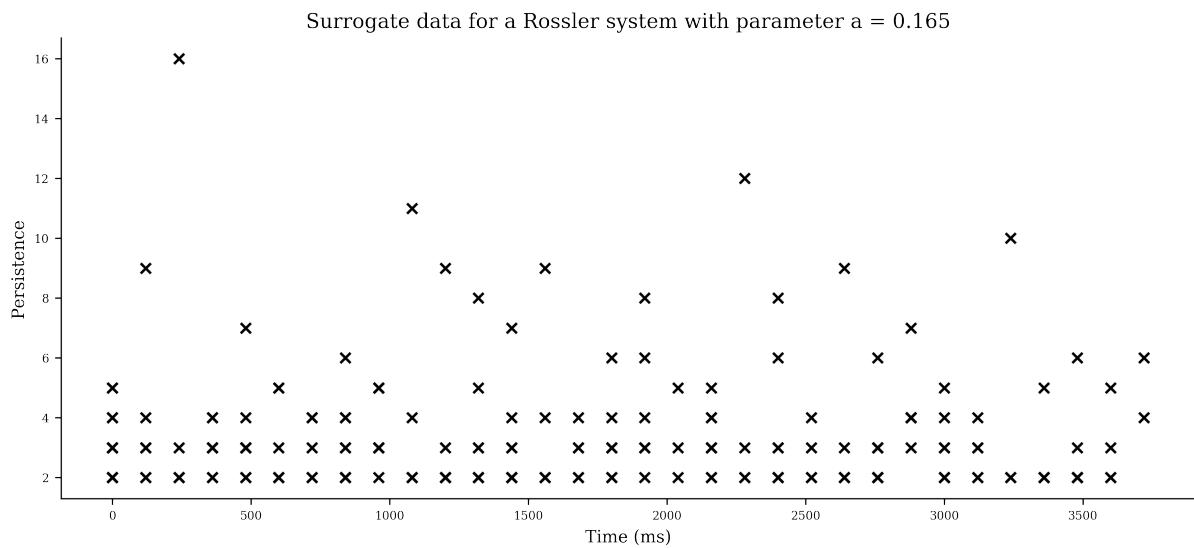


Figure 29: The persistence of the Figure 26 surrogate data, as described in Section 6.2.

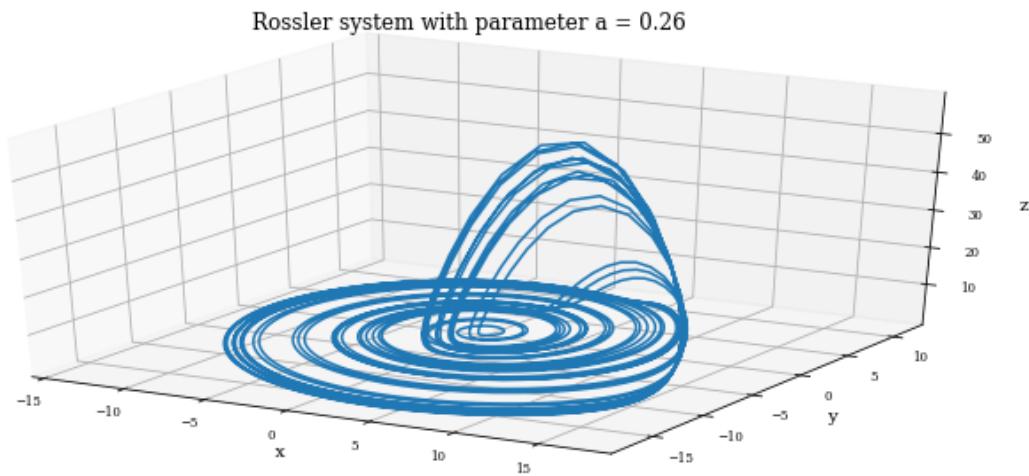


Figure 30: A non-phase coherent Rössler system as described in Section 6.2.

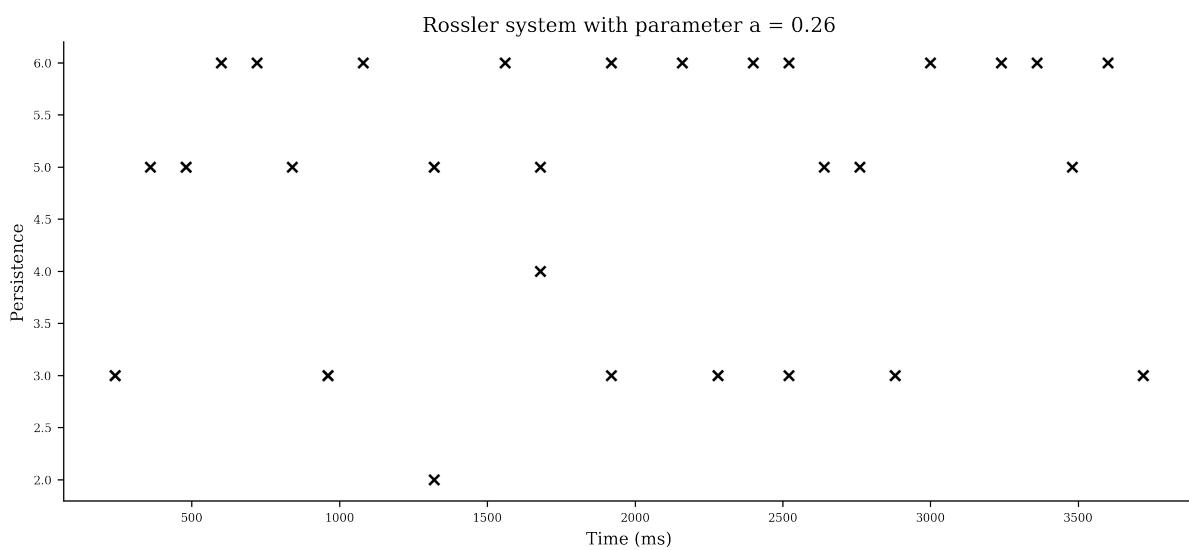


Figure 31: The persistence of the Figure 30 non-phase coherent Rössler system, as described in Section 6.2.

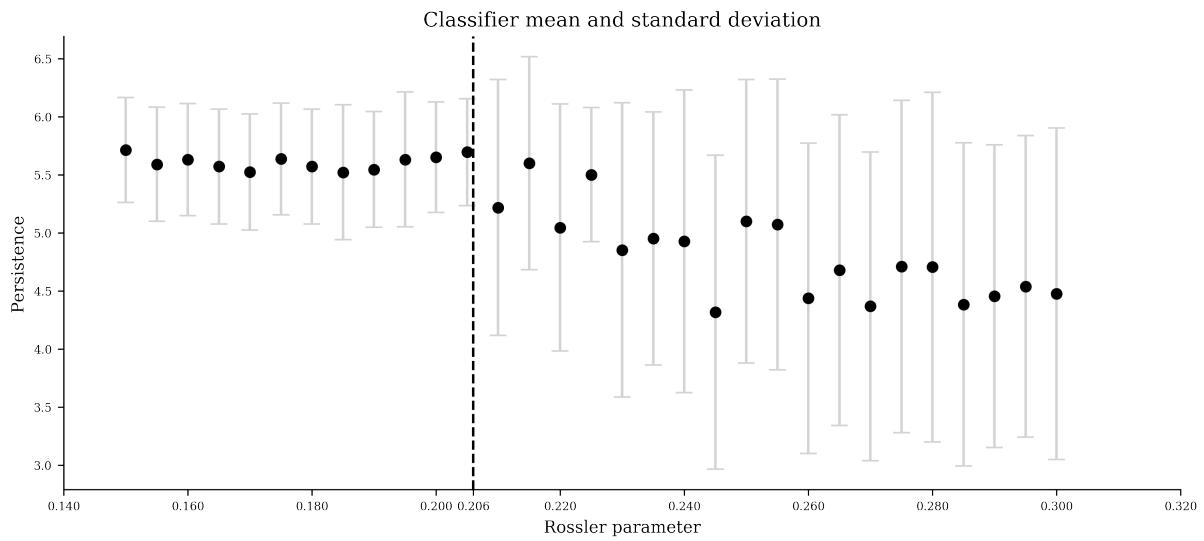


Figure 32: The classifier means and standard deviations of the Section 6 Rössler system for various parameter values. The dotted line indicates the boundary between phase coherent parameter values ( $\alpha < 0.206$ ) and non-phase coherent parameter values ( $\alpha > 0.206$ ).

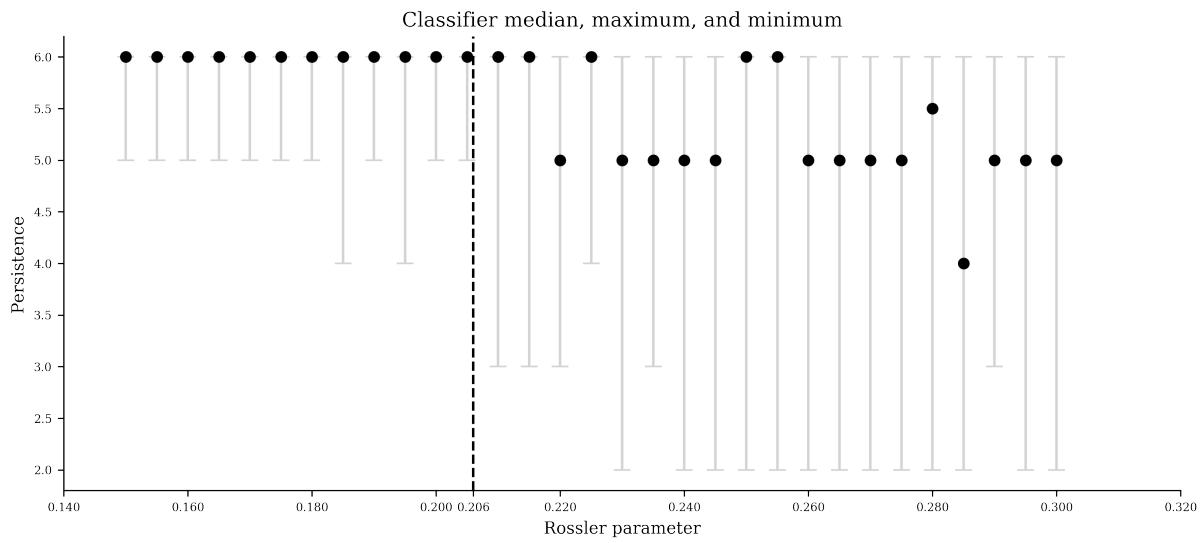


Figure 33: The classifier medians, maximums and minimums of the Section 6 Rössler system for various parameter values. The dotted line indicates the boundary between phase coherent parameter values ( $\alpha < 0.206$ ) and non-phase coherent parameter values ( $\alpha > 0.206$ ).

### 6.3 Seizure EEG Data

The seizure data was preprocessed as follows

- A bandpass filter between **0.5** Hz and **34** Hz was applied, and the signal was downsampled to **85** Hz; a similar preprocessing procedure was used by Faul et al. [37].
- Artefacts (e.g. eye blinks, muscle Artefacts, ECG) were removed using Independent Component Analysis and template matching.
- Each second of the seizure data was annotated as a seizure (1) or a non-seizure (0) window by three experts. The three sets of annotations were combined using a Kalman Smoother into a probability-like value  $0 \leq r \leq 1$  which was then thresholded so that  $r > 0.8$  was considered a seizure second while  $r \leq 0.8$  was considered a non-seizure second.

To evaluate the classifier the classifier was applied to 2 second sliding windows with 1.5 seconds overlap. These parameters were based both on trial-and-error and on the suggestion of Faul et al. [37] that the window be large enough so that the classifier may be effective, but small enough so that the signal is approximately stationary within the window.

Since, during a seizure event – especially focal seizure events – not all electrodes possess seizure signal properties [37] a channel selection procedure was employed. The permutation entropy similar to that of Mammone et al. [13] was used: For each sliding window the permutation entropy of each channel is computed so that the channels may be ranked and the **N** most seizure-like (i.e. periodic) channels can be selected.

Hence to evaluate the classifier, three parameters must be selected:

- The symbolization order **n**
- The number of channels **N** – that is, the number of variables in the multivariate time series

Referring to Figures 34 to 39, which all depict the application of the procedure described to the same recording for varying values of **N** and **n**, certain observations may be made:

- The classifier values of the figures may be interpreted similarly to the interpretation of the persistence diagrams with respect to time presented in Section 6.2
- For larger values of **N**, smaller values of **n** more effectively distinguish seizure from non-seizure classifier values, and
- both **n** and **N** must be small values lest the classifier computation become computationally intractable ( $(n!)^N$  graph nodes are produced per sliding window while computing the classifier),
- Non-seizure classifier values tend to dip below seizure classifier values. **This suggests that the classifier may in future be applied to the problem of seizure detection.**

To confirm that phase coherence, as opposed to some other property of the data, was being measured, the surrogate data procedure described in Section 6.1.2 and Section 6.2, was applied using a methodology similar to that of Breakspear and Terry [9]:

- The **z**-test was applied to seizure epochs (i.e. two second windows annotated by experts to contain a seizure vent).
- For each seizure epoch,  $m = 19$  realizations of surrogate data were used (Figure 40, Figure 41).
- 40 seizure epoch randomly selected from each of 3 patients.
- A base significance of  $\alpha = 0.05$  was used for the Šidák correction

$$p_{\text{corrected}} = 1 - (1 - \alpha)^{1/n}$$

as was done in Section 6.2. For each epoch  $n = m + 1 = 20$ , so that on the level of a single patient  $n = 40(m + 1)$  and on the level of all patients  $n = 3 \times 40(m + 1)$ .

The **z**-test null hypothesis was rejected for 22%, 16%, and 14% of all epochs for Šidák corrections on the level of an epoch, a patient, and all patients respectively. These percentages are significantly higher than the 5% value that would be expected by chance alone [4]. **Hence the results provide further confidence that phase coherence as opposed to some other property of the data is being measured by the mean classifier value.**

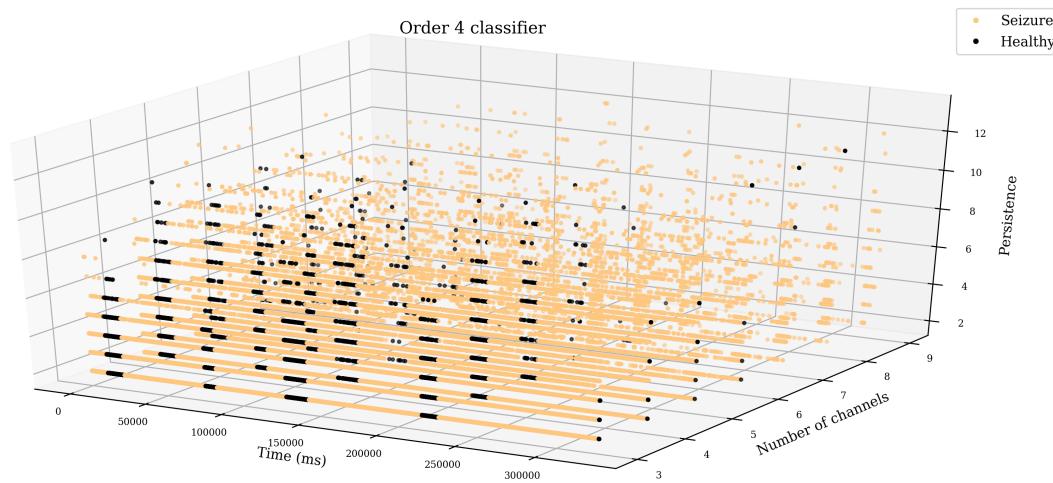


Figure 34: The classifier values with respect to time, for various values of  $N$ , where  $n = 4$ , as described in Section 6.3.

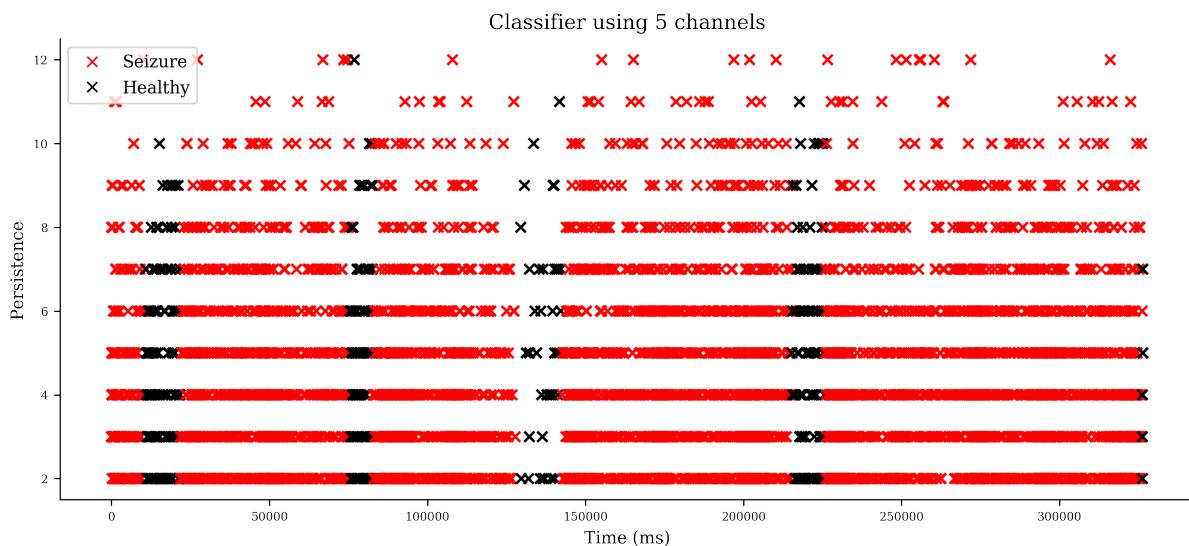


Figure 35: The classifier values with respect to time, for  $N = 5$ , where  $n = 4$ , as described in Section 6.3.

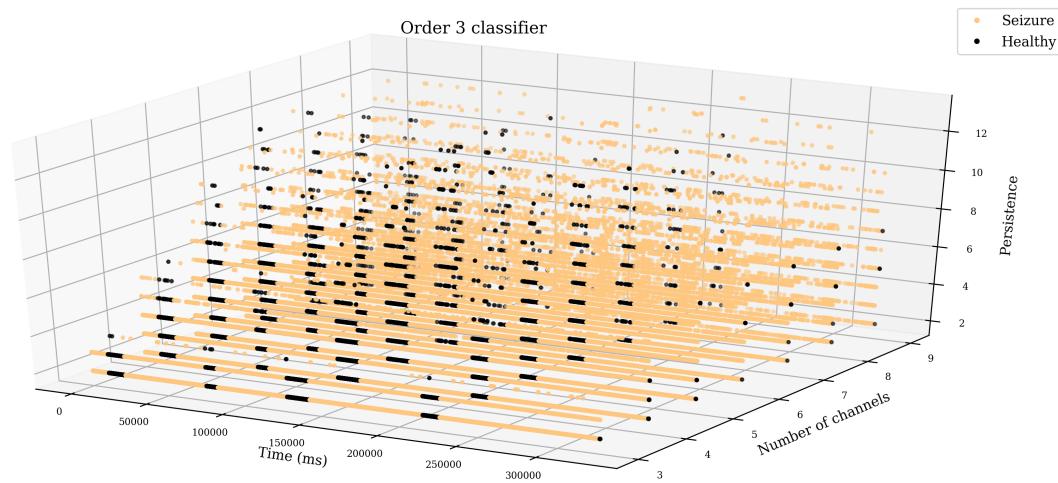


Figure 36: The classifier values with respect to time, for various values of  $N$ , where  $n = 3$ , as described in Section 6.3.

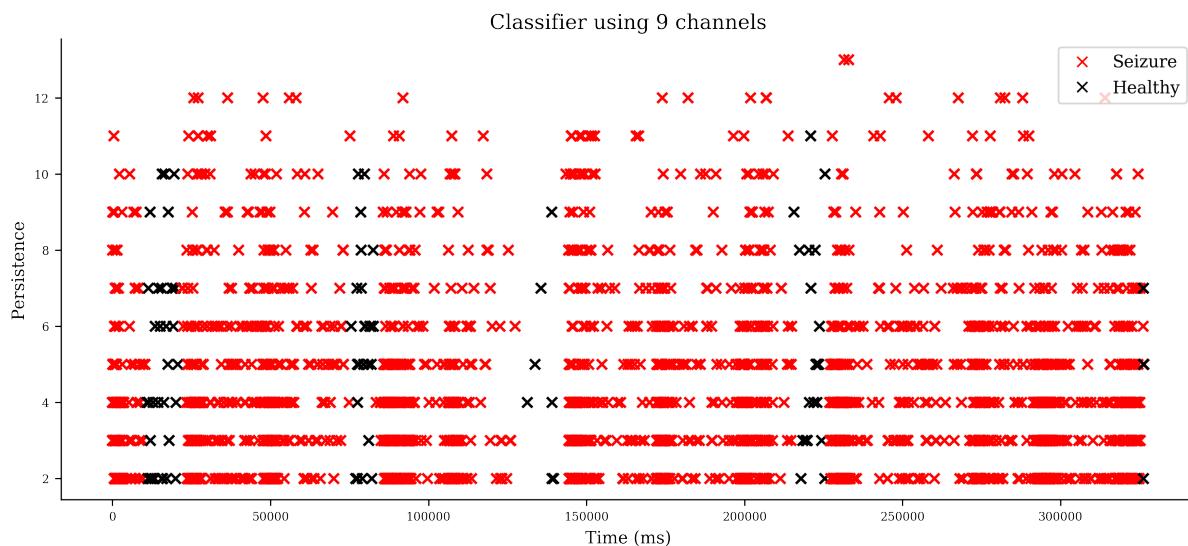


Figure 37: The classifier values with respect to time, for  $N = 3$ , where  $n = 9$ , as described in Section 6.3.

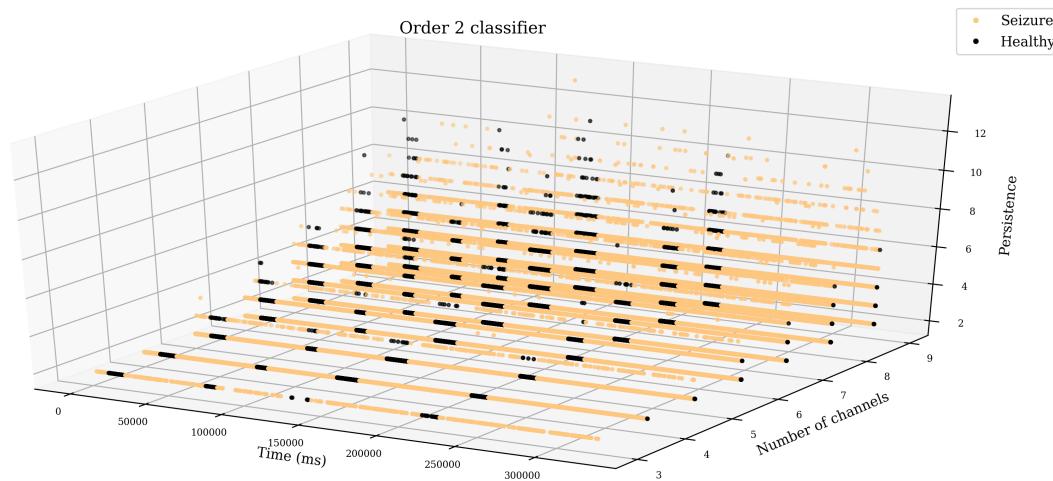


Figure 38: The classifier values with respect to time, for various values of  $N$ , where  $n = 2$ , as described in Section 6.3.

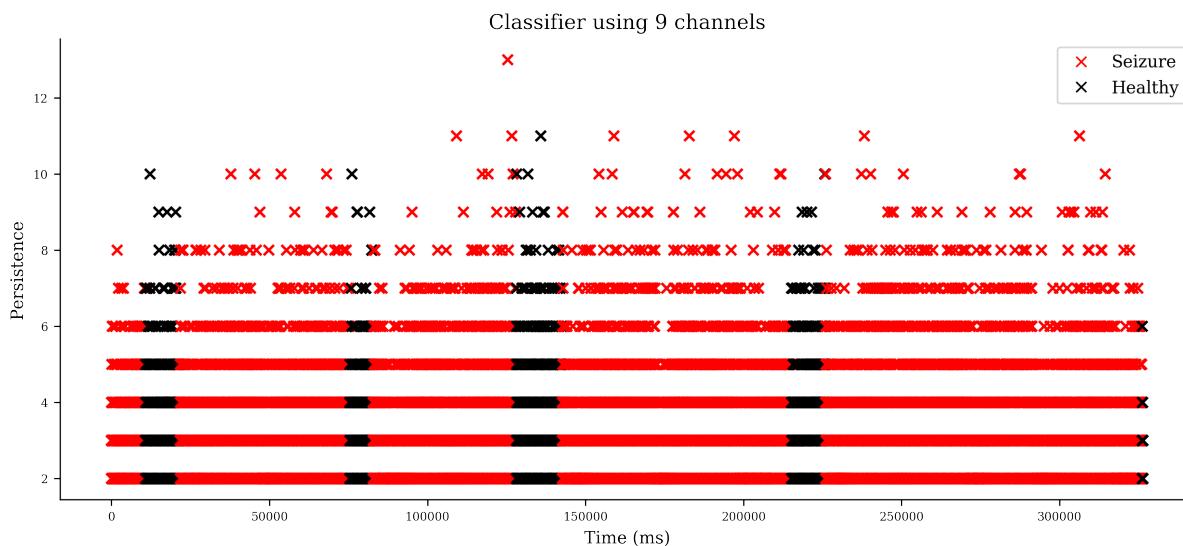


Figure 39: The classifier values with respect to time, for  $N = 9$ , where  $n = 2$ , as described in Section 6.3.

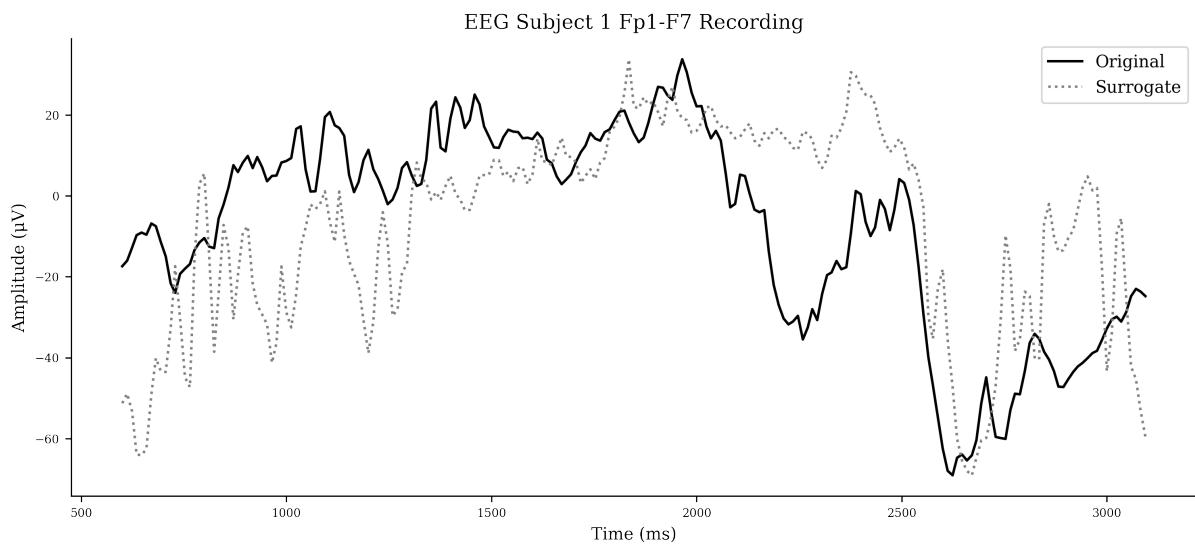


Figure 40: An example on of a seizure epoch from a single EEG channel together with a realization of the corresponding surrogate data. The patient from which the depicted data was collected differs from the patient of Figure 41

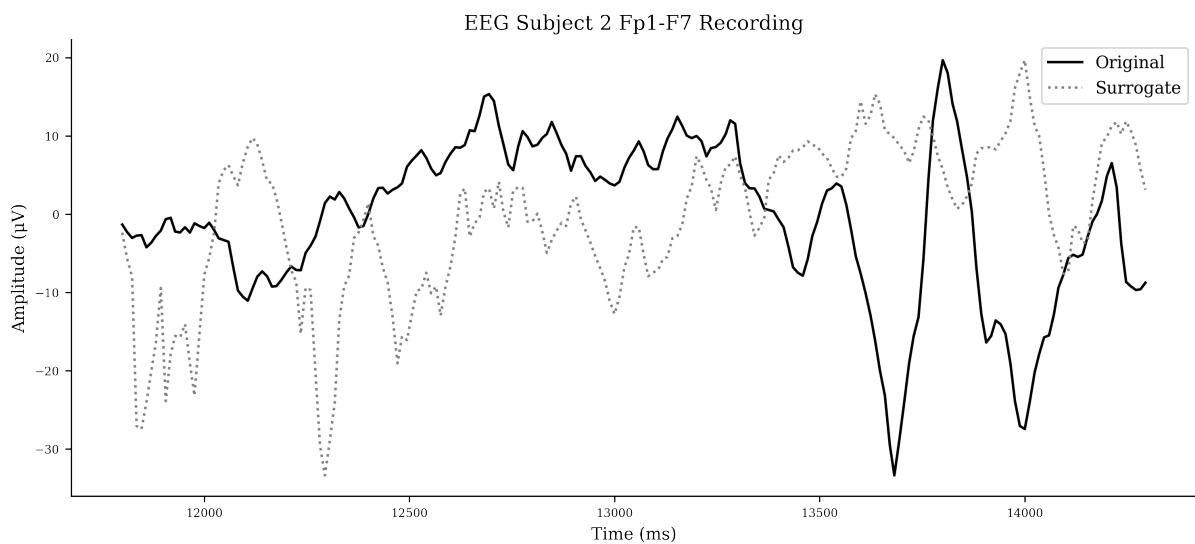


Figure 41: An example on of a seizure epoch from a single EEG channel together with a realization of the corresponding surrogate data. The patient from which the depicted data was collected differs from the patient of Figure 40



# 7 CONCLUSIONS AND REFLECTIONS

The conclusions drawn from Section 6 are described in Section 7.1. Then the project plan is evaluated by answering the questions

- Were the project objectives met?
- What aspects of the project were successful?
- What aspects of the project could have been improved?

in Section 7.2. Finally, suggestions for how the project may be extended are given in Section 7.3.

## 7.1 Conclusions

- Phase coherent and non-phase coherent classifier values may be considered to be drawn from distinct distributions (Section 6.2).
- The results provide further confidence that phase coherence as opposed to some other property of the data is being measured by the mean of the classifier (Section 6.2, Section 6.3).
- The standard deviation of the classifier value is a result of linear properties of the data, and may not be employed to measure phase coherence (Section 6.2)

## 7.2 Critical Appraisal of Project

The project undertaken was exploratory and research-based in nature: It involved the development of a novel method through trial and error in the face of uncertainty, as opposed to the implementation of an established method. Hence the fact that the method worked at all is a significant achievement.

In fact, as described in Section 4.1.1, the method

- draws most heavily from the paper of Myers et al. [31] released this year (2019),
- is a hybrid of the methods of Myers et al. [31] and Zhang et al. [8], and
- is able to measure phase coherence, as indicated by the results.

Hence the project achieves its objectives and falls within its defined scope.

The project could have been improved through the use of a larger number of patients, epochs, and iterations in Section 6.3. The long computation time required to do so was prohibitive, and hence this was not done by the author due to time constraints.

## 7.3 Future Work

### 7.3.1 Further Evaluation of the Method

The method may be further evaluated through

- comparing it systematically to other measures of phase coherence
- applying it to the problem of seizure detection (as suggested in Section 6.3, or the detection of other pathologies that may be identified through measures of phase coherence).

### 7.3.2 Development of a Mathematical Framework for the Method

Further, points mentioned in Section 5.5 are recapitulated. Section 3.3 described the multivariate symbolization procedure of Zhang et al. [8]. At the end of this section it was noted that for  $n > 2$  the interpretation described of symbolization as a process of partitioning the state space no longer holds. The heuristic argument that inspired the algorithm (Section 5.5) nonetheless employed this interpretation, so that it was assumed that the behaviour whereby loops in the phase portrait correspond to a periodic succession of symbols still held. A mathematical framework must be developed to

- justify this analogy,
- provide a context within which the robustness of the algorithm can be understood, and
- inspire a method to evaluate the robustness of the algorithm.

# REFERENCES

- [1] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press, 1961, vol. 25.
- [2] J. Gleick, *Chaos: Making a new science*. Open Road Media, 2011.
- [3] S. Strogatz, *Sync: How Order Emerges from Chaos in the Universe, Nature, and Daily Lif*. Hachette Books, 2003.
- [4] M. Breakspear, "Dynamic models of large-scale brain activity," *Nature neuroscience*, vol. 20, no. 3, p. 340, 2017.
- [5] A. Pikovsky, J. Kurths, M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*. Cambridge university press, 2003, vol. 12.
- [6] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, "Surrogate data for hypothesis testing of physical systems," *Physics Reports*, vol. 748, pp. 1–60, 2018.
- [7] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC Press, 2018.
- [8] J. Zhang, J. Zhou, M. Tang, H. Guo, M. Small, and Y. Zou, "Constructing ordinal partition transition networks from multivariate time series," *Scientific reports*, vol. 7, no. 1, p. 7795, 2017.
- [9] M. Breakspear and J. Terry, "Detection and description of non-linear interdependence in normal multichannel human eeg data," *Clinical neurophysiology*, vol. 113, no. 5, pp. 735–753, 2002.
- [10] C. Robinson, *Dynamical systems: stability, symbolic dynamics, and chaos*. CRC press, 1998.
- [11] J. M. Amigó, K. Keller, and V. A. Unakafova, "Ordinal symbolic analysis and its application to biomedical recordings," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 373, no. 2034, p. 20140091, 2015.
- [12] R. Shaw, "Strange attractors, chaotic behavior, and information flow," *Zeitschrift für Naturforschung A*, vol. 36, no. 1, pp. 80–112, 1981.
- [13] N. Mammone and F. C. Morabito, "Analysis of absence seizure eeg via permutation entropy spatio-temporal clustering," in *The 2011 International Joint Conference on Neural Networks*. IEEE, 2011, pp. 1417–1422.
- [14] B. K. Hillen, G. T. Yamaguchi, J. J. Abbas, and R. Jung, "Joint-specific changes in locomotor complexity in the absence of muscle atrophy following incomplete spinal cord injury," *Journal of neuroengineering and rehabilitation*, vol. 10, no. 1, p. 97, 2013.
- [15] F. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields: Basic Principles, Clinical Applications and Related Fields*. Lippincott Williams & Wilkins, may 2012.
- [16] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for eeg signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 1, p. 66, 2015.

- [17] H. Ghadyali, "Applications of topological data analysis and sliding window embeddings for learning on novel features of time-varying dynamical systems," Ph.D. dissertation, Duke University, 2017.
- [18] J. G. Webster, *Medical Instrumentation: Application and Design*. Wiley, feb 2009.
- [19] W. Klonowski, "Everything you wanted to ask about eeg but were afraid to get the right answer," *Nonlinear Biomedical Physics*, vol. 3, no. 1, p. 2, 2009.
- [20] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: the long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2006.
- [21] G. B. Boylan, N. J. Stevenson, and S. Vanhatalo, "Monitoring neonatal seizures," in *Seminars in Fetal and Neonatal Medicine*, vol. 18, no. 4. Elsevier, 2013, pp. 202–208.
- [22] M. Mula and F. Monaco, "Ictal and peri-ictal psychopathology," *Behavioural Neurology*, vol. 24, no. 1, pp. 21–25, 2011.
- [23] T. N. Alotaiby, S. A. Alshebeili, F. E. A. El-Samie, A. Alabdulrazak, and E. Alkhnaian, "Channel selection and seizure detection using a statistical approach," in *2016 5th international conference on electronic devices, systems and applications (ICEDSA)*. IEEE, 2016, pp. 1–4.
- [24] S. Roberts, M. Osborne, M. Ebden, S. Reece, N. Gibson, and S. Aigrain, "Gaussian processes for time-series modelling," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1984, p. 20110550, 2013.
- [25] H. Edelsbrunner and J. Harer, "Persistent homology-a survey," *Contemporary mathematics*, vol. 453, pp. 257–282, 2008.
- [26] H. Krim, T. Gentimis, and H. Chintakunta, "Discovering the whole by the coarse: A topological paradigm for data analysis," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 95–104, 2016.
- [27] M. Piangerelli, M. Rucco, L. Tesei, and E. Merelli, "Topolnogical classifier for detecting the emergence of epileptic seizures," *BMC research notes*, vol. 11, no. 1, p. 392, 2018.
- [28] E. W. Chambers, J. Erickson, and P. Worah, "Testing contractibility in planar rips complexes," in *Proceedings of the twenty-fourth annual symposium on Computational geometry*. ACM, 2008, pp. 251–259.
- [29] R. Ghrist, "Barcodes: the persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61–75, 2008.
- [30] J. Curry, "The fiber of the persistence map," jun 2017.
- [31] A. Myers, E. Munch, and F. A. Khasawneh, "Persistent homology of complex networks for dynamic state detection," *Physical Review E*, vol. 100, no. 2, aug 2019.
- [32] A. Sallaberry, F. Zaidi, and G. Melançon, "Model for generating artificial social networks having community structures with small-world and scale-free properties," *Social Network Analysis and Mining*, vol. 3, no. 3, pp. 597–609, 2013.
- [33] J. A. Perea and J. Harer, "Sliding windows and persistence: An application of topological methods to signal analysis," *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 799–838, 2015.

- [34] R. Wilds and L. Glass, "Contrasting methods for symbolic analysis of biological regulatory networks," *Physical Review E*, vol. 80, no. 6, p. 062902, 2009.
- [35] J. A. Urigüen and B. Garcia-Zapirain, "Eeg artifact removal—state-of-the-art and guidelines," *Journal of neural engineering*, vol. 12, no. 3, p. 031001, 2015.
- [36] N. Stevenson, K. Tapani, L. Lauronen, and S. Vanhatalo, "A dataset of neonatal eeg recordings with seizure annotations," *Scientific data*, vol. 6, p. 190039, 2019.
- [37] S. Faul, G. Gregorcic, G. Boylan, W. Marnane, G. Lightbody, and S. Connolly, "Gaussian process modeling of eeg for the detection of neonatal seizures," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2151–2162, 2007.