uc3m | Universidad Carlos III de Madrid

University Degree in Aerospace Engineering

Academic Year (2020-2021)

Bachelor Thesis

"Non-linear Data Analysis applied to Plasma Space Propulsion"

Clara Puerto Sánchez

Tutor: Mario Merino Martínez

Leganés, July 2021

SUMMARY

Plasma physics play a central role in today's technology and science development. From plasma fusion to space propulsion, plasma presents unique characteristics that make it a distinguished candidate in the search for clean and unlimited energy sources. In the field of plasma research tremendous amounts of data are generated on a daily basis. This growth in the amount of data available to analyze has triggered as a consequence the field of data analysis, which has also experienced a huge increase and development. In many cases, the problem does not lay on the lack or quality of the data, but on the inability to obtain meaningful information from the results obtained by simple visual inspection. Moreover, the presence of researcher bias may influence the obtained results towards more promising conclusions than what the data is actually providing.

Even though linear data-driven techniques present several advantages and are widely used by the scientific community, such as easy implementation and low-computational costs, there are some limitations on the use of these techniques. The study of complex systems, such as plasma, cannot be completely covered by just the application of these techniques.

The continuous study of plasma has demonstrated that its behaviors does not follow linear trends, containing several instabilities and turbulence that linear analysis tends to ignore. Even though non-linear analyses normally present a more difficult and costly implementation, they provide a more complete understanding of its dynamics.

The Research Group "Equipo de Propulsión Espacial y Plasma" (EP2) following the ZARATHUSTRA project from University Carlos III of Madrid investigates the physics behind the operation of plasma thrusters and this Bachelor Thesis serves as a starting point on the review and implementation of these type of algorithms, with a particular emphasis on "data-driven" techniques, which minimize research bias to a minimum. These techniques are influenced by several fields, ranging from linear algebra and calculus, to information and chaos theory.

A complete study and classification of each technique is presented along with its implementation against renowned non-linear systems, such as the Lorenz Attractor and the Ikeda Map. Finally, a selection of these techniques is used to explore the mechanisms in Helicon Plasma Thrusters and Hall-Effect Thrusters that lead to plasma turbulence in order to contribute to the understanding of the underlying plasma dynamics.

Keywords: Non-Linear Analysis, Plasma, Thruster, Aerospace Propulsion, Information theory, Chaos, Turbulence

DEDICATION

To my parents, for constantly believing in me and pushing me to become the best version of myself, even if that meant putting an ocean between us. Everything I do and everything I am is because of the support and love I always receive from you.

To my sister, Lucia, for showing me how important following your dreams is, no matter the complicated path that lays in front of you.

To Dr. Mario Merino, for trusting me with this project, and for his constant mentorship and guidance. The Plasma Propulsion field is in good hands.

To Dr. Adrián Domínguez-Vázquez for providing the HET simulations and to Davide Maddaloni and Dr. Jaume Navarro-Cavallé for the experimental HPT data on which this project is based.

To Marta and Patricia, you have been the luckiest coincidence of my life.

To all my friends, family, academic advisors and colleagues I have found along the way, thank you for accompanying me in this journey. It is time to begin a new chapter.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation

1.1.1. Data Analysis

In Plasma Aerospace Propulsion, and in all fields relating to Plasma Physics and Aerospace Engineering, tremendous amounts of data are generated on a daily basis through simulations and experiments. According to the experts from the International Thermonuclear Experimental Reactor, the world's largest fusion experiment, by 2035 ITER alone will produce 2 petabytes (1 petabyte = 10^{15} bytes) of data everyday [1].

The problems studied in these fields consistently present complex behaviors, such as the characterization of plasma turbulence and the identification of causal relationships between phenomena that occur in these systems [2]. Specifically, one of the main challenges of Plasma Physics revolves around the nonlinear character of plasma [3].

In many cases, the problem does not lay on the lack of data, but on the inability to obtain meaningful insights into the results obtained by simple inspection. Moreover, the presence of researcher bias may influence the obtained results towards more promising conclusions than what the data is actually providing [4] [5].

As a response to the growth in scientific raw data available, numerous data analysis techniques have been developed over the years in order to extract evidence and conclusions from experiments and simulations. The Research Group "Equipo de Propulsión Espacial y Plasma" (EP2) following the ZARATHUSTRA project from University Carlos III of Madrid investigates the physics behind the operation of plasma thrusters and this Bachelor Thesis serves as ZARATHUSTRA' starting point on the review and implementation of these type of algorithms, with a particular emphasis on "data-driven" techniques.

Data-driven techniques refer to all approaches to data-analysis in which no external information is needed in order to perform the exploration of the results. These type of interpretations present the advantages of removing research bias and detecting trends that may not be apparent from visual examination. As it will be shown throughout this project, these methods are influenced by several fields such as: linear algebra, calculus, information theory and chaos theory.

1.1.2. Hall-Effect Thruster

Hall effect Thrusters (HETs) are some of the most developed electric thrusters nowadays. Three different particles are used in HETs; a neutral gas, injected from the rear part of the channel, which flow axially towards the exit of the thruster; electrons, injected through the cathode just outside the channel and flow upstream towards the anode; and ions, created by the collision of neutrals and electrons [6].

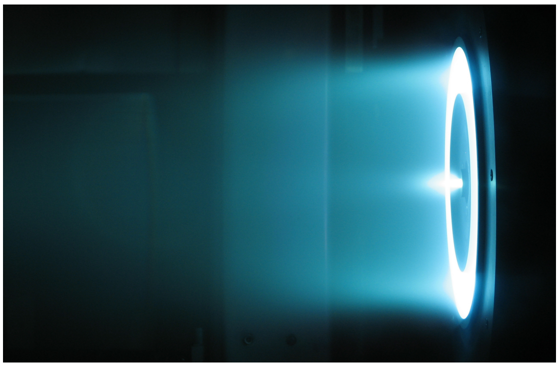


Fig. 1.1. JPL's 6 kW Hall thruster. Retrieved from [7].

A HET is composed of a discharge region, a cathode and a magnetic field generating system. The discharge region is surrounded by an insulating material and magnetic coils are normally used to induce a radial magnetic field. The cathode is located right outside the channel and the at the base of the discharge region the anode can be found. The anode is a ring located at the bottom part of the channel through which the neutral gas used as propellant in injected. Electrons coming from the cathode and trying to reach the anode encounter the radial magnetic field which eliminates their axial mobility almost completely and traps them in azimuthal orbits in the ExB direction. As mentioned before, ions are then generated by the collision of neutrals with electrons, to be finally accelerated by the electric field from the anode to the cathode towards the end of the thruster [8].

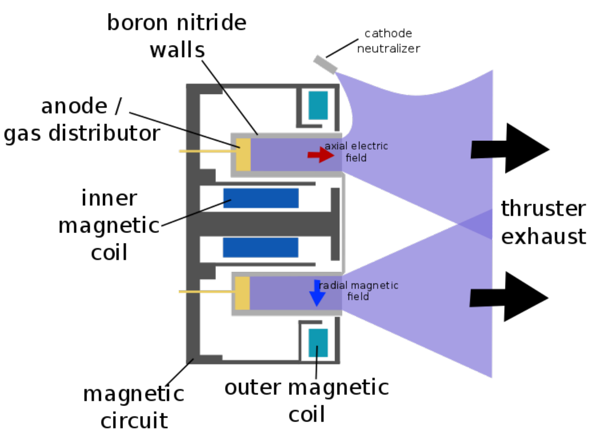


Fig. 1.2. Diagram of Hall Effect Thruster. Cross-section. Retrieved from [9].

The main applications of these thrusters are interplanetary probe missions requiring low-thrust, satellite maneuvers to reach specific orbits and north-south station-keeping operations of geostationary satellites [8].

1.1.3. Helicon Plasma Thruster

A Helicon plasma Thruster (HPT) is an electromagnetic thruster that is being developed and improved in order to be applicable to space missions in the near future. The HPT presents several advantages when compared with other electromagnetic thrusters, such as the Hall effect Thruster (HET), because it presents a much simpler assembly and it does not need additional elements such as grids, electrodes or neutralizers [10].

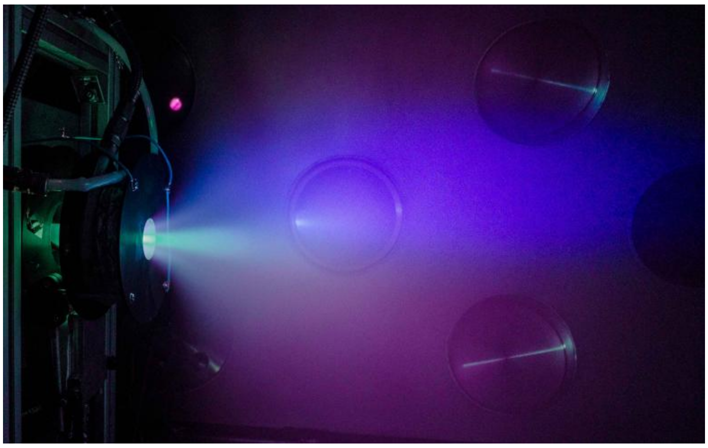


Fig. 1.3. HPT-05 prototype developed by EP2 Research Group and SENER Ingeniería y Sistemas. Retrieved from [11]

The architecture of a HPT typically consists of the following; a cylindrical chamber, where plasma is produced; a magnetic field generator, used to confine, guide and expand the plasma on the magnetic nozzle; a Radio Frequency (RF) system that emits RF waves; and an injector system to drive the inert gas into the chamber [12].

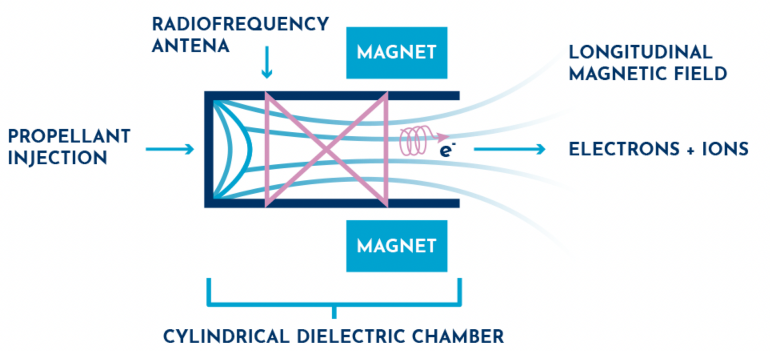


Fig. 1.4. Diagram of Helicon Plasma Thruster. Cross-section. Retrieved from [10].

The inert gas, after being injected into the chamber, is ionized by the electromagnetic RF field, producing hot plasma. A divergent magnetic topology is created and used to accelerate the plasma into a supersonic flow, transforming the thermic energy of the plasma into kinetic one, and thus generating thrust.

Some of the advantages that the HPT presents are; reduced costs, since it does not require large voltages for its operation, many elements are not necessary and manufacturing expenses are reduced, and in addition, less propellant is used during its operation, further reducing the overall costs; longer lifetime, due to the chamber being magnetically shielded and no electrodes being used, erosion is reduced favoring a longer operational life; possibility of using almost any gas as propellant and finally, higher thrust capacity per unit of power than other electric propulsion systems, such as the HET.

HPTs could potentially be used for missions involving orbit observations in VLEO, exploration operations where longer lifetimes are needed and telecommunications applications where reduced propellant consumption may be relevant [10].

1.1.4. Introduction to plasma turbulence

Even though there has been extensive research on the plasma propulsion and plasma fusion fields, many aspects of these thrusters are still not fully understood. The main challenge that scientists face is trying to explain the higher-than-expected electron conductivity found inside thruster channels, which cannot be explained by classical collisional theories alone. The term “anomalous transport” has been created to describe this phenomenon in which experimental electron conductivity is one or two orders of magnitude higher than the one expected from theoretical calculations.

There is still not a clear explanation as to why this occurs, however some of the most general theories point to plasma oscillations, also referred to as turbulent diffusion, or near-wall conductivity. Plasma oscillations in the azimuthal direction along with oscillations of the electric field could induce a net axial electron current and explain the anomalous transport. On the other hand, the near-wall conductivity hypothesis is not as widely accepted since some simulation codes that include near-wall conductivity models still do not correctly produce the electron conductivity found in real-life experiments [6].

This Bachelor Thesis aims to explore this phenomenon by means of non-linear analysis techniques and help throw some light into the mechanisms that lead to the appearance of turbulence in plasma thrusters.

2. OBJECTIVES

2.1. Objectives overview

Experimental measurements of plasma thrusters in the laboratory and advanced numerical plasma simulations of these devices produce vast quantities of data that need to be analyzed to yield physical insights. Many of the physical phenomena in plasmas are inherently nonlinear: simple linear cross correlations and modal analyses miss the nonlinearity and hidden casualty relations that may exist between different aspects of the operation of a plasma thruster. This project proposes to analyze experimental and simulation data using state of the art data-driven analysis techniques based on nonlinear system theory, chaos theory and information theory. To fulfil this purpose, several objectives were addressed:

1. Review of the state-of-the-art for Non-Linear Analysis Techniques and creation of algorithms catalogue

   Review and evaluation study of existing tools collected from a variety of sources regarding Non-Linear Analysis techniques, with an especial focus on those applicable to Plasma Space Propulsion. Creation of algorithms library with all the reviewed techniques explained, to be stored and further developed by the EP2 group.

2. Implementation of data-driven analysis algorithms

   All techniques selected will be implemented in MATLAB and validated with well-studied chaotic systems, such as the Lorenz system.

3. Application to laboratory data and simulation data. Comparison of different non-linear techniques

   Experimental and simulation data obtained from the EP2 group will be analyzed using the techniques mentioned before and the different results will be compared.

4. Analysis of results

   The last objective is to provide a physical insight from the results obtained, in particular the role of turbulence, instabilities, and trends in the anomalous transport of plasma particles

2.2. Content of the document

Throughout this Thesis, a methodology section will be introduced in which a brief review of traditional linear techniques will be provided along with all the non-linear techniques studied for this project, which will be implemented and tested. Following that, a chapter will be devoted to the data that will be employed in the Results and Discussion chapter, in order to give some background information regarding the context of the data selected. To conclude, a brief overview of the legal framework and the socio-economic impact of this project will be presented in the final pages.

3. METHODOLOGY

3.1. Linear Data-Driven Techniques

Linear Data-Driven techniques present several advantages. Primarily, they are fairly easy to implement and they normally do not present a high computational cost. However, by applying linear algorithms to some systems, especially those having complex dynamics such as the ones studied in this project, major relationships between the data may be ignored.

The study of plasma physics has demonstrated that its behaviour does not follow linear trends and contains several instabilities that linear analysis does not correctly capture [13]. Even though applying these linear techniques serves as a significant first step, more intricate approaches will be necessary to fully understand plasma dynamics.

3.1.1. Spectrum Analysis

The Fourier Transform of a system translates a signal expressed in the time domain into the frequency domain [14]. In other words, the Fourier Transform identifies what frequencies exist in a time series and with what magnitude they are present. Mathematically it is expressed as:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi f t}dt \tag{3.1}$$

This technique is widely used because any waveform can be expressed as a combination of sines, and the Fourier Transform extract from a given sample the frequency and the magnitudes of the sines present in said waveform.

If a signal contains N points, the frequency bins studied in the frequency domain have a width of:

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

Being f_s the sampling frequency of the time-series. The lowest frequency studied is normally $0Hz$ and the highest frequency is the Nyquist frequency, $f_s/2$.

Additionally, even if noise is present in the time-series studied (up to a certain level), the Fourier Transform still recovers the underlying structure of the waveform. In order to illustrate the performance of the Fourier Transform, a simple example has been created. The created waveform and its corresponding Fourier Transform with and without noise are presented below:

$$y(t) = sin(2\pi 50t) + 3sin(2\pi 65t) + 2sin(2\pi 100t) \tag{3.3}$$

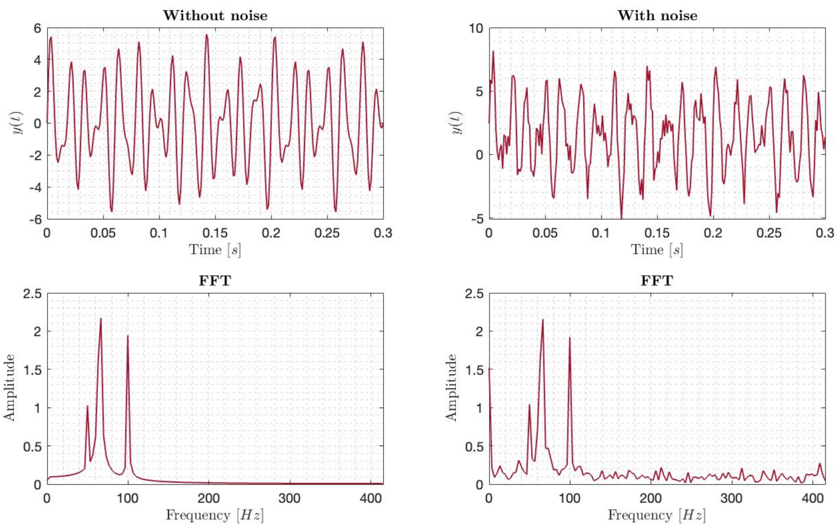


Fig. 3.1. Fourier Transform of a simple waveform: without noise (left plots) and with noise (right plots). $f_s = 830Hz$.

As it can be seen on the figure, the Fourier Transform perfectly identifies both the frequencies of the corresponding wave and the magnitudes of said frequencies, even if noise is present on the sample.

3.1.2. Correlation Analysis between signals

Correlation describes the existing relationship between two signals. It provides a quantification about how much the signals resemble each other.

There are two different types of correlation, depending on the inputs provided: autocorrelation and cross-correlation [15].

- Cross-correlation

  For the cross-correlation, the two signals studied are different. The objective of this analysis is to find how much similarity there is between them. Mathematically this is expressed as:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y^*(t-\tau)dt \tag{3.4}$$

  Being $y^*(t)$ the complex conjugate of $y(t)$.

- Autocorrelation

  Autocorrelation is an especial case of cross-correlation in which the signal is correlated with itself. Normally this involves a time-shift between the signal and itself, or a transformation similar to this. The mathematical formula to express this is as follows:

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt \tag{3.5}$$

An illustrative example of the correlation analysis has been generated in order to exemplify its use:

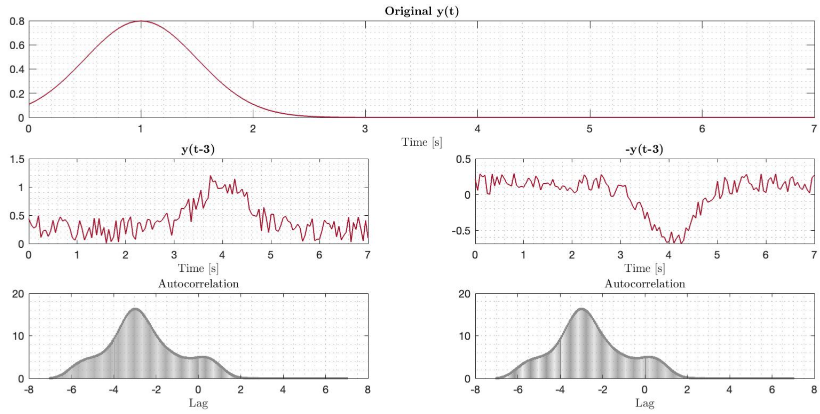


Fig. 3.2. Correlation analysis of simple signals: with finite time-lag (left plots) and inverted with finite time-lag (right plots).

Furthermore, Correlation can also be used to determine the periodicity of a signal by means of the autocorrelation function. This task can be performed by autocorrelating a function with a time-shifted version of itself and accounting for the repetitions of maximums found in the resulting graph.

The correlation analysis have some clear advantages highlighted here, however there are some set-backs in these type of analyses, such as the lack of directionality in its results, that will be illustrated along this project.

3.1.3. Modal Analysis techniques

Modal analysis techniques seek to find relevant features or oscillations, known as modes, from the structure or fluid studied [16]. Experimental modes are obtained by exciting the system and recording its response. This excitement can be done through several paths, the most common ones involving shakers or impact testing [17].

Each mode is characterized by a natural frequency, a modal damping and a mode shape. Near the natural frequency of a certain mode, that mode will often dominate the whole behaviour of the system. This is one of the main advantages of these type of techniques, by finding the natural frequencies, the responses of the system under specific conditions can be accurately predicted beforehand. This can be appreciated in the following figure, when the natural frequency on one of the modes is reached, the overall system behaves according to that specific mode.

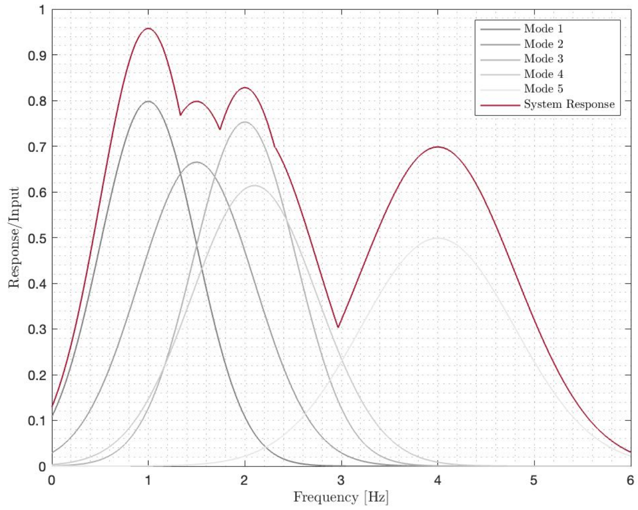


Fig. 3.3. System response and individual modes contribution. Example retrieved from [17]

The modes depend greatly on the specific conditions of both the system and the exciting inputs. Changing these conditions will modify the modes and hence, the response obtained. However, characterizing the response under general or critical conditions still adds a lot of value to the overall study of a system.

Some examples of modal analysis techniques are the eigenvalue and the singular value decomposition. These techniques will not be explained in this thesis, since they lay outside the scope of study of this project, but they have been thoroughly studied and explanations on them can easily be found on literature [18] [19].

3.2. Non-linear Data-Driven Techniques

Because of the nonlinear coupling and the complexity of non-linear systems, such as plasma, linear analyses are not sufficient to fully encompass the whole behaviour of their dynamics [20]. Non-linear behaviours, such as the anomalous transport presented in Section 1.1.4., are one of the main focus of plasma physics nowadays, and the application of non-linear analysis techniques is starting to become one of the most powerful tools to understand these phenomena [21] [22].

Even though non-linear analyses normally are more costly computer-wise, this theses aims to attain a more complete understanding of the underlying dynamics of plasma propulsion with the application of these techniques.

3.2.1. Time Delay Embedding techniques

A time series is defined as a sequence of measurements taken over time, where the ordering of time carries part of the information contained on the data. In other words, if a time series was to be shuffled, some of the information would be lost. Due to the simplicity of recording these type of events, time series represent a large volume of experimental and numerical results obtained nowadays. Following that reasoning, a high number of techniques have been developed in order to extract the maximum information from them.

Normally the data obtained corresponds to a projection of the dynamics of the event studied from a higher dimensional phase space. In order to study said dynamics, the signal can be represented in a higher dimension, this process is known as embedding the data. The two most common ways of embedding data is using the derivatives of the state studied or using time lags from the time series. Both approaches are pretty much equivalent, but the use of time lags constitutes a less expensive process and has a lower sensitivity to noise in the sample. Furthermore, the reconstruction of the phase space using time delays translates trends buried in the time series into regions in the embedded space, making it easier to analyze and visualize them. All the phase spaces reconstructed throughout this project follow the second approach presented here, the Delay Reconstruction [23].

Time Delayed Mutual Information

For an ideal noise-free infinite data set, the original phase space and the embedded one present the same characteristics regarding their topology and differentiability. However, experimentally infinite data sets cannot be recorded and the correspondence between the original and the reconstructed attractor depends greatly on the choices of τ and m.

If the time lag selected is too small, a phenomenon known as redundancy will be encountered. This is due to the fact that all coordinates will almost coincide and the trajectories in the reconstructed phase-space will virtually follow a line. On the other hand, if the time lag selected is too large, especially when dealing with series that are governed by chaotic motion, the points may seem as independent even if the system is deterministic [24].

The method to find an appropriate embedding dimension will be detailed in the following section, but first a suitable time lag should be determined. The technique presented in this section to find the optimum τ, the Time Delayed Mutual Information, was introduced by Fraser and Swinney in 1986 [25].

For a general case in which neither the past history of the time series nor the coupling with the current state is assumed to be deterministic, the amount of information contained in a variable of the series, I, is given by Shannon Entropy [26]:

$$H_I = -\sum_i p(i) log_2 p(i) \tag{3.6}$$

The base 2 of the logarithm is just an indication of the units chosen to measure the gain in information [27] and the associated probability distribution p(i) represents the probability of the state to take that specific state over the period studied [27]. The mutual information technique was created. Mutual information between two processes is defined as the excess amount of data used to predict a state due to the wrong assumption that it is independent from its past-history. In other words, it quantifies how much information is gained by observing the current (I) and past (J) states of the system instead of just the current one. Mathematically it can be expressed according to Kullback entropy [28]:

$$M_{IJ} = \sum p(i, j) log_2 \frac{p(i, j)}{p(i)p(j)} \tag{3.7}$$

By perfoming the study of M_{IJ} for a range of τ, the first minimum found represents the time lag which adds maximal information and is selected as the optimal one [29]. One advantage that this technique has when compared to others, is that it takes into account the nonlinear nature of the system, something critical for the study of plasma propulsion [24]. In order to illustrate this method, the Lorenz system, a widely known chaotic system, has been used as an example:

$$\dot{x} = \sigma(y - x) \tag{3.8}$$

$$\dot{y} = x(\rho - z) - y \tag{3.9}$$

$$\dot{z} = xy - \beta z \tag{3.10}$$

With $\sigma = 10, \beta = 8/3$ and $\rho = 28$. The result obtained is shown below:

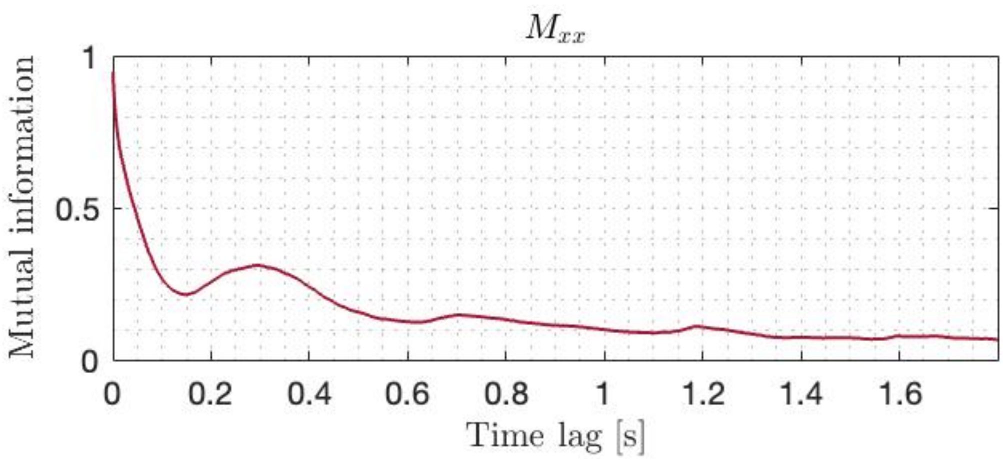


Fig. 3.4. Average Mutual Information: Lorenz system

As it can be appreciated from the figure, the first minimum lays around $t = 0.16s$, which agrees with the optimum time lag for the Lorenz system found in the literature [30].

False Nearest Neighbours

The method of False Nearest Neighbours (FNN) was introduced by Kennel, Brown and Abarbanel in 1992 [31]. This method seeks to find the minimum embedding dimension, m, conditional that the optimal time delay, τ, is provided.

FNN is based on the notion that the evolution of a dynamical system in state-space is defined by a smooth vector field. According to that concept, two points that are originally close, after some time t has passed, should remain close, even if the system is chaotic. However, if the system is represented in a dimension lower than its minimum embedding dimension, points that are not close may seem as adjacent neighbours due to the reduction of effective dimensions used. In other words, if a lower embedding dimension is used, that means that some of the coordinate axes necessary to fully represent the system are being ignored, and the points will be projected onto the remaining ones. This may lead to points that were originally far away having images that make them look like nearest neighbours [29].

In order to detect the false neighbours, the dimension m is varied and the system evolution is studied according to [31]:

$$R_m^2(i) = \sum_{i=0}^{m-1} [x(i + m\tau) - x_{n(i,m)}(i + m\tau)]^2 \tag{3.11}$$

$$R_{m+1}^2(i) = R_m^2(i, r) + [x(i + m\tau) - x_{n(i,m)}(i + m\tau)]^2 \tag{3.12}$$

$$\sqrt{\frac{R_{m+1}^2(i) - R_m^2(i)}{R_m^2(i)}} > R_{tol} \tag{3.13}$$

Being $x_{n(i,m)}$ the nearest neighbour of x. The False Nearest Neighbours approach depends greatly on tolerance values decided by the user, such as R_{tol}, which may lead to significant differences on the minimum embedding dimension found.

Cao introduced in 1997 a modification to this method in order to reduce the chosen thresholds and provide more reliable outcomes. Instead of studying the number of false neighbours found, Cao proposed to study how the distance between these neighbours changes when increasing the embedding dimension [24].

In a similar way to the original FNN method, the following variable is calculated [32]:

$$a(i, m) = \frac{||y_i(m + 1) - y_{n(i,m)}(m + 1)||}{||y_i(m) - y_{n(i,m)}(m)||} \tag{3.14}$$

Being $y_{n(i,m)}$ again the nearest neighbour of y_i in embedding dimension m and representing $|| * ||$ the maximum norm between them. The mean value of all $a(i, m)$ is given by:

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} a(i, m) \tag{3.15}$$

The variation of E from m to $m + 1$ is then defined as follows:

$$E1(m) = \frac{E(m+1)}{E(m)} \tag{3.16}$$

Provided that the time series studied has an underlying attractor, $E1$ will stop changing once an embedding dimension m_o is surpassed. The minimum embedding dimension is then: $m_o + 1$ [32].

If a time series of random numbers is studied, it is expected that $E1$ may never reach a stable value, however, there might be cases in which this occurs. To resolve this situation, the following additional variable was introduced by Cao [32]:

$$E^*(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} |x_{i+m\tau} - x_{n(i,m)+m\tau}| \tag{3.17}$$

$$E2(m) = \frac{E^*(m+1)}{E^*(m)} \tag{3.18}$$

For stochastic systems, the lack of dependency from past states guarantees the independence from the embedding dimension, which means that $E2$ will always be 1. However, deterministic systems do depend on the embedding dimension, so $E2$ will not be constant for all values of m. Cao suggested to study both $E1$ and $E2$ together to be able to correctly find the minimum embedding dimension of a system and to distinguish if said system is deterministic or random [32]. Following the example of the Lorenz system to illustrate this method, the following results were obtained:

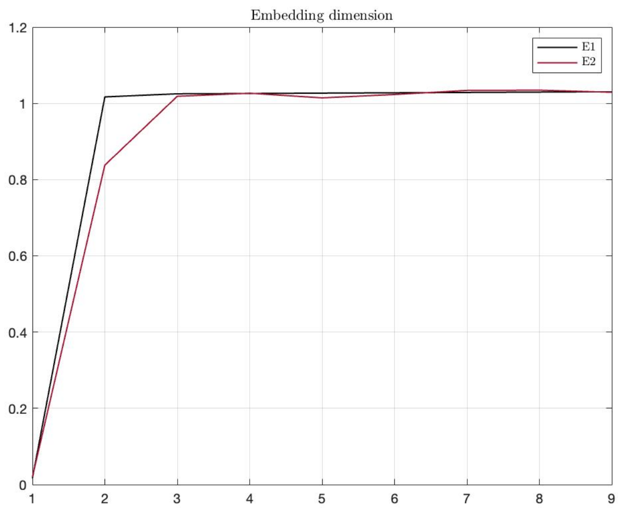


Fig. 3.5. CAO method: Lorenz system

From the figure it can be reasoned that $m_o = 2$, so the minimum embedding dimension for Lorenz is $m = m_o + 1 = 3$, which is known to be true since Lorenz can be simply described by x, y and z.

3.2.2. Visual techniques

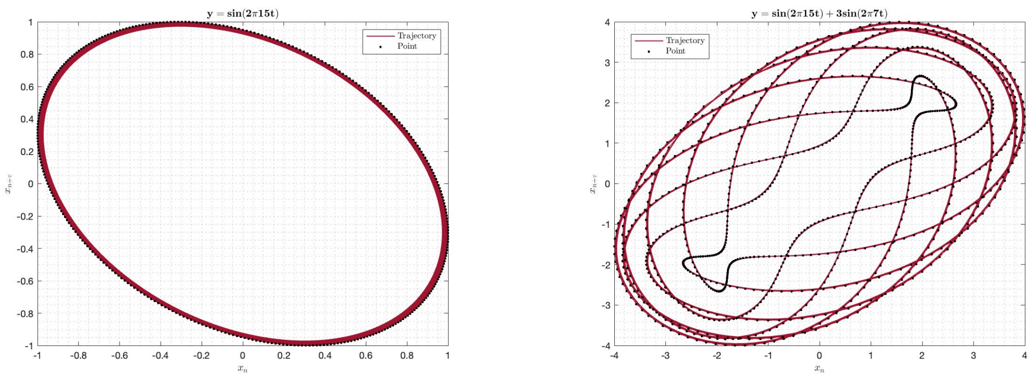
Phase Portraits

The Poincaré graph, also known as Phase Portrait, was developed by Henri Poincaré in 1890 [33]. There are several ways of presenting these graphs, but in this thesis the focus will be on those who graph an interval of the time-series provided, x_{n+1}, against its preceding points, x_n. This simple technique allows to analyze visually the correlation between consecutive points in the data set. Different behaviours can be found depending on the underlying dynamics of the set, the main general ones will be explained below.

The majority of the systems will always present a bounded Phase Portrait, only if the system tends to infinity the Phase Portrait will grow unboundedly as the number of points studied is increased.

Periodic systems normally present a uniform bounded behaviour. In these cases, the optimum lag found from the Mutual Information minimum explained in Section 3.2.1 presents more advantages for this technique instead of directly plotting the Poincaré graph for the previous state.

A circle is recovered for the case of pure single periodic systems, such as sines. For a more complicated periodic system, such as those composed by more than one sine, a closed curve is recovered, but it does not necessarily need to be a circle. Some examples for a simple sine and an addition of sines have been produced in order to illustrate this:



(a) Pure sine. (b) Combination of sines.

Fig. 3.6. Poincaré graph for a sinusoidal functions.

One of the main advantages of the Poincaré graphs is that they help discover the underlying attractors of dynamic systems. If the embedding dimension of the systems is too high, the recovered Poincaré graph will be a projection of said attractor on a plane, instead of the full attractor. However, said projections may still contain useful information regarding the dynamics of the studied system.

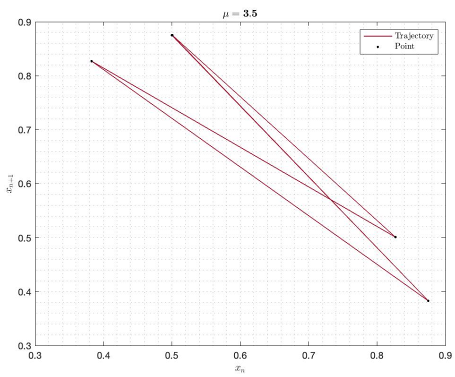
Different types of attractors may be found; for the case of systems with fixed-points, the attractor will just be a point in space, with an associated dimension zero; limit cycles will be represented as curves, as the one presented in Figure 3.6, with a dimension one; and finally, even "strange attractors" may be found. Strange attractors are irregular structures towards which the system tends, that normally present non-integer dimensions and self-similarity, they are characteristic of chaotic systems [34].

It should be also mentioned that if the dimension of a deterministic system is equal to 1, this system will be single-valued. In other words, it will only have one ordinate per abscissa. However, if the dimension is higher than one, the function will be multi-valued, with several ordinates per abscissa.

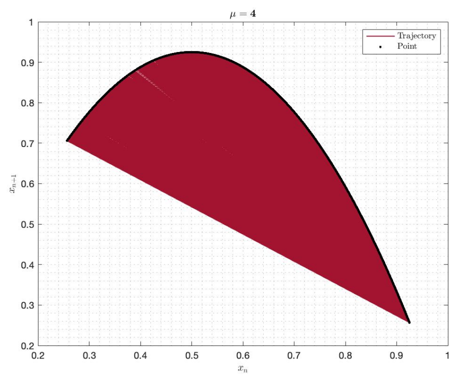
Let's use the Logistic Map example [35] to illustrate this technique for chaotic and non-chaotic systems:

$$x_{n+1} = \mu x_n(1 - x_n) \tag{3.19}$$

The Logistic Map is a polynomial map that is normally presented when studying chaos because of the simplicity of the formula. By varying the value of μ, the system goes from non-chaotic to chaotic behaviour really quickly at around $\mu = 3.6$.



(a) Non-chaotic.



(b) Chaotic.

Fig. 3.7. Phase portraits for the Logistic Map.

For the lower value of μ represented in this graph, corresponding to 3.5, this regression is not chaotic. No complex attractor was found for this case, the system is completely defined by the four points shown in the picture. However, once the transition to the chaotic regime has been performed, as in the case of Figure 3.7(b), the parabolic behaviour shown is found. This is due to the fact that as the transition to chaoticity occurs, the orbits which defined the dynamics of the system, 4 in the case of Figure 3.7(a), start exponentially increasing until the full parabolla shown on 3.7(b) is completed.

The Lorenz system has also been studied with this technique and its corresponding optimum time lag in order to further exemplify the potential of this technique, as shown in the following page.

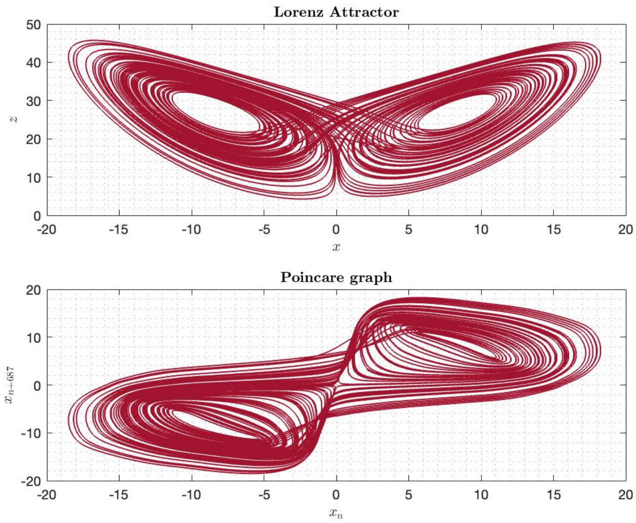


Fig. 3.8. Poincaré graph for Lorenz system. Optimum lag: $\tau = 0.16s$.

Here, an approximation to the real Lorenz attractor was found by using the correct time lag, which further supports the potential that this technique has when studying complex systems, since it recovers a projection of the full attractor, even if said system has a really high dimension.

Recurrence Plot

A Recurrence Plot (RP) is a graphical technique used to extract information about temporal correlations inside a time series. This tool was developed by J.P. Eckmann, O. Kamphorst and D. Ruelle in 1978 [36]. RPs rely on repetitions inside the data studied to extract meaningful information that may not be evident from direct inspection. Recurrence is defined as the return of a trajectory in state space to a neighbourhood that had already been visited before.

Recurrences are expected for any type of motion which is not transient, such as systems with; a fixed point, which will be recurrent for any studied time t; or a limit cycle, which presents a recurrence to its initial points when a complete period is performed. For chaotic systems, a point on a chaotic attractor will return to a neighbourhood of any of its points, as it happens with the Lorenz attractor. This recurrence is guaranteed by the invariance of the set, which is the basis of the attractor. If the time series studied never returns to any of its points, this can be due to non-stationarities included in the system or to the fact that the series studied was a transient. A transient represents points that are outside the invariant set but will tend to it as time goes to infinity [29].

Since RP represent a signal against itself, the obtained matrix and graph corresponding to it are symmetrical with respect to its main diagonal. The main diagonal represents the case in which each individual point is compared to its counterpart from the other signal. Following this reasoning, it would be enough to just represent either the upper half or the lower half of the matrix, but it is customary to present the whole plot.

As explained in a previous section, one of the advantages of reconstructing the phase space with time delays, is the translation of temporal trends to spatial features. For the time series $x(t)$ studied, which contains N points, the time-delayed phase space is reconstructed and the recurrence plot will consist on a NxN array (symmetric by construction), where a step function (Θ) will be used to obtain the graphical representation of the system. A 1 will be placed at (i, j) if $x(j)$ is closer in the phase space than a tolerance (ϵ) to $x(i)$, and a 0 otherwise. The tolerance is normally chosen according to the recurrence rate, the number of 1s present to the size of the matrix (total number of possible states for the system), and is normally set in order to obtain a recurrence rate of $5 - 10\%$ [37]. Mathematically this can be expressed as:

$$M_{ij} = \Theta(\epsilon - |x_i - x_j|) \tag{3.20}$$

The obtained array, as mentioned before, will be symmetric due to the fact that a constant tolerance is used for each point studied and the signal is compared with itself. Eckmann et all. [36] also used varying tolerances for each point until a minimum number of neighbours were found, in this case the array will not present a perfect symmetry but will still provide symmetric tendencies.

The Recurrence Plot obtained will significantly depend on the embedding dimension m and, in a less relevant way, on the time lag τ used to construct the state space from the time series $x(t)$. In deterministic systems, two points which are close should have projections under its dynamics which are also close, even if dynamical instabilities are present. Typically, short line segments appear parallel to the diagonal due to this fact. If scattered dots are found, these can appear because of noise in the time series, or because an incorrect low embedding dimension was being used [29]. The following figures serve as an illustration of this phenomena:

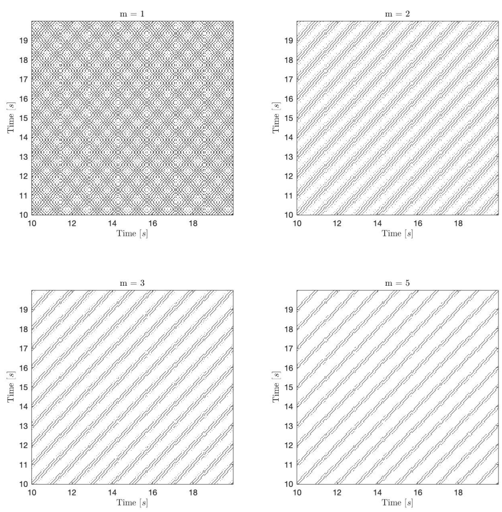


Fig. 3.9. Evolution of Recurrence Plot for increasing values of embedding dimension, m. Study performed for periodic system, whose real embedding dimension is $m = 3$. $y(t) = sin(2\pi * 150t) + 3sin(2\pi * 43t)$

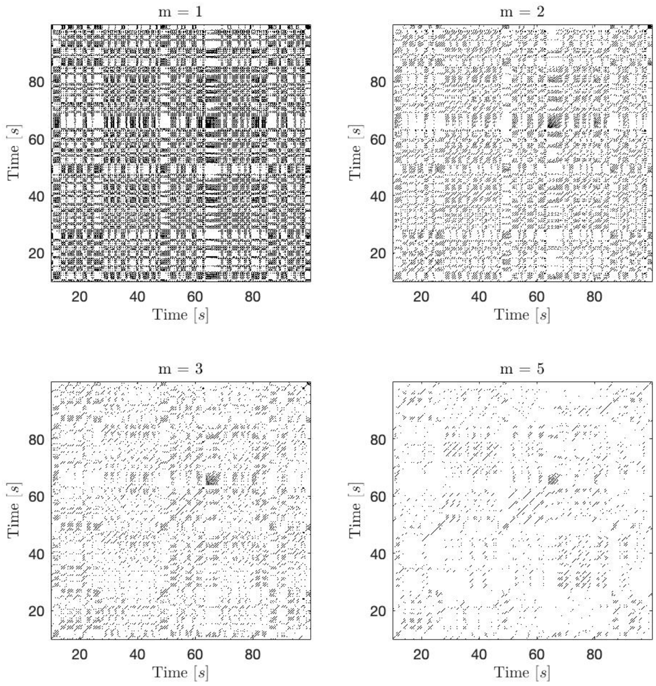


Fig. 3.10. Evolution of Recurrence Plot for increasing values of embedding dimension, m. Study performed for Lorenz System behavior, whose real embedding dimension is $m = 3$.

The sampling rate is another important characteristic that should be closely monitored to ensure the correct Recurrence Plot is being retrieved. If a sampling rate too close to the natural frequency of a periodic system is used, the pattern obtained will not be the one corresponding to the periodic topology, and might lead to wrong conclusions [38]. An illustrative example of this is represented in figure 3.3.

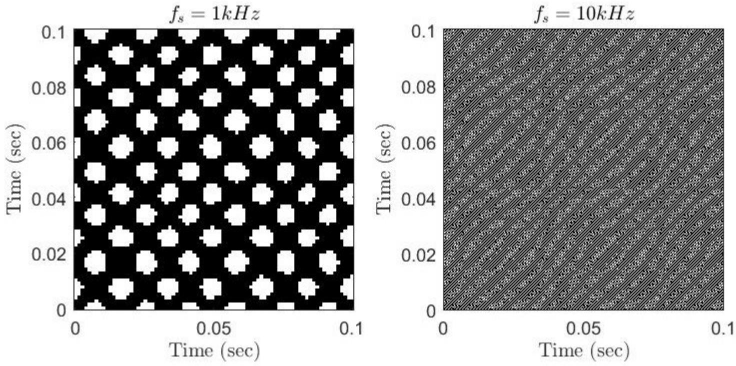


Fig. 3.11. Recurrence plots for time series: $y(t) = cos(2\pi 1000t + 1/2sin(2\pi 25t))$. Left plot: Sampling frequency equal to frequency of harmonic signal. Right plot: Sampling frequency bigger than frequency of harmonic signal. State-space reconstruction variables: $m = 3, \tau = 7$

Some of the most characteristic patterns found in Recurrence Plots are listed below and illustrated in figure 3.2 [36] [38]:

1. Periodic topology

   Periodic systems are captured by Recurrence Plots as lines parallel to the diagonal. The distance between successive lines is equivalent to the period of the equivalent periodic signal. An adequate sampling rate must be used in order to be able to recover perfect parallel lines with respect to the main diagonal. This has been illustrated on. Figure 3.11.

2. Chaotic topology

   Chaotic systems represented by Recurrence Plots are defined as irregular checkerboard textures. This designation is due to the horizontal and vertical white lines that appear, which correspond to transitions inside the system. The inverse of the longest diagonal, without taking into account the main diagonal, is proportional to the largest positive Lyapunov exponent [36]. Lyapunov exponents will be introduced later on in this Thesis.

3. Brownian motion

   Brownian motion is defined as the path followed by a particle moving randomly without any big jumps or disruptions [39]. This type of motion is characterized by the appearance of lines both parallel and perpendicular to the main diagonal. In addition, curved lines may appear joining some of the clusters of the recurrent trajectories.

4. White noise

   Noise normally appears as a heterogeneous pattern of scattered points with no apparent structure visible. As shown previously, by increasing the embedding dimension, most of the noise can be effectively eliminated. Another way of eliminating said scatter points would be to filter the time series studied before processing it.

5. Non-stationary data

   Non-stationary data is represented by Recurrence Plots as a fading scheme towards the upper left and the lower right corners. Patterns may be found close to the main diagonal, but they start fading away as the corners are approached. This indicates a trend or drift in the time series being studied.

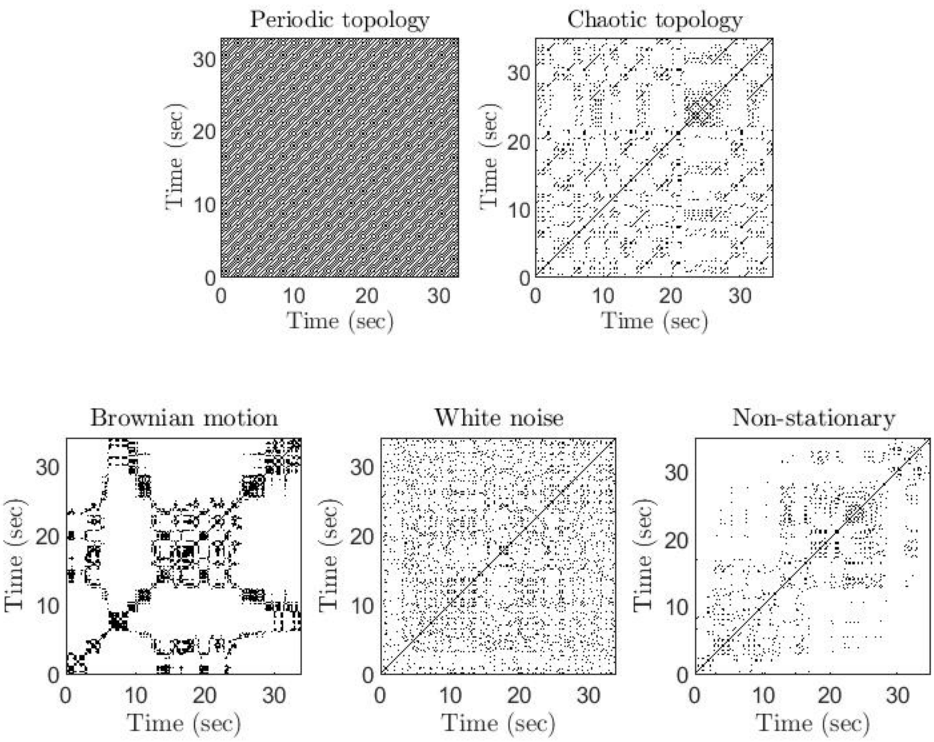


Fig. 3.12. Representation of most common Recurrence Plots found in literature.

Recurrence plots have proven its value in the field of plasma physics in several projects, being used as a graphic tool to analyze low-temperature discharge plasma [40] and to perform dynamic analysis on the JET Tokamak [22] to cite some examples.

The advantage that Recurrence Plots present when compared with other techniques, such as Phase Portraits, is that apart from using the temporal information of the signal, it uses the embedding dimension of the system, m. The recurrences of the data-points are studied in the higher dimensional phase space, and illustrated in a 2D grid. So even if the output of this technique is two-dimensional, it provides information regarding the distribution of points in the real Phase Space dimensions of the system.

3.2.3. Model reduction techniques

Hankel Alternative View of Koopman (HAVOK)

The HAVOK technique is built based on The Koopman operator, which lay outside the scope of this project, but information regarding this mathematical operator can be found in the following reference [41].

The HAVOK method seeks to find this Koopman invariant measurement space through the construction of a Hankel matrix, H. H is built by time-delayed coordinates, as follows:

$$H = \begin{bmatrix} x(t_1) & x(t_2) & ... & x(t_p) \\ x(t_2) & x(t_3) & ... & x(t_4) \\ ... & ... & ... & ... \\ x(t_q) & x(t_{q+1}) & ... & x(t_m) \end{bmatrix} \tag{3.21}$$

Once the Havok matrix is built, the Singular Value Decomposition (SVD) of this matrix is performed: $H = U\Sigma V^T$. SVD provides the Eigen-time delayed coordinates ordered from most to least important regarding their ability to describe the studied data. The first r columns of U represent a Koopman invariant measurement space. The number r can be found by different techniques that lay outside the scope of this project [42] [43].

The HAVOK method relies on the assumption that chaos is not random, it has a structure and patterns contained within it. Chaos can be understood as a "intermittently forced linear system" [41]. Following this, a forced linear system can be constructed from the time-series contained in V as follows:

$$\frac{d}{dt}\mathbf{v}(t) = \mathbf{A}\mathbf{v}(t) + \mathbf{B}v_r(t) \tag{3.22}$$

Where $v(t)$ are the first r-1 columns of V, and the last one, v_r, is imposed as the forcing element. Equation 3.22 represents the reduced model of the system, with all the non-linearity of its dynamics being encompassed in the forcing term v_r.

This method has been successfully applied to the Lorenz system, demonstrating its value to the modeling of chaotic systems.

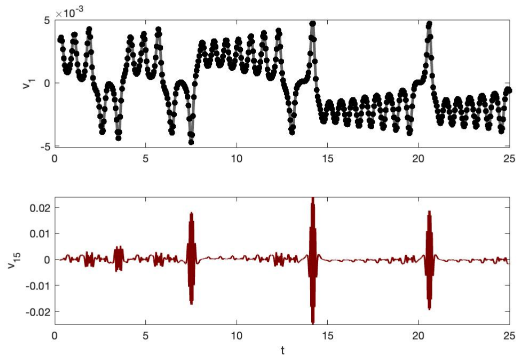


Fig. 3.13. HAVOK method: Lorenz system

For Lorenz, r = 15. The plot on the bottom of Figure 3.6. corresponds to the intermittent forcing term. It has been shown that the peaks on this figure corresponds to the change of lobes in the Lorenz system, which further supports the assumption that chaotic systems can be interpreted as linear systems in which an intermittent force is applied.

HAVOK is a technique useful for identifying reduced models, as in the example with Lorenz system shown before. By using this technique the most important properties of the dynamics are preserved and defined into a mathematical formula, that can later be used to study or even control the studied system.

Sparse Identification of Nonlinear Dynamics (SINDy)

SINDy is a method developed by Steven Brunton et all. [44] that seeks to discover governing equations of dynamical systems purely from measurement data, specifically from time series measurements.

From the time series obtained, a library of possible functions up to a certain power is built in the form of a matrix. Each column of said matrix corresponds to each one of the functions selected, evaluated in each row at a different time instant. The assumption in which this method stands is the fact that the derivatives of the time series obtained can be described by a few of the functions contained in the matrix multiplied by a series of constants.

As an illustrative example, let's assume 3 different time series, wich n data points each, have been recorded during an experiment: x, y and z. The need to also record the derivatives of the time series can be relaxed by calculating them directly from the time series itself by numerical differentiation. The user chooses a set of polynomials up to order five as its example library of functions, namely: $[1, x, y, z, x^2, xy, xz, y^2, yz, ..., z^5]$.

SINDy assumes the derivatives of the variables recorded can be described by a combination of some of the functions inside the library multiplied by some constants. The output of this algorithm is the set of constants which, when combined with the functions inside the library, provide an accurate description of the dynamics studied through a reduced model. Mathematically, it can be expressed as:

$$\begin{bmatrix} \dot{x}_1 & \dot{y}_1 & \dot{z}_1 \\ \dot{x}_2 & \dot{y}_2 & \dot{z}_2 \\ ... & ... & ... \\ \dot{x}_n & \dot{y}_n & \dot{z}_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & z_1 & x_1y_1 & ... & z_1^5 \\ 1 & x_2 & y_2 & z_2 & x_2y_2 & ... & z_2^5 \\ ... & ... & ... & ... & ... & ... & ... \\ 1 & x_n & y_n & z_n & x_ny_n & ... & z_n^5 \end{bmatrix} \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ ... & ... & ... \\ \epsilon_{n1} & \epsilon_{n2} & \epsilon_{n3} \end{bmatrix} \tag{3.23}$$

In order to ensure the sparsity of the constants obtained, the following procedure is implemented:

$$\epsilon = argmin||\Theta\epsilon' - x|| + \lambda||\epsilon'|| \tag{3.24}$$

Where λ is the sparsity constraint, and is left for the user to choose, depending on how sparse they desire the system to be.

For the Lorenz system, the SINDy algorithm with provide the following result:

$$\begin{bmatrix} \dot{x}_1 & \dot{y}_1 & \dot{z}_1 \\ \dot{x}_2 & \dot{y}_2 & \dot{z}_2 \\ ... & ... & ... \\ \dot{x}_n & \dot{y}_n & \dot{z}_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & z_1 & x_1y_1 & ... & z_1^5 \\ 1 & x_2 & y_2 & z_2 & x_2y_2 & ... & z_2^5 \\ ... & ... & ... & ... & ... & ... & ... \\ 1 & x_n & y_n & z_n & x_ny_n & ... & z_n^5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ -10 & 28 & 0 \\ 10 & -1 & 0 \\ 0 & 0 & -8/3 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \\ ... & ... & ... \end{bmatrix} \tag{3.25}$$

As in the case of HAVOK, SINDy also serves as a tool that finds a reduced model for the system. However, the advantage of SINDy is that it looks for the smallest reduced model available for the dynamics. As an illustration, HAVOK provided a mathematical formula with 15 variables, including the forcing parameter, for the Lorenz system. On the other hand, SINDy, from the extensive library of possible functions that it received, only kept those who surpassed the sparsity constraint set by the user for significance in the model. The disadvantage that SINDy presents, however, is the fact that the library of functions is selected by the user, and if the wrong library is introduced, the algorithm only provides the trivial solution with all the constants set to zero.

3.2.4. Chaos quantification techniques

There is not a general definition of chaos. However, some specific characteristics are attributed to chaotic systems, one of the most important ones being the sensitivity to initial conditions of the system. Small variations in those conditions for chaotic systems result in the trajectories differing exponentially as time goes by.

Numerous experiments whose anomalous behaviour was thought to be due to errors or noise in the samples have been restudied to determine that these non-linearities were due to the chaotic behaviour of their dynamics [45]. The application of these techniques to the field of plasma physics presents a huge potential in the study of turbulence.

Maximal Lyapunov Exponent

Chaotic systems, as explained before, present an extreme dependence on initial conditions, which normally results in unpredictability by using traditional techniques even if said systems are deterministic. This means that trajectories that were originally close in phase-space will diverge exponentially fast as time increases. The average exponent of this exponential increase is denominated as the Maximum Lyapunov exponent.

Let two different particles, that were originally close in the phase-space, be at a distance $\delta_o << 1$. After some time Δt, the distance will have grown to $\delta_{\Delta t}$ exponentially fast. The Maximum Lyapunov exponent can then be calculated according to [29]:

$$\delta_{\Delta t} = \delta_o e^{\lambda \Delta t} \tag{3.26}$$

Different types of motion can be identified according to the Maximum Lyapunov exponent obtained according to the following table:

TABLE 3.1. Different types of motion according to its Maximum Lyapunov exponent [Extracted from [29]]

Type of motion	**Maximum Lyapunov exponent**
Stable fixed point	$\lambda < 0$
Stable limit cycle	$\lambda = 0$
Chaos	$0 < \lambda < \infty$
Noise	$\lambda = \infty$

Lyapunov exponents are invariant under phase-space reconstruction, rescaling or shift. Based on this phenomenon, this technique can be applied to time-series obtained experimentally by using its phase-space reconstruction. Even though the reconstructed attractor will be based on a single trajectory, by choosings points whose temporal separation is bigger or equal to an estimated orbital period of the time-series, these points will be assumed to lie on different trajectories [46].

Lyapunov exponents are widely used to characterize chaotic systems, and in the field of plasma they have been employed in experiments regarding distributed beam-plasma systems [47] and Hall thruster electromagnetic field profiles [48] among others.

The 0-1 test for Chaos

The 0-1 test for Chaos is a simple technique introduced by Gottwald, G. A., and Melbourne, I. in 2004 [49] that separates chaotic systems from regular dynamics. One of the best features of this technique is, apart from its simplicity, the fact that it only requires a time series as an input in order to perform the analysis.

Imagine we have a time series $x(t)$ with N points, from it we construct the following 2 vectors:

$$p_c(i+1) = p_c(i) + x(i) * cos(c * i) \tag{3.27}$$

$$q_c(i+1) = q_c(i) + x(i) * sin(c * i) \tag{3.28}$$

Being c a fixed value between 0 and 2π. If the system contained regular dynamics, the motion of p and q would remain bounded in a regular pattern, however, if the dynamics were chaotic, a Brownian-like motion would be observed. To illustrate this notion, the Logistic Map presented in the Phase Portraits section will be used.

If $\mu = 3.55$ the dynamics of the Logistic Map would be regular non-chaotic, however, if the parameter was chosen to be $\mu = 3.97$ the system would behave chaotically. This can be clearly appreciated in the following figure:

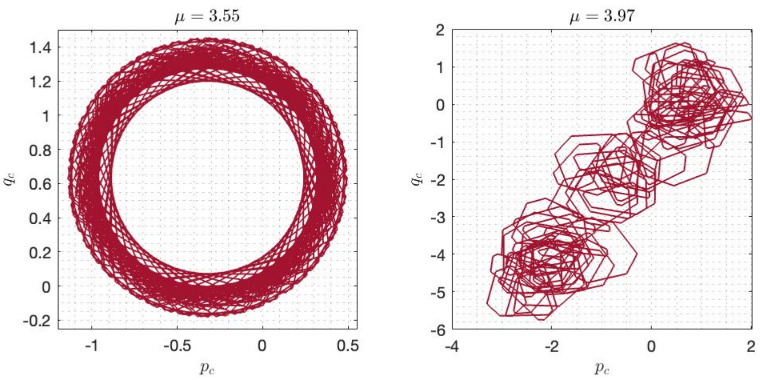


Fig. 3.14. Behaviour of q_c as a function of p_c for: a regular system (left plot, $\mu = 3.55$) and a chaotic system (right plot, $\mu = 3.97$).

Once p_c and q_c have been calculated, the time-averaged mean square displacement can be obtained according to:

$$M_c(i) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} ([p_c(j+i) - p_c(j)]^2 + [q_c(j+i) - q_c(j)]^2) \tag{3.29}$$

This variable presents a lot of oscillations, so a modified mean-square was proposed:

$$D_c(i) = M_c(i) - V_{osc}(c, i) \tag{3.30}$$

$$V_{osc}(c, i) = (Ex)^2 \frac{1 - cos(c * i)}{1 - cos(c)} \tag{3.31}$$

$$Ex = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} x(j) \tag{3.32}$$

A comparison showing both M_c and D_c is shown below to illustrate this:

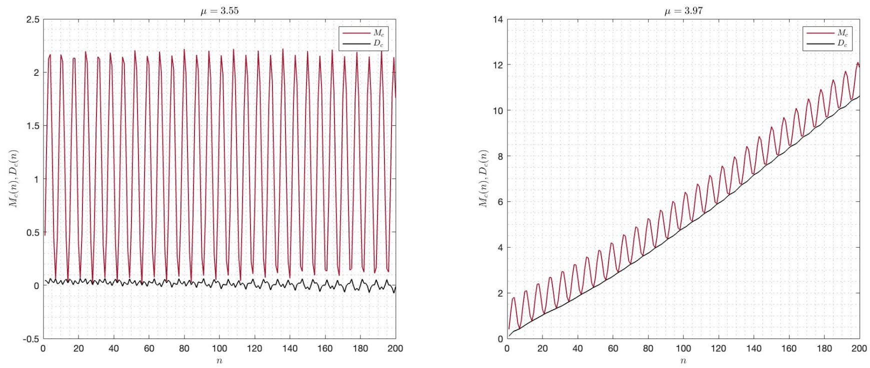


Fig. 3.15. Behaviour of M_c and D_c as a function of n (n = i in the formulas): a regular system (left plot, $\mu = 3.55$) and a chaotic system (right plot, $\mu = 3.97$).

As it can be seen in these graphs, if the system is regular, there is no growth on the mean of both M_c and D_c. On the other hand, chaotic systems are characterized by a linear growth of M_c and D_c. Having shown the advantage of studying D_c instead of M_c, the test for chaotic behaviour will be performed using only D_c.

Finally, in order to better illustrate if the system were regular or chaotic, the following variable was introduced:

$$K_c = \frac{cov(\epsilon, \Delta)}{\sqrt{var(\epsilon)var(\Delta)}} \tag{3.33}$$

$$var(x) = cov(x, x) \tag{3.34}$$

$$cov(x, y) = \frac{1}{q}\sum_{j=1}^{q}(x(j) - \hat{x})(y(j) - \hat{y}) \tag{3.35}$$

$$\hat{x} = \frac{1}{q}\sum_{j=1}^{q} x(j) \tag{3.36}$$

Being q the length of the vectors studied, $\epsilon = (1, 2, ..., i)$ and $\Delta = (D_c(1), D_c(2), ..., D_c(i))$. The variable K_c will be 0 if the system is regular, and 1 if the system is chaotic. This can be explained by examining the trajectories of the studied systems. In the regular dynamics case, the trajectories are normally bounded and its mean-squared displacements remains constant, but for the chaotic case the system normally behaves like a Brownian-motion and its mean-squared displacement grows on time. K_c captures this behaviour and computes a 0 if there is no growth, and a 1 is there is [35]. To continue with the example studied in this section, the following graphs have been created:

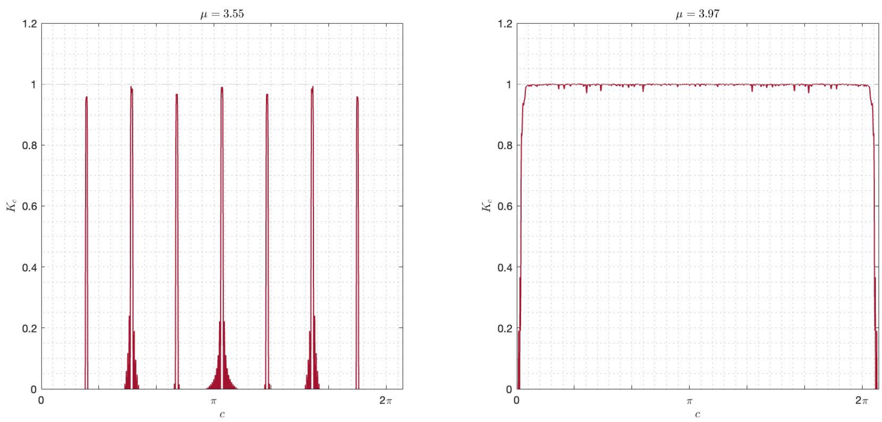


Fig. 3.16. Behaviour of K_c as a function of c for: a regular system (left plot, $\mu = 3.55$) and a chaotic system (right plot, $\mu = 3.97$).

Some resonances are captured by the method for the regular case, but overall the trend of $\mu = 3.55$ tends to stay around the value 0. On the other hand, the case for $\mu = 3.97$ clearly represents the chaotic behaviour of the system as K_c takes the value 1 for almost all the values of c.

3.2.5. Statistical techniques

Intermittency

A signal is intermittent when it displays some activity only during a fraction of the time, occurring in bursts. The Lorenz attractor is again a good example to illustrate this phenomenon. It displays almost two periodic orbits in which the particles travel between them, which could be defined as an intermittent periodic cycle. This can be seen in the following figure, in which the different orbits corresponding to the two lobes of the Lorenz attractor are clearly separated in the top and bottom fluctuations pictured.

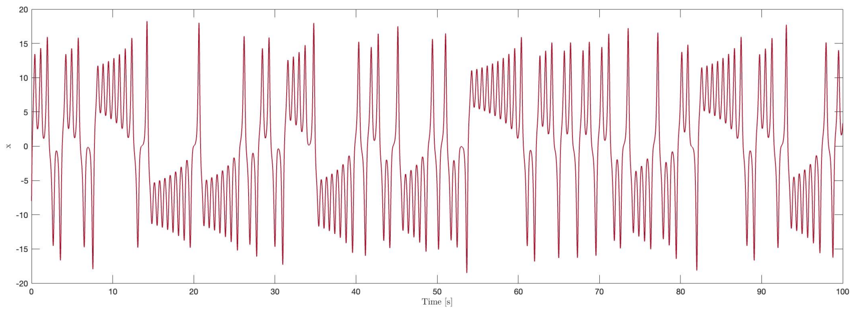


Fig. 3.17. Intermittency on Lorenz system

Having defined what intermittency is, now let's focus on how to quantify it, different approaches have been developed over the years in order to do this. Batchelor postulated in 1953 [50] that one could model such systems as a still phase, interrupted by burts, modeled as Gaussian random fluctuations for a fraction of time γ. Kurtosis was easy to calculate for such signals, being Kurtosis a description of the shape of the probability distribution of said signals according to the formula [51]:

$$K = \frac{\mu_4}{\sigma} \tag{3.37}$$

Being σ its standard deviation and μ_4 the fourth moment about the mean of the sample. Assuming the variables studied, X, to be random, with a probability distribution defined by F, the fourth moment of X is defined as:

$$\mu_4(X) = \int_{-\infty}^{\infty} x^4 F(x) \tag{3.38}$$

Relating K to the variable γ it was obtained that: $K = 3/\gamma \geq 3$, so Kurtosis was proposed as a general measure of intermittence. However, since K is also a measure of the shape of the pdf of the signal, it was unclear when K different than 3 corresponded to a non-Gaussian pdf and when it corresponds to intermittence [50].

Frisch noted in 1995 [52] that in real systems, signals are rarely still. He suggested applying a high-pass filter to the data (thus assuming that the bursts are high-frequency).

Then, K would depend on the cut-off frequency of the filter employed. It was postulated that if $K(fc)$ was not constant, then the signal was intermittent.

A more reliable measure of intermittency was later introduced based on chaos theory by Meneveau [53] [54] and has already been proven useful for the study plasma physics [21]. This approach is based on the "box-counting" method. Given a time series $X = (x_1, x_2, ..., x_N)$, the following quantity can be calculated:

$$\epsilon(1, i) = \frac{(x_i - < x_i >)^2}{< (x_i - < x_i >)^2 >} \tag{3.39}$$

Where $< x_i >= (\sum_{i=1}^{N} x_i)/N$. This measure can be averaged over sub-blocks of data with length n<N, as follows:

$$\epsilon(n, i) = \frac{1}{n} \sum_{j=0}^{n-1} \epsilon(1, i + j) \tag{3.40}$$

By calculating the q-moments of $\epsilon(n, i)$, for a given range of n-values, these moments are expected to behave according to:

$$< \epsilon(n, i)^q > \propto n^{-K(q)} \tag{3.41}$$

Where $K(1) = 0$. The parameter $C(q)$ can then be defined as:

$$C(q) = \frac{K(q)}{q - 1} \tag{3.42}$$

And finally, the intermittency parameter, C(1), can be defined, due to the singularity that occurs for $C(q)$ at $q = 1$ as:

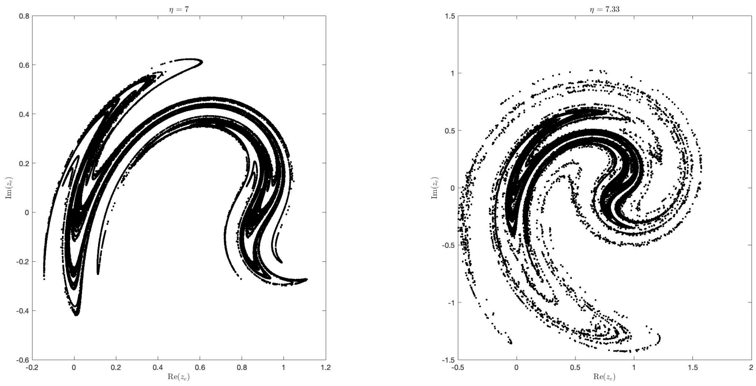
$$C(1) = \frac{dK(q)}{dq}\Big|_{q=1} \tag{3.43}$$

K and C(1) present the improvement when compared to Kurtosis that they contain information related to time, which in the statistical treatment of the data relating the Kurtosis was completely lost.

Both K and $C(1)$ serve as quantifiers of the intermittence present in a sample, however, it has been shown that $C(1)$ is more reliable than K. A simple example was presented for intermittency study by B. P van Milligen et all. [55] for the Ikeda map. This map is defined as:

$$z_{n+1} = a + b z_n e^{ik - \frac{i\eta}{1+|z_n|^2}} \tag{3.44}$$

With a = 0.85, b = 0.9, k = 0.4. By varying the values of the parameter η it was discovered that at $\eta_c = 7.26884894$, the system went from a regular behaviour between 0 and 1 to showing bursts of intermittency. The following figures illustrate this phenomenon for two different values of η: one below the critical value, 7, and one above, 7.33.



(a) Non-intermittent behaviour. (b) Intermittent behaviour.

Fig. 3.18. Ikeda map behaviour for different values of η.

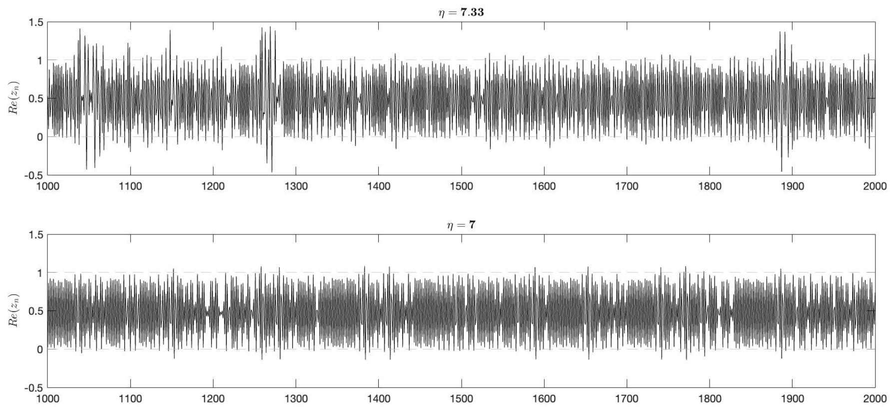
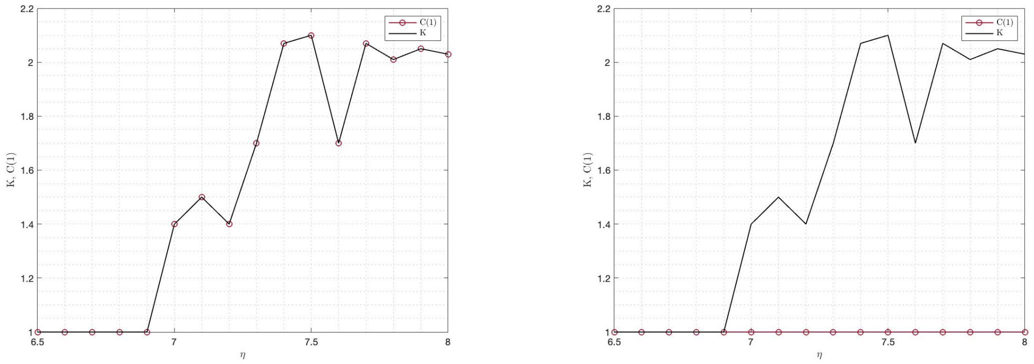


Fig. 3.19. Time series for Ikeda map. Bottom plot shows clear range defined between 0 and 1. Top plot represents intermittent behaviour, with bursts occuring at different times during sample.

To demonstrate the superiority of C(1) over K, a sample obtained from the Ikeda map was studied for both K and C(1), then, the data points were re-ordered in ascending order removing all intermittency from the set. C(1) dropped to 1 as expected, however K still showed the same pattern as in the previous case. The results are shown in the following figures:



(a) Experimental sample.

(b) Re-ordered sample with values in increasing order.

Fig. 3.20. Comparison of K and C(1) as intermittency quantifiers. Example extracted from [55].

Turbulence may contain big turbulent eddies, which break up and become smaller ones in a cascade sense. These smaller eddies will behave then as intermittent systems and can be monitored using the $C(1)$ parameter. The field of intermittence in plasma propulsion is yet to be explored, but seeing the success rate that the $C(1)$ parameter has provided when detecting intermittent behaviours in fusion plasma experiments, these technique presents promising prospects to study the turbulence in plasma thrusters [56] [57].

Transfer Entropy

Transfer Entropy is defined as the gain in information about a future state of time series $x(t)$ when both the past history of $x(t)$ and an additional time series, $y(t)$, are considered comparing it with the gain that would be obtained by studying the past history of $x(t)$ alone. This technique was developed by Thomas Schreiber to quantify the statistical causality between two signals and additionally identify the directionality of said influence [28].

Before Transfer Entropy was introduced, Standard Time Delayed Mutual Information was used to study the coherence between systems. Based on the explanation about Mutual Information from section 3.1 and specifically on equation 3.7, it can be concluded that the Mutual Information technique presents a symmetry under the interchange of processes I and J; if the order was to be reversed, the result would be exactly the same and no information would be obtained regarding the direction of influence between them. In addition, Transfer Entropy has the advantage of providing directionality when compared with other techniques such as the correlation. By providing directionality, not only the causality between two signals can be found but in addition, which signal is producing the influence can also be identified.

Furthermore, Transfer Entropy is also different from the cross-correlation technique explained in Section 3.1.2. Cross-correlation presents a maximum when both signals studied are equal, whereas the Transfer Entropy returns a 0 value if both signals are the same. This is due to the fact that if the signals are equal, no additional information is gained by studying the second signal in comparison to the first one alone.

The entropy rate is defined as the average number of bits needed to encode one additional state of the system if all previous states are known [28]:

$$h_I = -\sum p(i_{n+1}, i_n^{(k)}) log_2 \frac{p(i_{n+1}^{(k+1)})}{p(i_n^{(k)})} \tag{3.45}$$

Being $i_n^{(k)} = (i_1, i_2, ..., i_n)$. By generalizing h_I to two different processes, I and J, Schreiber arrives at the final equation to describe the Transfer Entropy:

$$T_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(k)}) log_2 \frac{p(i_{n+1}|i_n^{(k)}, j_n^{(k)})}{p(i_{n+1}|i_n^{(k)})} \tag{3.46}$$

If the processes are independent, then the first conditional probability should not depend on J and the 2 expressions inside the logarithm would be identical, yielding a null value for the Transfer Entropy [29].

For the purpose of data analysis, the expression used for the Transfer Entropy when studying two different time series, x(t) and y(t), is the following:

$$T_{Y\to X} = \sum p(x_{n+1}, x_n^{(k)}, y_n^{(k)}) log_2 \frac{p(x_{n+1}|x_n^{(k)}, y_n^{(k)})}{p(x_{n+1}|x_n^{(k)})} \tag{3.47}$$

Schreiber defined the kernel estimation as a suitable approximation for the probability distribution of the time series studied. A step kernel is used that yields 1 when the state studied is found and 0 elsewhere, this procedure is performed for all possible states and is then divided by the total number of samples studied [28]. In this Thesis, 3 bins of probability will be used: positive, negative and null.

A simple approach that estimates the statistical significance of the Transfer Entropy, as explained before, is to compute the TE between two random noise signals of the same length. This computation is performed as a function of the length of the signal, N. Figure 3.21 shows this calculation.

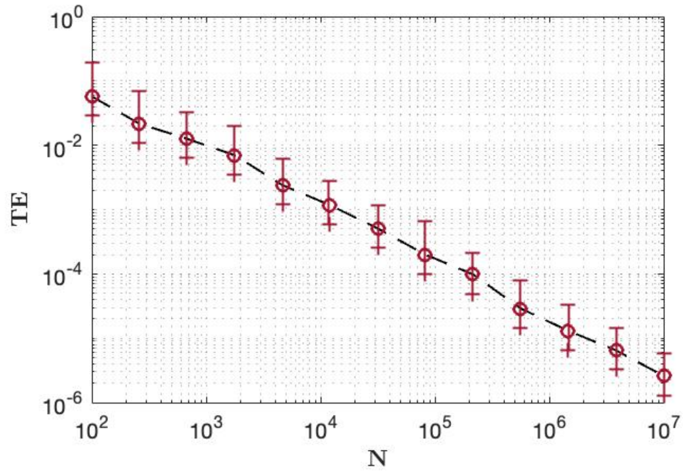


Fig. 3.21. Transfer Entropy for two random noise signals as a function of the sample length, N. Each point is calculated as the average of 100 independent simulations, with the error bar representing the variation over said simulations.

As stated before, Transfer Entropy provides a quantification of the impact of a variable on another and the direction of said impact. This feature is relevant in the study of Plasma Propulsion because it provides a way to track small propagations in highly turbulent systems. Additionally, Transfer Entropy has already been employed and proven useful in the study of plasma fusion for heat transfer in experiments conducted in the TJ-I and W7-X stellarators and in the Joint European Torus (JET) [58] [59].

4. DATA OVERVIEW

All the pertaining data used for the analyses shown on this project was freely-provided by the EP2 research group. The HET simulation constitutes the main focus of study of this Thesis, and the HPT experimental data has also been included and analyzed in order to extend the implementation of these techniques as an additional exploration.

It should be noted that the author of this Thesis did not intervene neither in the HPT Experiment nor in the HET simulation, while all analyses presented in the following chapter have been performed originally for this project. A description of both the experiment and the simulation is presented in this section in order to provide a better understanding of the source of the data analyzed.

4.1. Helicon Plasma Thruster Experiment

The data obtained for the Helicon Plasma Thruster represents the first experiment carried by the EP2 Research Group on this motor. This experiment was performed with the objective of performing a preliminary study of the plume of the HPT. It should also be kept in mind that this data will be mainly used by the EP2 group for calibration porpuses and is just a first attempt into the characterization of this new motor.

The HPT experiment was performed in the EP2-UC3M laboratories. The thruster was set with Krypton at $450W$ of power and $30sccm$ of mass flow rate.

Two non-compensated Langmuir probes facing the nozzle were employed. They were positioned symmetrically with respect to the vertical centerline of the thruster and $20mm$ over the horizontal centreline. The probe tips, with a length of $2mm$ and a radius of $0.25mm$, were positioned at $8 - 10mm$ from each other.

Two types of data were obtained, even though the floating potential will be the focus of the analysis in this project:

1. The floating potential

   The configuration explained before was directly used in this case. Seven measurements at different positions as shown in Table 4.1 were obtained. In order to record this, the probes are left floating and the electric potential is recorded.

2. The ion saturation current

   Data was acquired by using the configuration previously mentioned alongside a $100k\Omega$ resistor for each probe. The positions where the measurements were taken are also shown in Table 4.1, in this case only two different positions were studied.

To record the ion saturation current, the probes are biased at low voltages and then the current between the probe and the ground is recorded.

TABLE 4.1. Measurement positions, from thruster to nozzle, performed for the different experiments. All values expressed in *mm*.

	1	2	3	4	5	6	7
Float	400	350	300	250	200	150	100
Isat	200	100	-	-	-	-	-

The wiring, both in-chamber and out-of-chamber, was chosen to be unshielded in order to avoid shortcutting all the high-frequency oscillations through the wire itself and all the shielding of the coax.

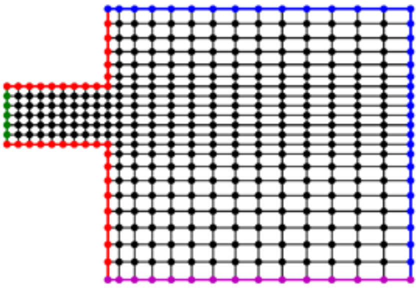
4.2. Hall-Effect Thruster Simulation

The Hall-Effect Thruster simulated was generated by means of a hybrid algorithm composed of Particle-In-Cell (PIC) and a fluid model. This hybrid code was created by the EP2 Research Group with the objective of reducing computation time and costs when handling plasma simulations and has already been proven successful in the simulation of Hall-Effect Thruster environments [60].

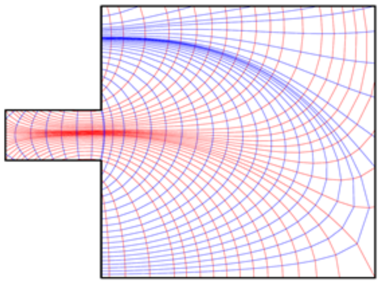
The PIC algorithm follows trajectories of charged particles in electromagnetic fields for a provided mesh. The code simulates the motion of each particle and calculates all associated plasma parameters, such as density or electric potential, from the position and velocity of these particles. PIC simultaneously solves Newton-Lorentz's force and Maxwell's equations [61]. The PIC algorithm is used for heavy particles, such as neutrons. The mesh used in this case, shown on Figure 4.1(a), presents a higher concentration of points inside the thruster discharge region than in the rest of the nozzle in an attempt to reduce the noise of the simulation.

On the other hand, electron particles would be very costly to compute using the PIC method, especially due to their short-time characteristic behaviour. Following this reasoning, the electrons inside the simulation are modelled following a fluid model. For this case, the mesh employed is aligned with the magnetic field of the thruster, as shown on Figure 4.1(b). By making the mesh follow the magnetic lines, numerical diffusion is minimized and the accuracy of the simulation is greatly improved.

Visual images of the meshes generated for the Hall-Effect thruster simulation are provided below:



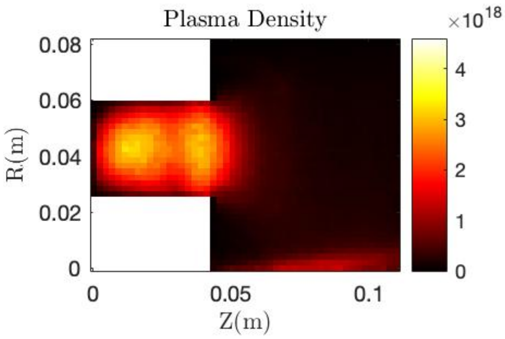
(a) PIC model.



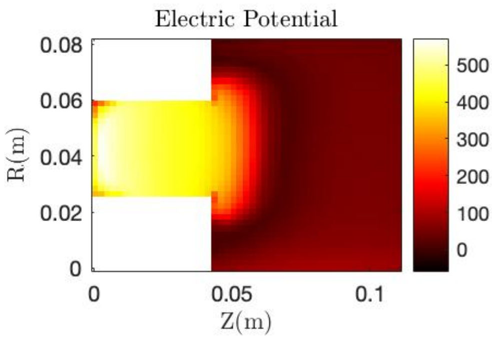
(b) Fluid model.

Fig. 4.1. Generated meshes for HET simulation.

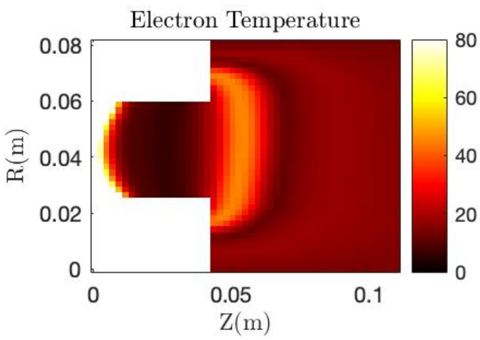
Some of the results generated for the 2D grid are also provided here in order to better define the type of outputs obtained from this simulation:



(a) Plasma density.



(b) Electric potential.



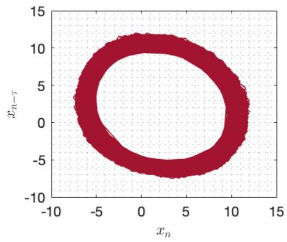
(c) Electron temperature.

Fig. 4.2. Outputs of HET simulation.

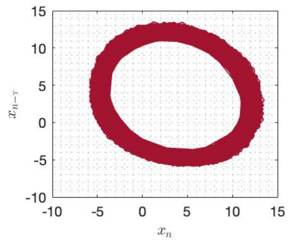
5. RESULTS AND DISCUSSION

5.1. Helicon Plasma Thruster Experiment

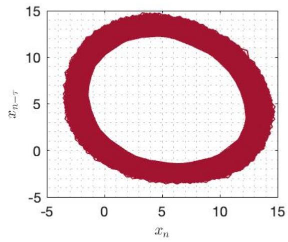
By performing a preliminary study of the data to be analyzed, it was discovered that the probes at positions 4 through 7 were saturated during the experiments, so these sets of data had to be discarded. The saturation can be observed in the following figures:



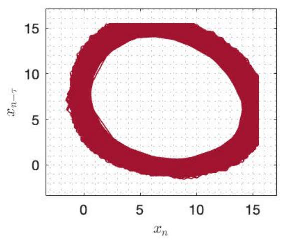
(a) Position 1.



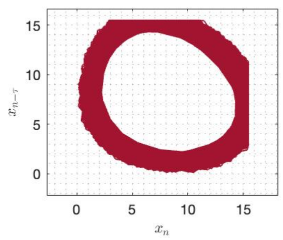
(b) Position 2.



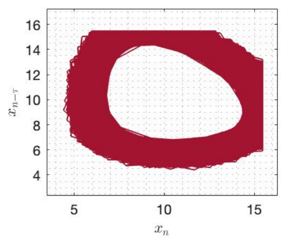
(c) Position 3.



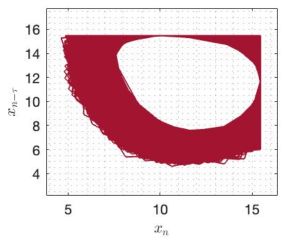
(d) Position 4.



(e) Position 5.



(f) Position 6.



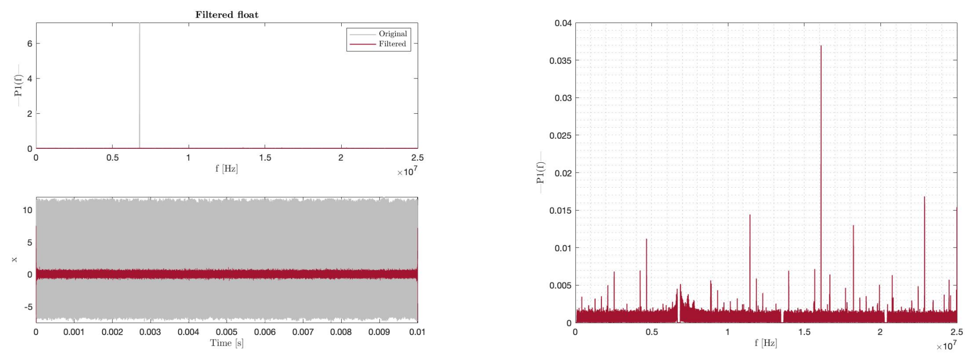
(g) Position 7.

Fig. 5.1. Saturated probes identification through Poincaré graphs. Analysis performed on unfiltered data sets.

Based on this reason, only positions 1 through 3 will be analyzed in this section. Additionally, all data sets have been filtered twice, as it will be explained in the following section.

5.1.1. Pre-processing

It was known before-hand that the frequency of the Helicon Plasma Thruster, $f = 13.56MHz$, along with the frequency corresponding to the wiring inside the chamber, $f = 50Hz$ would be present in the recorded signals. Since these frequencies would bury all the non-linearities of the plasma, these frequencies along with their corresponding harmonics were filtered out as shown in the figures below. Furthermore, after this processing a second filter was applied in order to reduce the amount of noise present in the measurements.



(a) Time series and FFT of signal before and after first processing.

(b) FFT of signal after first processing.

Fig. 5.2. Initial pre-processing of HPT signals.

5.1.2. Embedding dimensions

The first step prior to applying any analysis on the data sets was to obtain the embedding dimensions corresponding to the different variables to be studied with the Time Delayed Mutual Information and the False Nearest Neighbours techniques.

The remaining variables that will be studied for the floating potential are those corresponding to position 1, 2 and 3, with channel 1 and 2 for the 3 cases. The results are illustrated in the following table:

TABLE 5.1. Embedding dimensions for Helicon Plasma Thruster Experiment

	Position 1		**Position 2**		**Position 3**	
	Ch. 1	**Ch. 2**	**Ch. 1**	**Ch. 2**	**Ch. 1**	**Ch. 2**
m	5	4	4	3	4	5
tau [ns]	10	20	10	40	10	10

In order to better compare the results obtained, and since using a higher embedding dimension does not greatly affect the accuracy of the results obtained, all analyses presented in this section will use $m = 5$.

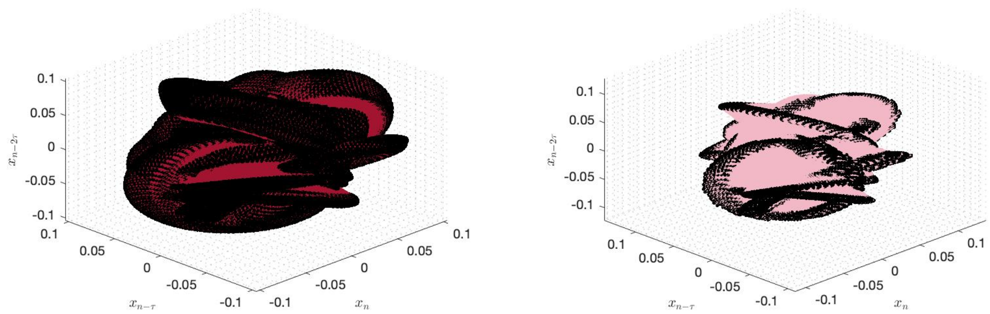
5.1.3. Phase Portraits

Phase Portraits for the different positions and channels were generated in order to see if any defined structure could be seen for the floating potential. By analyzing the Poincaré graphs obtained for the different positions and channels, some complex attractors were found. Since no simple structure which could be used for modelling or control purposed was found, only a few illustrative cases will be shown.

Channel 1

What is clear from inspecting the figure obtained for Channel 1 is that the dynamics of the floating potential seems to be confined to this complex attractor, with no points laying outside of it. Figure 5.3(a) shows defined orbits in which trajectories encompass each other.

The red lines represent the trajectories that join the points studied, and said points are represented in black, it can be seen that there are surfaces in the complex attractor where points are never found.



(a) Position 1. N = 500003 points.

(b) Position 1. N = 20000 points.

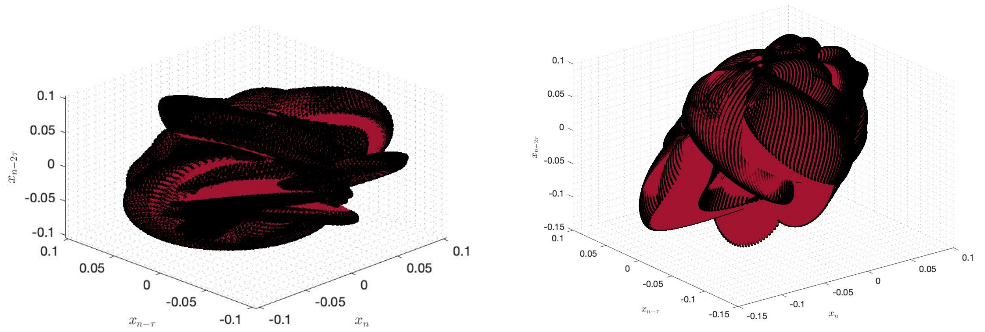
Fig. 5.3. 3D Poincaré graphs for floating potential in channel 1. Time lag used for study: $\tau = 2x10^{-8}s$.

Since the attractor found presents such a complicated behaviour, a second figure with less points studied was generated in order to better illustrate the paths followed by the particles (Figure 5.3(b)).

The data points seem to be confined to specific paths, travelling from one to the other as represented by the pink lines, but never found outside of them. This can be either due to the fact that because of the sampling frequency used the points always happened to be in these positions, or to the possibility that the floating potential can only exist in these particular states represented on the figure.

Channel 2

Since the probes corresponding to channel 1 and channel 2 were placed symmetrically with respect to the center of the thruster, the results obtained from them should be expected to be the same, or at least pretty close. However, this was not the case. As it can be appreciated from Figure 5.7 the resultant dynamics of the floating potential recorded are not the same.

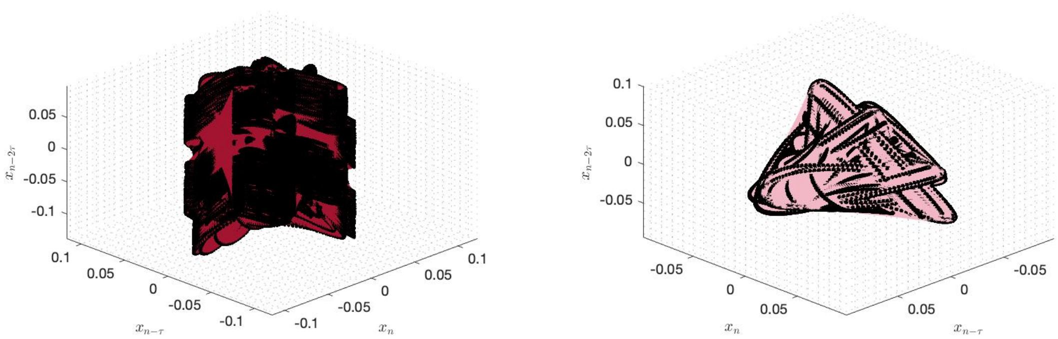


(a) Position 1. Channel 1. (b) Position 1. Channel 2.

Fig. 5.4. Comparison of 3D Poincaré graphs for floating potential in position 1 (400 mm from thruster). Time lag used for study: $\tau = 2x10^{-8}s$.

This difference in the results obtained is likely due to noise, a defect in one of the probes used for the experiment or even some interference between the wiring and the probes for one of the channels.

An intereseting result was found for Position 3 in channel 2, shown on Figure 5.7. A "tower" of triangles seemed to appear, being the triangles formed by the trajectories of the studied points.



(a) Position 3. N = 500003 points. (b) Position 3. N = 20000 points. Top view.

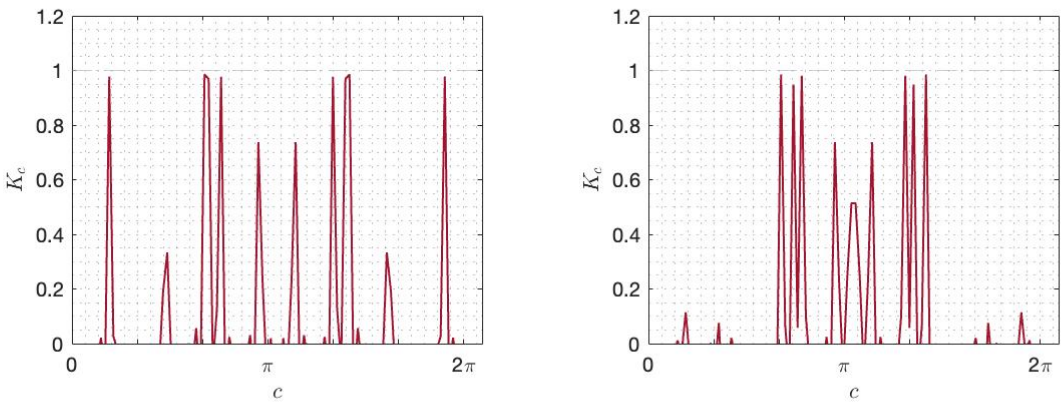
Fig. 5.5. Comparison of 3D Poincaré graphs for floating potential in channel 2. Time lag used for study: $\tau = 2x10^{-8}s$.

The orbits shown on figure 5.5(b) clearly represent the states studied being confined to specific configurations inside this reconstructed phase space. This view corresponds to the "top" of the tower of orbits that appeared at position 3. Trajectories followed by the data points being organized in a complex and self-similar geometry is characteristic of chaotic dynamics [62], so the 0-1 test for Chaos was applied to these data sets to further study this possibility.

To conclude this section, the finding of an attractor which is bounded and shows structure, even if that structure is really complicated, provides some knowledge about the dynamics of the floating potential dynamics. Even if the dynamics of the system were to be chaotic, meaning there would be local unstabilities subject to initial conditions, the existance of an attractor ensures global stability. In other words, even if variations on the initial conditions were introduced, making the evolution of the system diverge exponentially, the existence of the attractor ensures the divergence would not be infinite, it would be subjected to the size of said attractor [34].

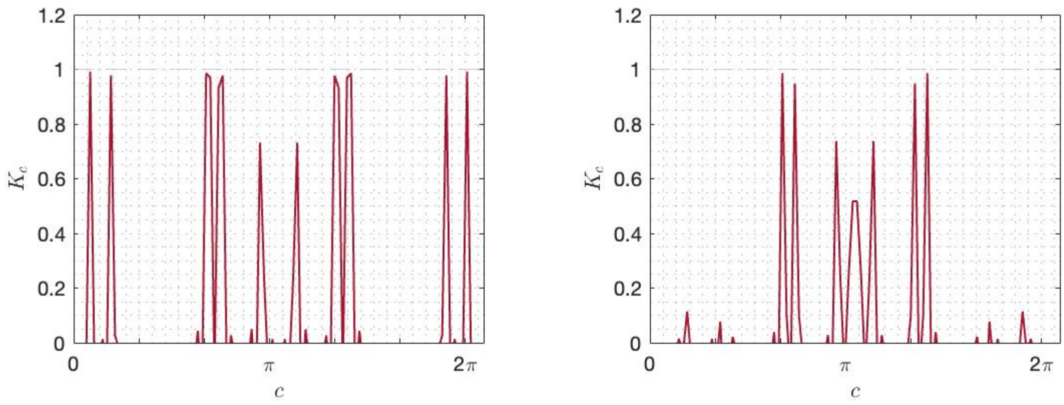
5.1.4. The 0-1 test for Chaos

Folling the findings of the previous sections, the 0-1 test for Chaos was applied to all the positions and channels studied, in order to see if chaos was present in the dynamics of the floating potential.



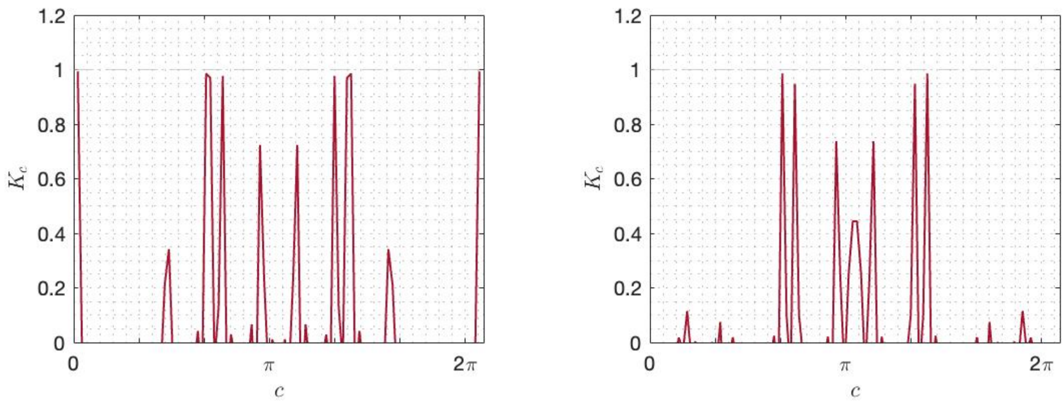
(a) Channel 1. (b) Channel 2.

Fig. 5.6. The 0-1 test for Chaos performed for the floating potential. Position 1 (400 mm from thruster).



(a) Channel 1. (b) Channel 2.

Fig. 5.7. The 0-1 test for Chaos performed for the floating potential. Position 2 (350 mm from thruster).



(a) Channel 1. (b) Channel 2.

Fig. 5.8. The 0-1 test for Chaos performed for the floating potential. Position 3 (300 mm from thruster).

As it can be seen in the different figures, the floating potential only shows values of $K_c = 1$ for specific values of c, so chaos is unlikely to be present in the studied data. This is reasonable, since even though the complex attractors recoverered through the Poincaré method resembled chaotic behaviour in their geometry, clear defined orbits could be seen inside them. What this means is that while the point is contained in one of these defined orbits, it is not behaving chaotically or suddenly changing from a state to the other. This can be further explained by thinking of the Lorenz attractor. This chaotic system contains two quasi-periodic orbits in which the particles travel. Lorenz is one of the most well-known examples of chaotic systems in the literature, however it can also be thought of as two quasi-periodic systems being intermittently forced to change, as explained in Section 3.2.3. In fact, if the 0-1 test for Chaos is applied to the Lorenz system, the following result is retrieved:

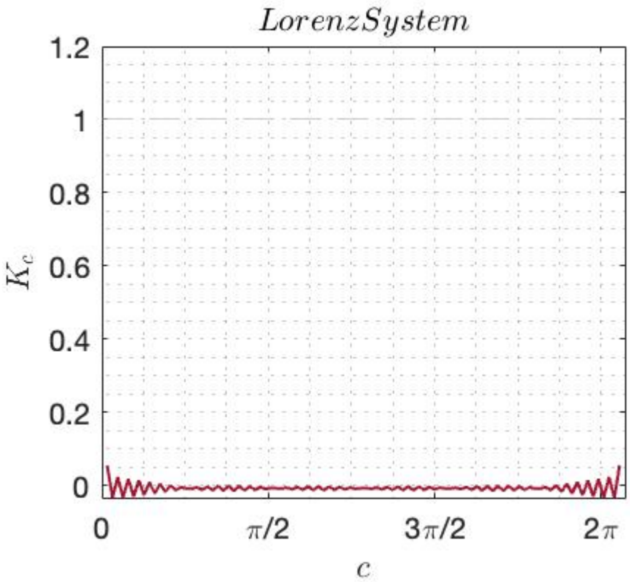
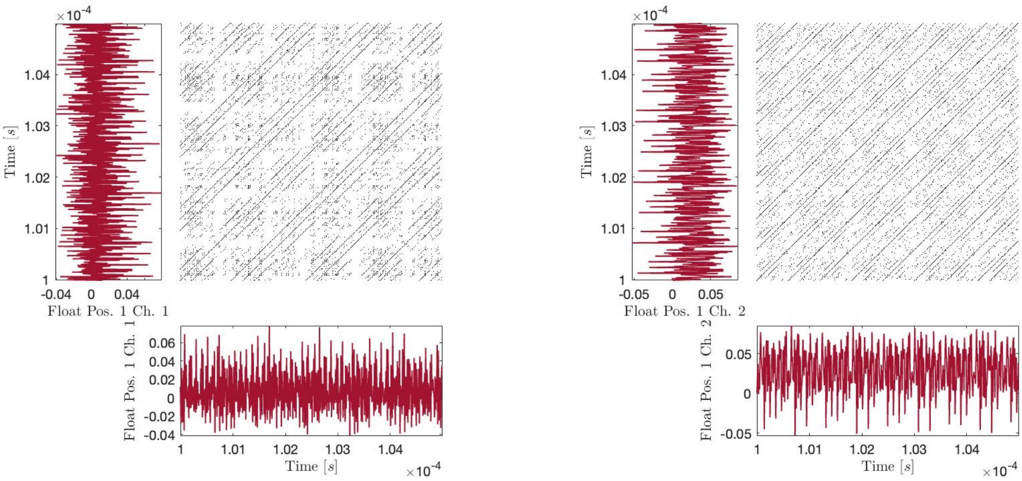


Fig. 5.9. The 0-1 test for Chaos performed for Lorenz system.

This results point in the direction of the 0-1 test for Chaos only detecting a specific type of chaos, more disorderly than the one presented for Lorenz System.

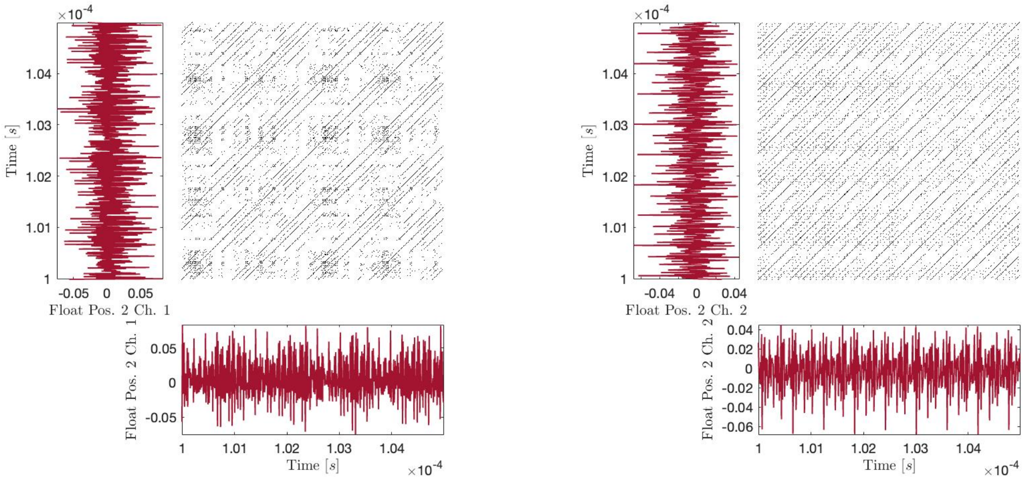
5.1.5. Recurrence Plots

Having chosen the optimal embedding dimension chosen to be $m = 5$, the following recurrence plots for the different positions and channels were generated.



(a) Channel 1.
(b) Channel 2.

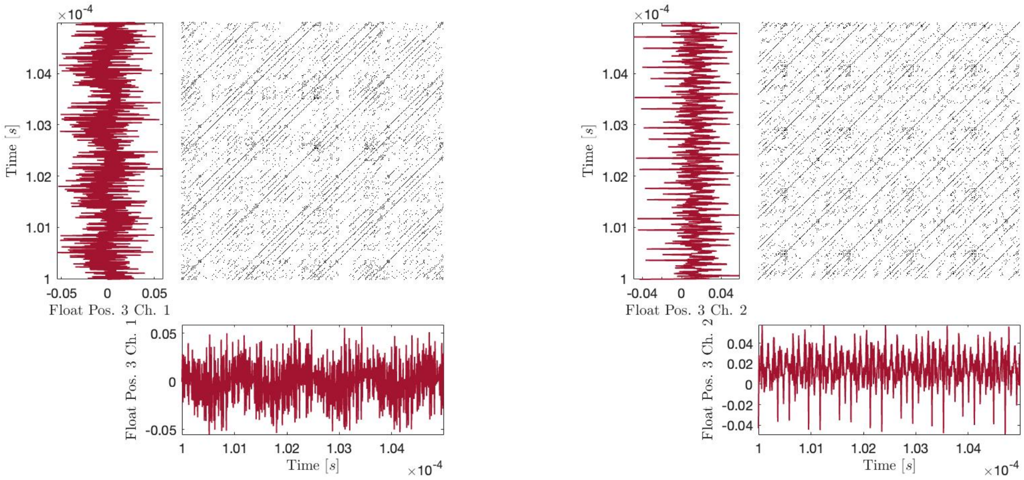
Fig. 5.10. Recurrence plots for floating potential at Position 1 (400 mm from thruster).



(a) Channel 1.

(b) Channel 2.

Fig. 5.11. Recurrence plots for floating potential at Position 2 (350 mm from thruster).



(a) Channel 1.

(b) Channel 2.

Fig. 5.12. Recurrence plots for floating potential at Position 3 (300 mm from thruster).

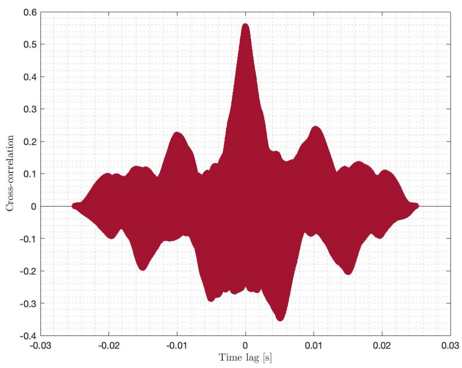
As it can be seen on the figures, there are a lot of scattered points in the plots, which may be due to bad filtering of the signals. However, it is interesing to note that, especially the samples corresponding to channel 1, resemble the irregular chessboard texture characteristic of chaotic systems explained in Section 3.2.2. These square clusters where the recurrences appear (represented in black on the figures), correspond to the particles coming near each other in the 5th dimensional reconstructed space. The white vertical and horizontal lines that can also be appreciated on the figures corresponding to channel 1 correspond to time spans in which almost all particles seem to be separated from each other in the embedded space.

The fact that different dominant frequencies seem to be dominant in channels 1 and 2 is also worth highlighting. For the case of channel 1, a periodic pattern of $1x10^{-6}s$ seems to appear, and in the case of channel 2, a faster periodic pattern of $7x10^{-7}$ becomes apparent. These different behaviours align with the findings in the last two sections, where discrepancies between the data recorded by Channel 1 and Channel 2 were found.

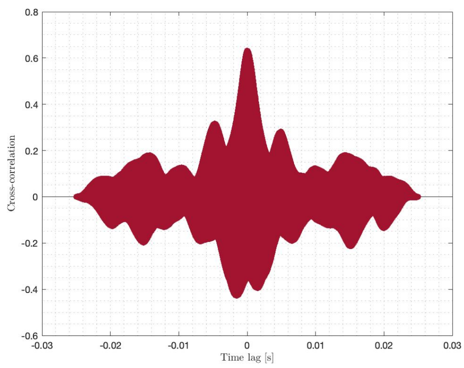
5.1.6. Cross-Correlation

The cross-correlation, introduced in Section 3.1.2., was employed to look for any relationship between the channels at the different positions. For each position, the experiments were repeated several times, so the cross-correlation shown in the figures corresponds to the concatenation of these repetitions, in order to make the statistical calculations for that specific location more accurate.

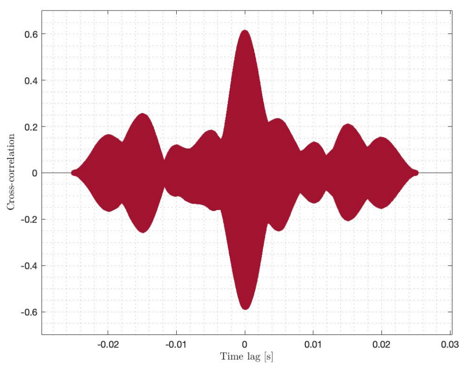
As it can be seen from the figures, some degree of similarity was found between the different channels for all the positions studied. The discussion of the relationship between the channels, along with a comparison with the Transfer Entropy in order to show the advantages of using the later one, is developed in the following section.



(a) Position 1.



(b) Position 2.



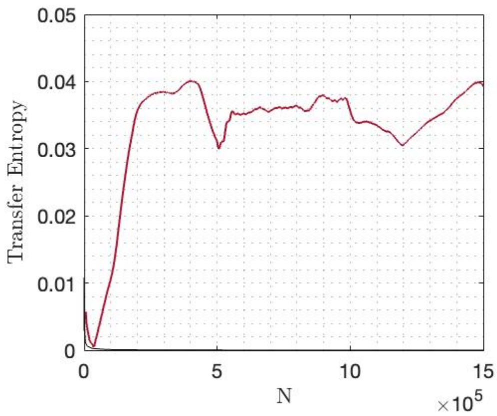
(c) Position 3.

Fig. 5.13. Normalized cross-correlation between probes at the three positions studied.

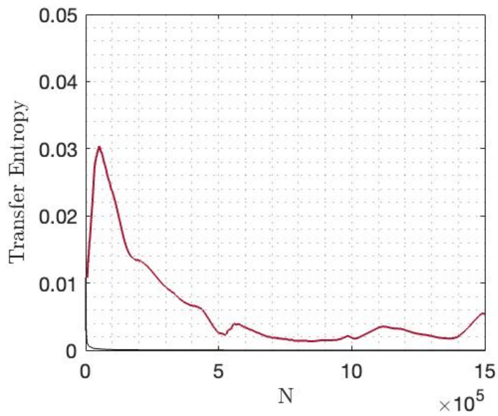
5.1.7. Transfer Entropy

In order to see if there was any azimuthal transfer of information, the Transfer Entropy technique was applied to both probes at the three different positions available.

The length of just one sample for the Hall-Effect Thruster Simulation is $N = 500003$. For the case of position 1 the experiment was repeated three times, and for position 2 and 3 five times. For the three positions this produces a statistical relevance threshold around 10^{-5} according to Figure 3.21. It should be noted that all transfer entropies obtained are well above the statistical significance threshold.

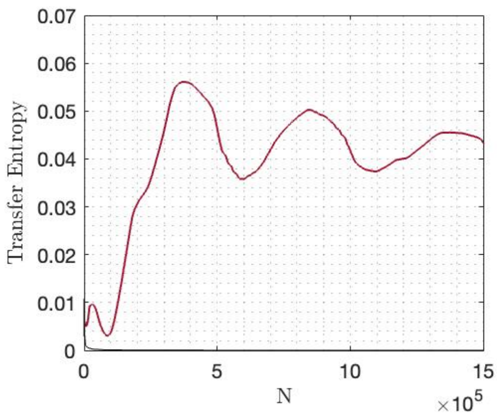


(a) Influence of Channel 1 over Channel 2.

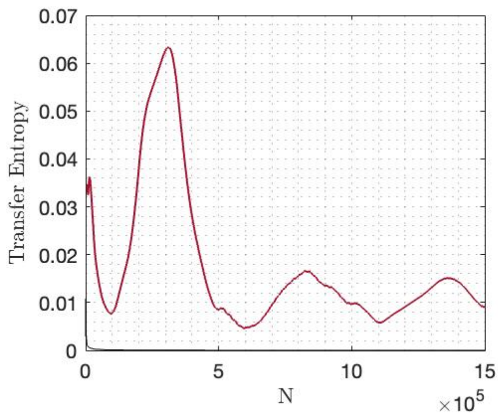


(b) Influence of Channel 2 over Channel 1.

Fig. 5.14. Transfer Entropy between probes at Position 1 (400 mm from thruster).

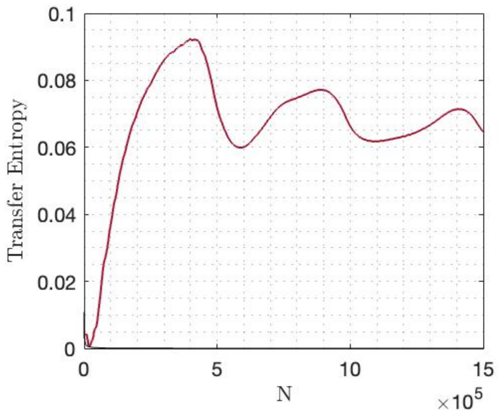


(a) Influence of Channel 1 over Channel 2.

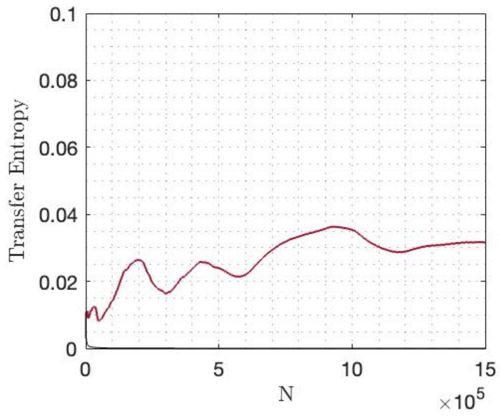


(b) Influence of Channel 2 over Channel 1.

Fig. 5.15. Transfer Entropy between probes at Position 2 (350 mm from thruster).



(a) Influence of Channel 1 over Channel 2.



(b) Influence of Channel 2 over Channel 1.

Fig. 5.16. Transfer Entropy between probes at Position 3 (300 mm from thruster).

Confirming the findings of the cross-correlation shown before, the transfer entropy also suggests that there is a statistical causality between the studied probes.

It is clear from comparing the graphs, that Channel 1 seems to have a bigger influence on Channel 2 than Channel 2 over Channel 1. However, the influences in both directions are well above the statistical threshold set for the length of these signals.

This study has been performed by varying the number of data points studied in each iteration. There is a minimum number of points that the Transfer Entropy technique in order to provide meaningful results, as it can be appreciated from the initial sudden changes on the figures when N was still too low. If the influence of one of the channels over the other was fixed in time, it should be expected that as N grows, the Transfer Entropy reaches a saturation value in which no matter how many points are included in the sample, the result does not vary.

In this case, however, as N is increased the Transfer Entropy seems to oscillate around a value in each case, but it never reaches that saturation state. This can be due to the length of the signals not being long enough, but this reasoning seems unlikely since for the case of Position 3 the final length of the concatenated is $N = 2.5x10^6$. These variations could also be due to noise or some external factor during the data recollection that altered the results. Additionally, there exists the possibility that the variations observed from the Transfer Entropy results could be due to the azimuthal flow of information switching from one side of the thruster to the other, being the time when the information flow comes from Channel 1 longer, based on the bigger magnitude of the TE coming from this Channel.

To conclude this section, it should be highlighted how the Transfer Entropy allows to study the directionality of the relationship between the signals, as opposed to the cross-correlation shown in the previous subsection.

5.2. Hall-Effect Thruster Simulation

5.2.1. Embedding dimensions

Again, the first step before performing any analysis on the data sets was to obtain the embedding dimensions corresponding to the different variables to be studied with the Time Delayed Mutual Information and the False Nearest Neighbours techniques. From here on the variables studied will be referred to as: plasma density (ρ), ion current density (J_i), electric potential (ϕ) and electron temperature (T_e). The results are illustrated in the following table:

TABLE 5.2. Embedding dimensions for Hall-Effect Thruster Simulation

	ρ	$\mathbf{J_i}$	ϕ	$\mathbf{T_e}$
m	4	4	4	3
τ [μs]	4.8	2.7	3.3	24.6

No filtering of the data was performed in this case, since the time-series analyzed were generated by a numerical simulation.

5.2.2. Phase Portraits

The Phase Portraits presented here have been created for different positions along the central channel in order to asses the differences between the spatial references. The reference points used are shown in the following figure, and will be used in different sections along this results chapter for the HET.

These points were selected in order to study the different regions of interest inside the central channel. The first point corresponds to the dynamics inside the inner channel of the thruster, the second and third points have been chosen to lay right before and after the cathode bound, and finally the fourth point corresponds to a position near the end of the HET once the cathode bound has already been surpassed.

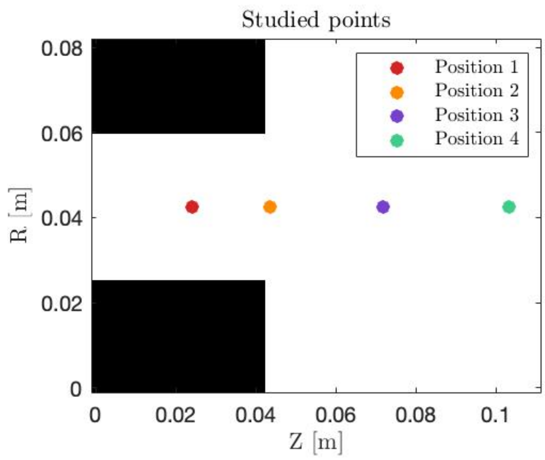
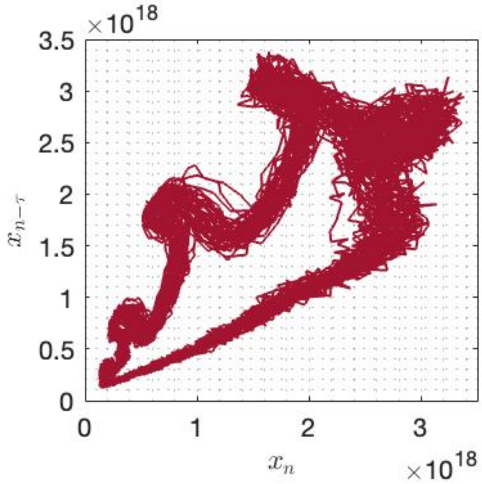
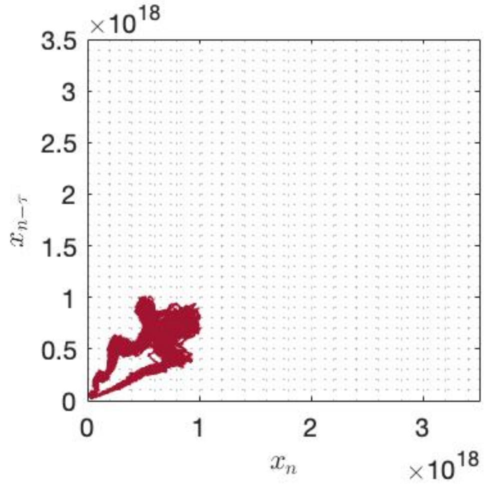


Fig. 5.17. Reference points studied along central channel on 2D grid.

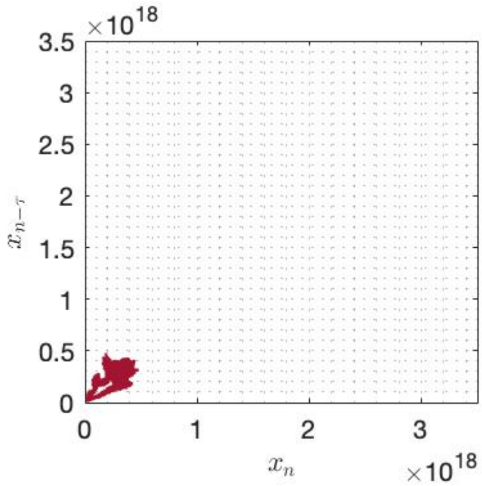
For the Phase Portraits shown in this subsection, the same grid has been used to compare the four positions in order to show the expansion of the studied systems. Additionally, a close-up of the first and fourth position is also provided for each case to better study their structures.



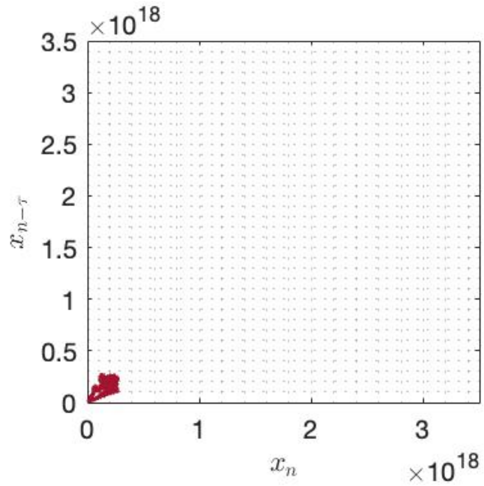
(a) Position 1



(b) Position 2



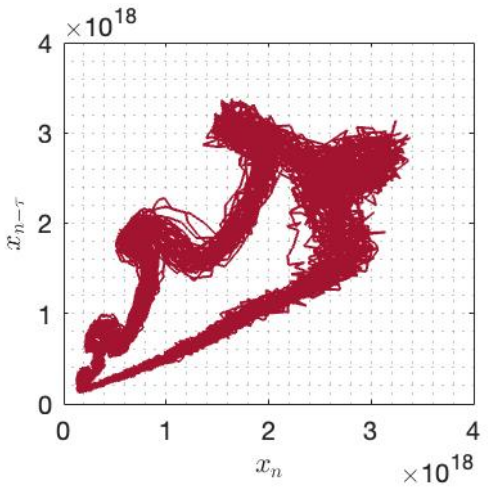
(c) Position 3



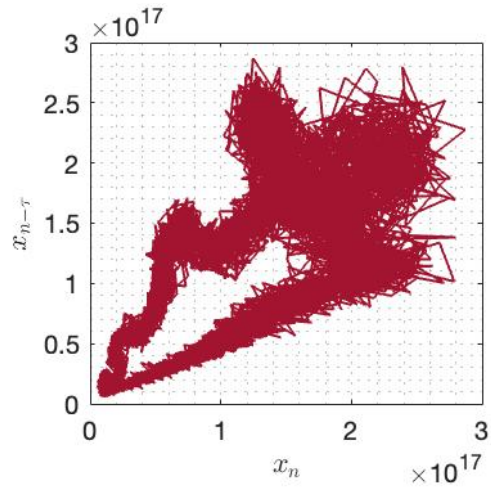
(d) Position 4

Fig. 5.18. Poincaré graph for plasma density along different positions. Time lag used for study: $\tau = 4.8x10^{-6}s$.

In the case of Figure 5.18, which represent the plasma density for the different positions, it can be appreciated that initially the 2D-attractor found has a huge scale that starts reducing as the nozzle end gets closer. Physically, this means that between the current state and the past state there are smaller differences in magnitude as the plasma advances through the thruster.



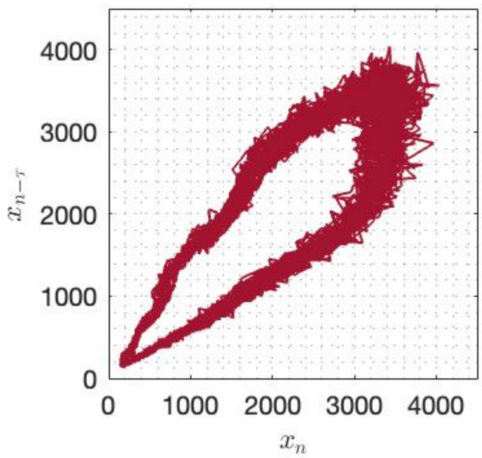
(a) Position 1



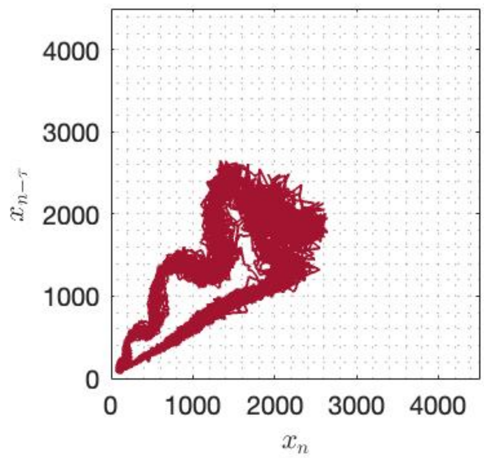
(b) Position 4

Fig. 5.19. Poincaré graph for plasma density along different positions. Close-up.

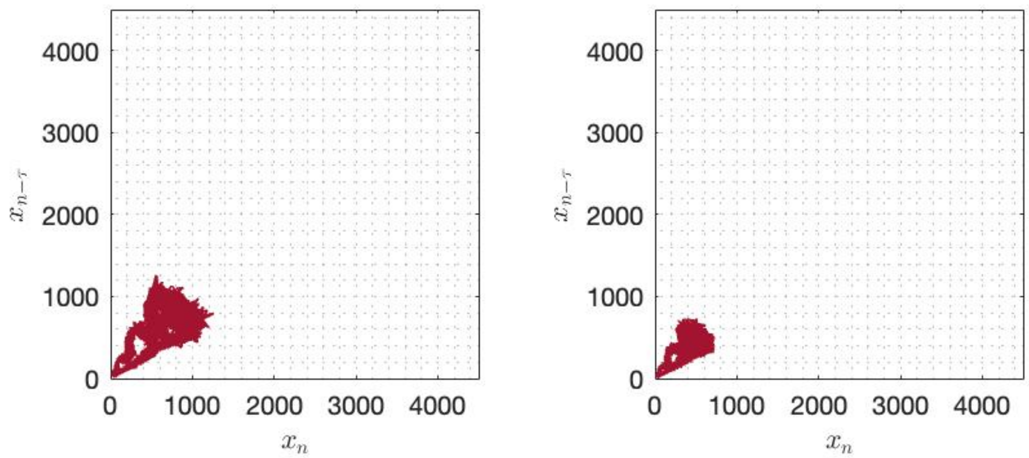
Even though the scales of the 2D-attractors differ greatly for the plasma density, by inspecting the close-ups presented in Figure 5.19, the same overall structure can be recognized. For smaller values, represented in the graph towards the lower left corner, the dynamics of the plasma density seem to be contained, but as bigger values are approached, represented in the graph towards the upper right corner, the system becomes more unbounded. It should also be highlighted that the chaoticity of the attractor is bigger for Position 4 than for Position 1.



(a) Position 1



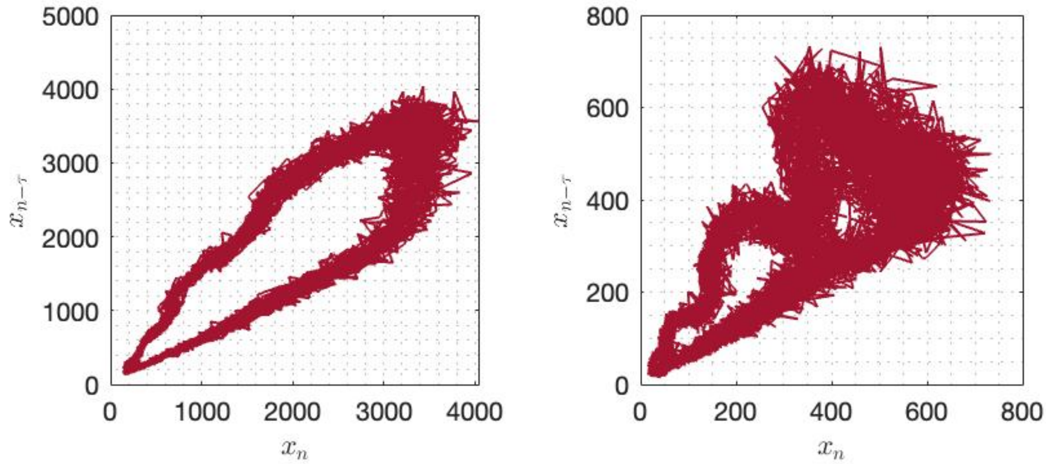
(b) Position 2



(c) Position 3 (d) Position 4

Fig. 5.20. Poincaré graph for ion current density along different positions. Time lag used for study: $\tau = 2.46x10^{-5}s$.

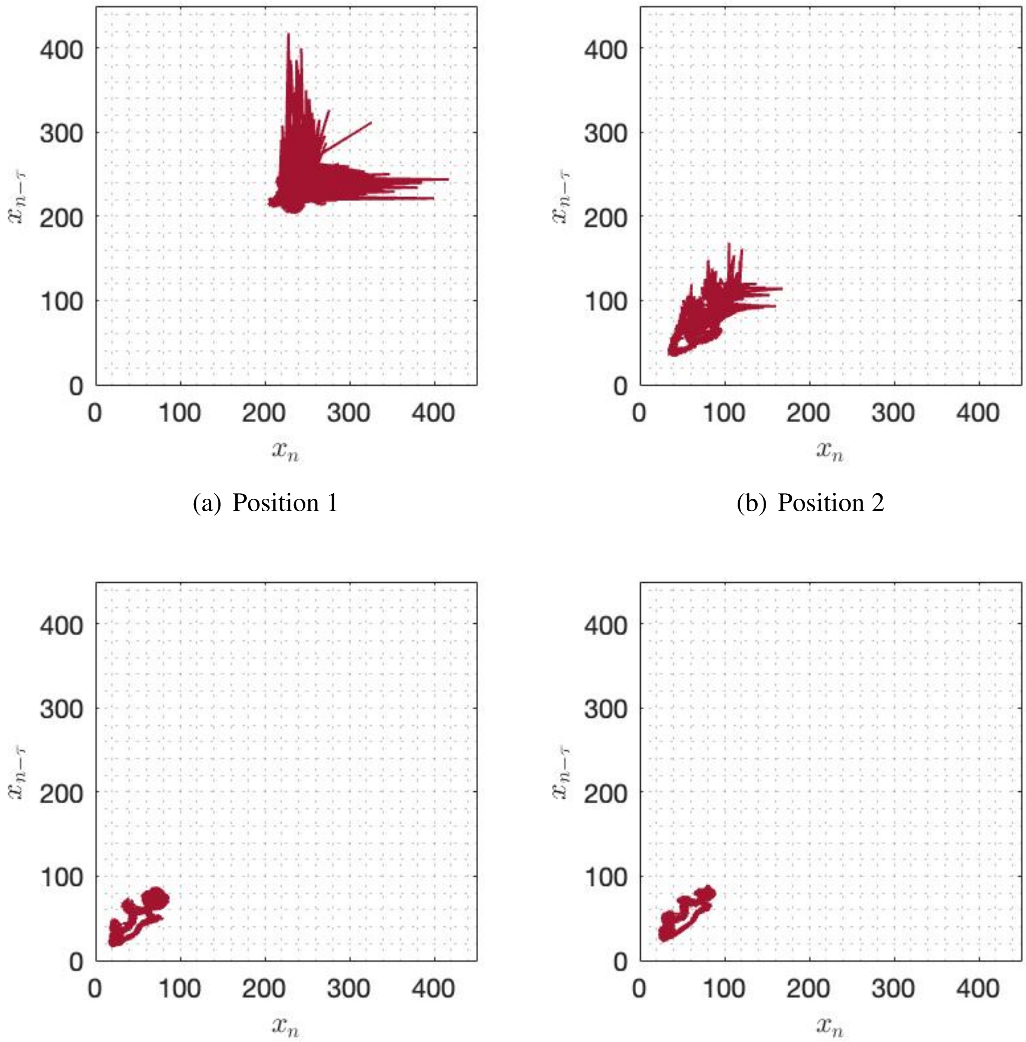
For the ion current density, as it happened for the plasma density case, the scale of the attractor also decreases as they system goes from position 1 to position 4. Additionally, the same overall structure is obtained for all cases, but more irregularities are present in Position 4 than in Position 1, as appreciated on Figure 5.21.



(a) Position 1 (b) Position 4

Fig. 5.21. Poincaré graph for ion current density along different positions. Close-up.

An exception to the pattern found for the variables studied before was found for the case of the electric potential. For the rest of the cases, the 2D-attractors recovered presented the same overall structure at different scales and with the inclusions of some irregularities.

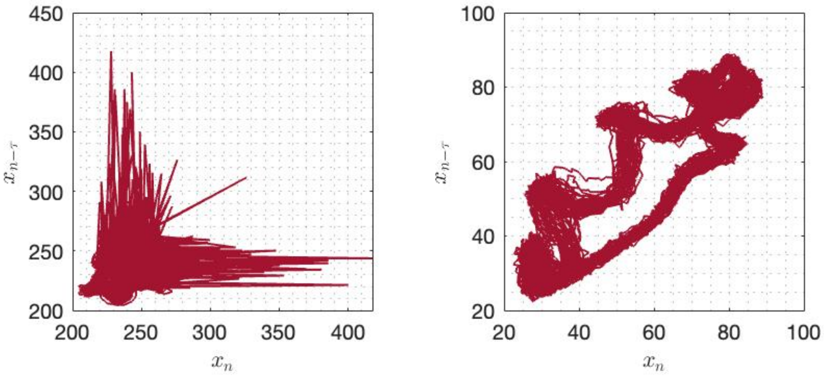


(a) Position 1 (b) Position 2

(c) Position 3 (d) Position 4

Fig. 5.22. Poincaré graph for electric potential along different positions. Time lag used for study: $\tau = 3.3x10^{-6}s$.

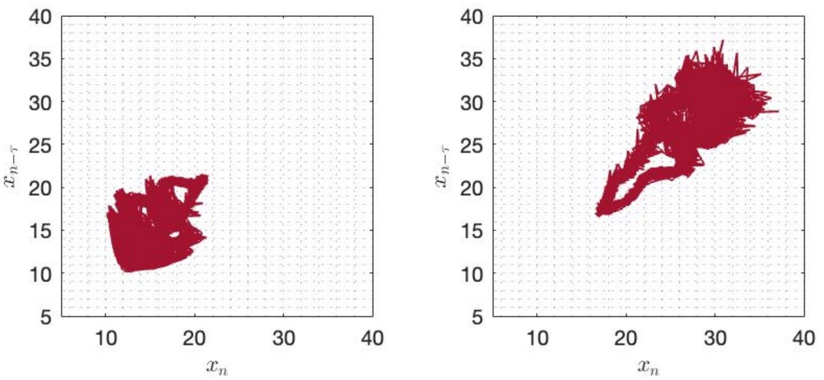
For the case of the electric potential, as shown in Figure 5.22, the structure recovered for positon 1 presents a lot of irregularities and is displaced on the grid with respect to the other 3 figures. Position 2 still presents some of the irregularities seen on Pos. 1, but on a smaller scale, and the figure is also shifted towards the starting point of the rest of the positions.



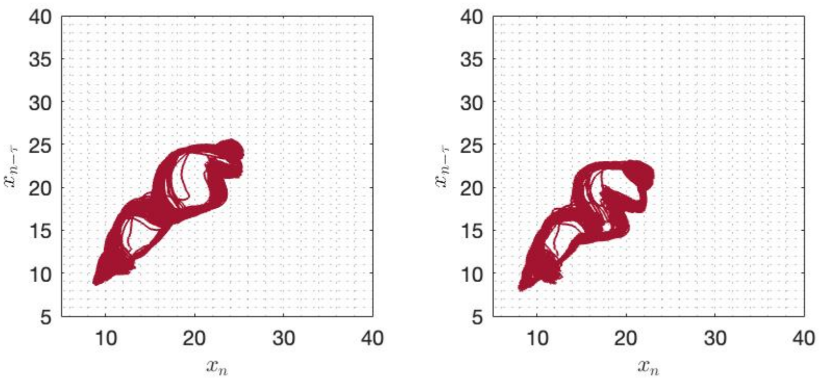
(a) Position 1 (b) Position 4

Fig. 5.23. Poincaré graph for electric potential along different positions. Close-up.

In this case, as previously stated, the structures recovered are nothing alike due to the inclusion of disturbances that make the 2D-attractor for position 1 blow up. This agrees with the result obtained in the following Recurrence Plots section, in which for Position 1 some chaotic-like behaviour was obtained, but it was lost as Position 4 was reached.



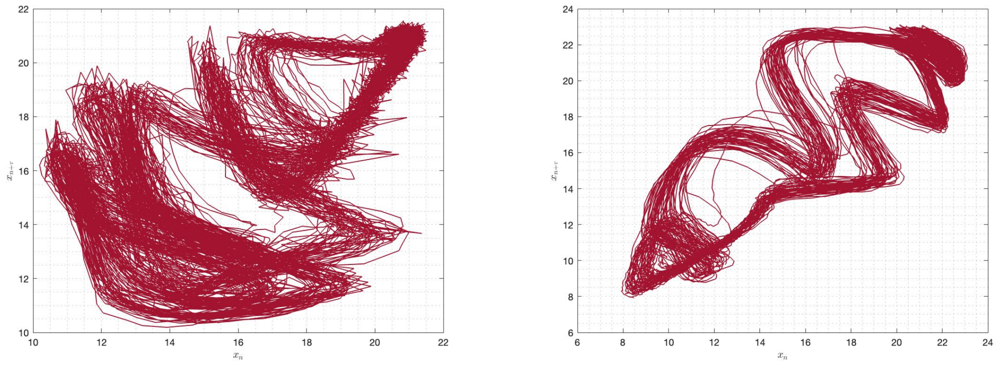
(a) Position 1 (b) Position 2



(c) Position 3 (d) Position 4

Fig. 5.24. Poincaré graph for electron temperature along different positions. Time lag used for study: $\tau = 2.7x10^{-6}s$.

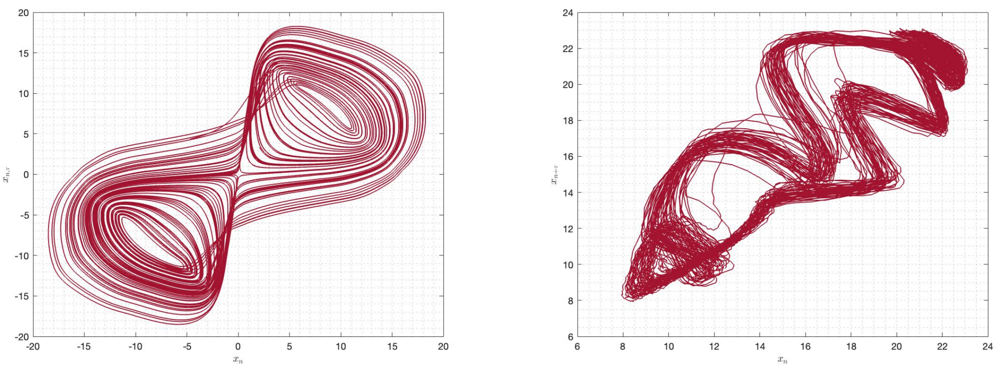
Regarding the electron temperature case, the chaotic-like behaviour of the attractor found is not as prominent as that found on Subfigure 5.23(a), but the pattern found for both plasma and ion current density was not recovered either. As for the electric potential case, some "noisy" attractor was recovered for position 1, and then at position 2 the attractor recovered resembled the one corresponding to Pos. 3 and 4 more, but with the difference that in this case Pos. 2 seems to be displaced towards the upper right corner of the grid, indicating bigger values.



(a) Position 1 (b) Position 4

Fig. 5.25. Poincaré graph for electron temperature along different positions. Close-up.

By examining the close-up of both positions 1 and 4, it can be appreciated in Subfigure 5.26(a) that a closed-loop structure seems to form towards the upper right corner, and another closed-loop form appears towards the lower left corner. However, many irregularities and outburts are also included which deform the recovered attractor making it differ a lot from the one corresponding to position 4.



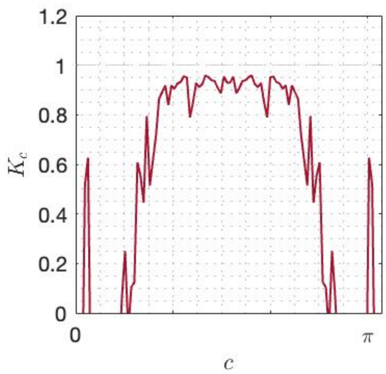
(a) Lorenz System (b) Electron Temperature

Fig. 5.26. Comparison of Poincaré graphs for Lorenz system and electron temperature.

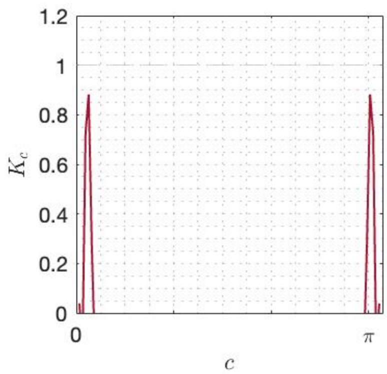
As in the case of Lorenz, it was found that the optimum embedding dimension for the electron temperature data was $m = 3$. From Figure 5.26, a clear paralellism between the two projections of the attractor can be seen. For Lorenz, two periodic states are interchanged in a chaotic manner. For the electron temperature, a similar behaviour could be expected after analyzing this graph.

5.2.3. The 0-1 test for Chaos

After finding two indicators for possible Chaos for the electric potential and the electron temperature in Position 1, the 0-1 test for Chaos was performed in order to continue studying this possibility.



(a) Position 1.



(b) Position 4.

Fig. 5.27. The 0-1 test for Chaos performed on the electric potential for different positions.

The test for chaos came back with $K_c = 1$ for most of the values of c, which is an indicator of chaos present in the system. A comparison also representing the electric potential but at position 4, where no indication of Chaos was present, has been included in the figure to further illustrate this. In comparison, the test for position 4 came back with $K_c = 0$ for almost all values of c, indicating no Chaos present.

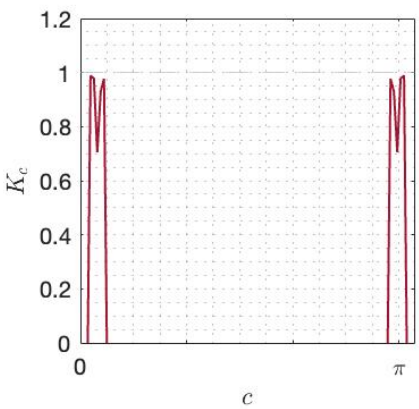
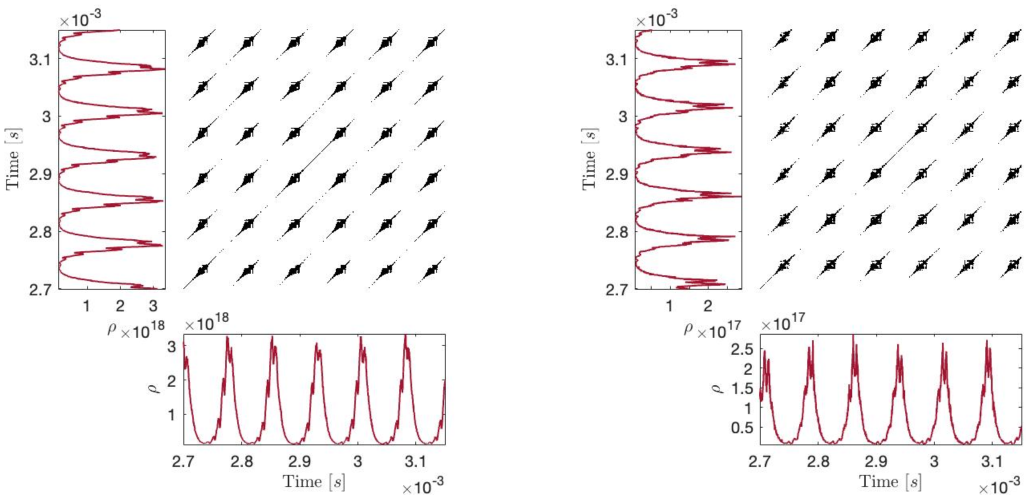


Fig. 5.28. The 0-1 test for Chaos performed on the electron temperature for Position 1.

On the other hand, the test performed for the electron temperature at position 1 came back with $K_c = 0$ for almost all values of c, contradicting the hypothesis of chaos being present made from the results of the Poincaré graph for that specific position.

5.2.4. Recurrence Plots

The next step of the analysis was to create the Recurrence Plots corresponding to the different positions studied. After obtaining all the results, it was observed that the major differences occured between positions 1 and 4, so in order to avoid redundancies, only there 2 positions are shown in this section. It should be also kept in mind that the embedding dimension was chosen to be $m = 4$ for all variables and positions in order to better compare the results.



(a) Position 1. (b) Position 4.

Fig. 5.29. Recurrence plots for plasma density at different positions of central channel.

By studying Figure 5.29, no significant difference is appreciated between Position 1, corresponding to a point inside the thruster, to Position 4, almost at the end of the nozzle. A periodic pattern with the addition of some possible noise can be appreciated, with a corresponding period of $7.5x10^{-5}s$, which corresponds to the period of the discharge current, as illustrated on Figure 5.30.

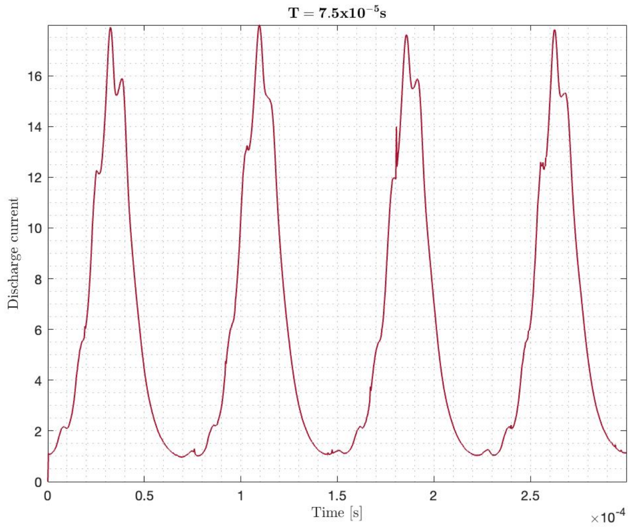
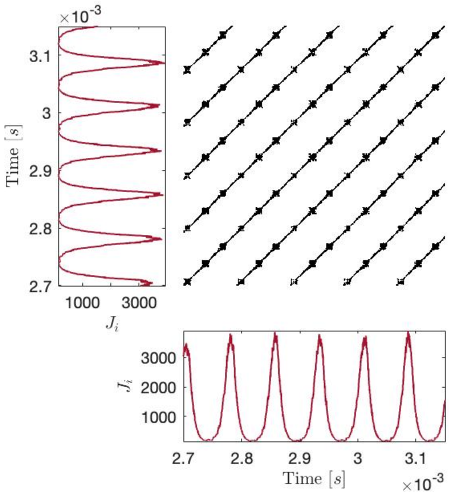
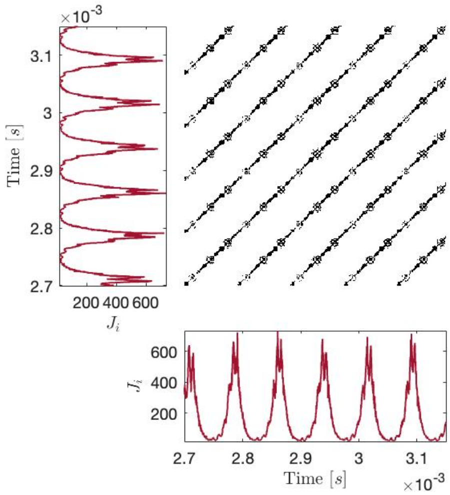


Fig. 5.30. Discharge current of Hall-Effect Thruster simulation.



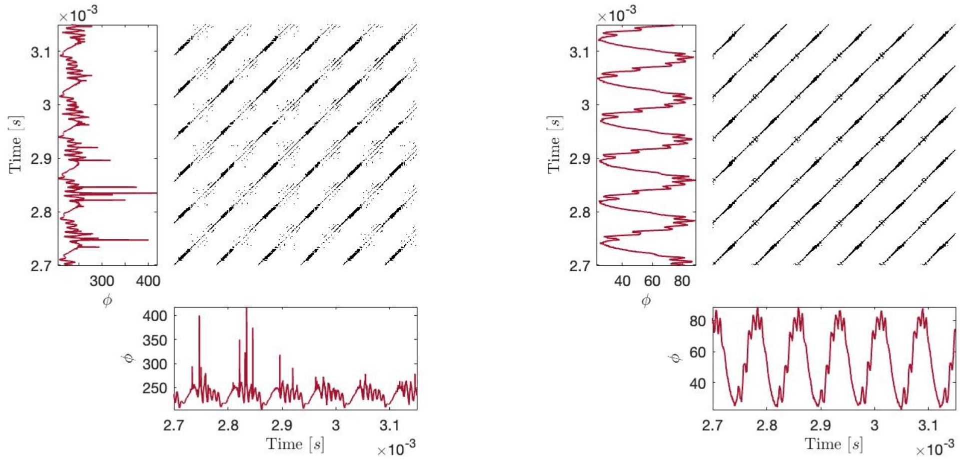
(a) Position 1.



(b) Position 4.

Fig. 5.31. Recurrence plots for ion current density at different positions of central channel.

For the case of the ion current density, the same periodic pattern seems to be present. However, a difference can be appreciated for this case between position 1 and position 4. For position 4, the inclusion of some disturbances can be appreciated, both in the time-series and in the recurrence plot. In the recurrence plot this irregularities are illustrated in the form of clusters along the main diagonals.

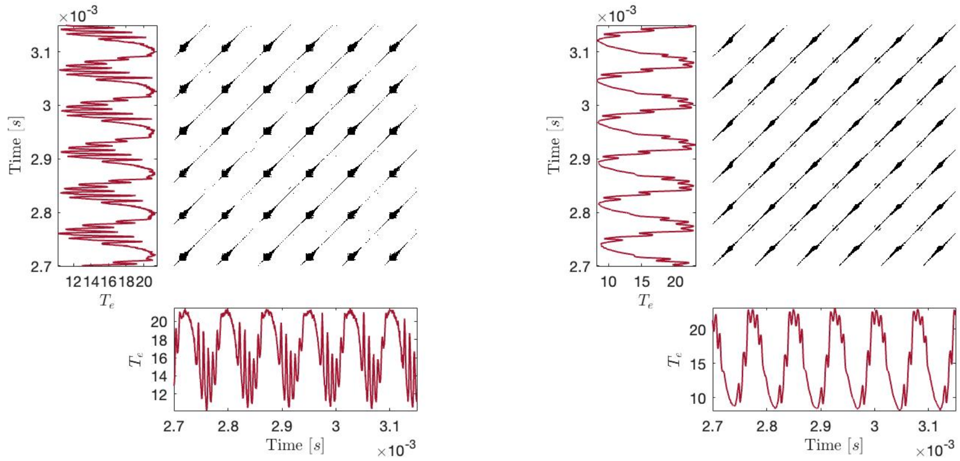


(a) Position 1.

(b) Position 4.

Fig. 5.32. Recurrence plots for electric potential at different positions of central channel.

Regarding the electric potential, again the same overall periodic pattern is found for this case. A clear difference between the pattern shown in Subfigure 5.32(a) and the rest of the Recurrence Plots present here can be identified. Even though the periodicity is present in this sample, some chaotic-like behaviour seems to appear, but as the position is increased until 4 is reached, this pattern fades away leaving only the periodic pattern with some small irregularities. This further supports the results obtained from the Poincaré graph and the 0-1 test for Chaos performed in previous sections for this specific variable and position. Additionally, it could be argued that some intermittency is present for the electric potential, regarding the peaks that can be observed around the time range $t = (2.7 - 2.9)x10^{-3}s$.



(a) Position 1.

(b) Position 4.

Fig. 5.33. Recurrence plots for electron temperature at different positions of central channel.

Finally, for the electron temperature, a similarity with the behaviour of the electric potential can be drawn. Even though the electron temperature presents a cleaner form, some irregular inclusions can be seen. It can also be appreciated that the irregularities present on Position 1 fade away as Position 4 is approached.

5.2.5. Transfer Entropy

In order to analyze the spatial correlations inside the HET, the Transfer Entropy technique was employed. The statistical significance was chosen to be $TE = 10^{-3}$ according to Figure 3.21 since the length of the studied signals is $N = 12000$. All results below this threshold have been suppressed from the figures since they do not offer ay relevant information to this analysis.

Transfer Entropy was applied between a reference point in the inner channel of the thruster, located at position $Z = 0.0105m$ and $R = 0.0425m$, and the rest of the points in the 2D mesh.

One of the most relevant results found which will be presented here, is the comparison between the electron temperature and itself. The spatial propagation of the electron temperature fluctuations will be completed by the temporal propagation analysis of the electron temperature, presented in the following subsection, using the same reference value.

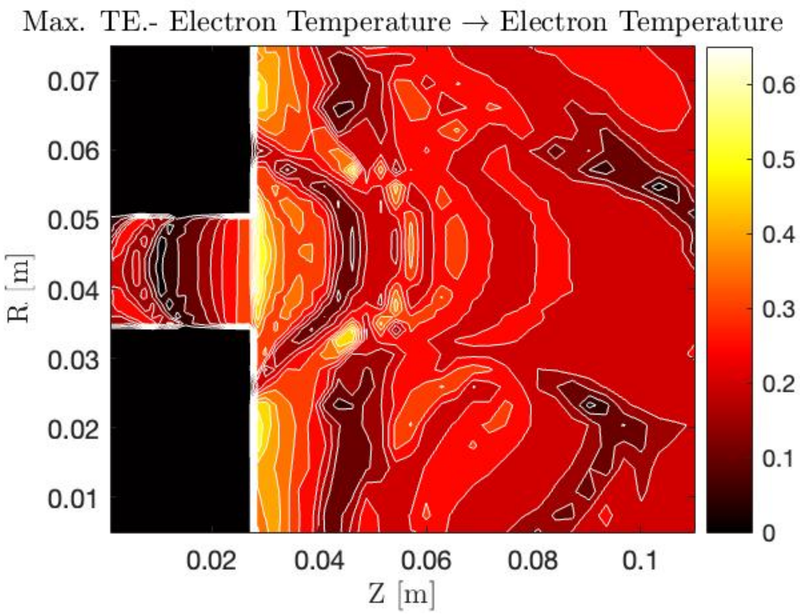


Fig. 5.34. Transfer Entropy studied over the 2D mesh of the Hall-Effect Thruster. Lines in white represent the different levels of TE present in the simulation. Reference T_e used for study: $Z = 0.0105m$, $R = 0.0425m$.

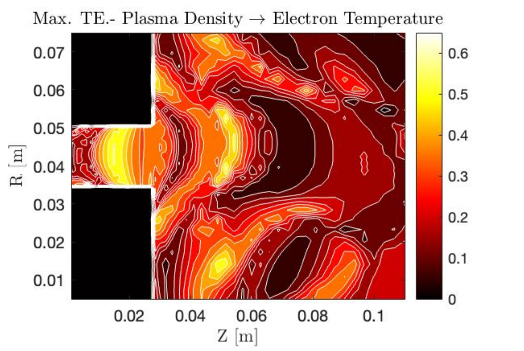
On Figure 5.34 what is presented is the maximum entropy found for a range of time lags studied, τ. What this represents is the maximum flow of information between the studied reference point and the rest of the grid.

There seems to be a high information flow coming from the reference value to the end of the discharge region of the HET, around $Z = 0.027m$. What this means is that there seems to be a high statistical causality between the electron temperature of the reference point studied and the electron temperature of the particles in this area.

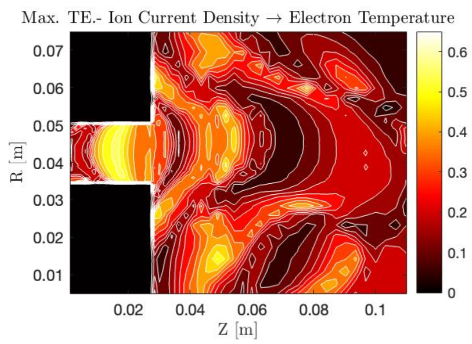
On the other hand, there are regions in the thruster where there seems to be no flow of information coming from the reference point. The round semicircle that appears right outside the thruster discharge area coincides with the cathode bound of the HET. This lack of information flow indicates that the state of the reference point studied does not affect the electron temperature of the particles contained by the cathode bound.

However, this lack of information flow in the cathode bound seems to be jumped over, since statistical causality seems to be found again after that point. It is interesting to note that right after the cathode bound, where a minimum of the transfer entropy was observed, another round semicircle seems to appear formed by small circular areas indicating high statistical causality.

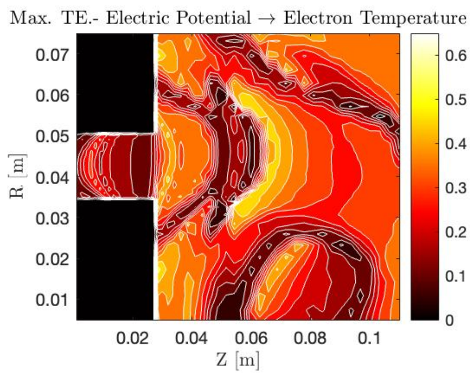
There are some additional areas where no information flow appears spread around the nozzle of the HET. Even though the HET presents a symmetrical geometry around its axis, these areas are not evenly spread with respect to this symmetry axis. As in the case of the cathode bound, the areas of no information flow appear to form aligned with the magnetic field of the HET, as appreciated on Figure 4.1(b). Similar statistical causality behaviours were found around the cathode bound area for plasma density, ion current density and electric potential, as shown in the following figure:



(a) Information flow between reference plasma density and electron temperature



(b) Information flow between reference ion current density and electron temperature



(c) Information flow between reference electric potential and electron temperature

Fig. 5.35. Maximum Transfer Entropy interchanged between the reference studied variable and the electron temperature of its surroundings. Lines in white represent the different levels of TE present in the simulation.

Temporal correlations

To study the temporal correlations occurring in the plasma, the Transfer Entropy technique was again employed. The statistical significance for the HET simulation was chosen to be $TE = 10^{-3}$ again as in the spatial correlations case. All results below this threshold have been surpassed from the figure shown in this section, since they do not offer any relevant information to this analysis.

Transfer Entropy was applied between a reference point in the inner channel of the thruster, located at position $Z = 0.0105m$, and the rest of the points constituting the central channel of the HET for a range of time lag values, τ. It should be noted that the figure presented here has been selected for a certain range of time lags, if this range were to be extended, the same patterns would be found again in a repetitive manner.

One of the most relevant results obtained was the one studying the propagation of electron temperature against itself. By plotting the Transfer Entropy obtained against the time lags for this case, the propagation of the temperature variations could be identified. The transfer entropy between the reference point and itself is zero by definition, on the Figure represented by a black line along the line $Z = 0.0105m$. This is due to this point being chosen for the reference value, and has no further physical meaning.

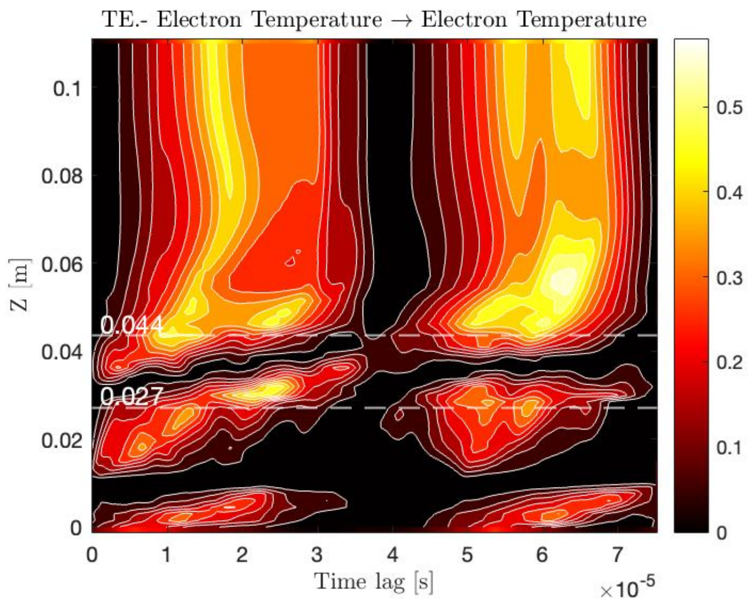


Fig. 5.36. Transfer Entropy studied over several time lags for the inner channel of the Hall-Effect Thruster. Lines in white represent the different levels of TE present in the simulation. Horizontal white dashed lines represent respectively the end of the thruster (0.027 m) and the cathode bound (0.044 m). Reference T_e used for study: $Z = 0.0105m$.

Temperature propagation can be observed on Figure 5.36. The propagation coming from the reference point seems to move across the thruster as time goes by until $t = 3.5x10^{-5}s$, where it stops. This can be seen on the figure as a wide line that start from the reference point and grows diagonally until the mentioned time. After a lag of around $0.5x10^{-5}s$, the perturbation seems to break through the barrier that was containing it and grows both axially towards the end of the thruster, as represented from the vertical column expanding at $Z = 0.044 - 0.11m$, and temporally, illustrated by the width of the column mentioned before going from $t = 4.5x10^{-5}s$ to $t = 7.5x10^{-5}s$. This jump seems to coincide with the cathode bound of the Hall-Effect Thruster, represented by a white dashed line on the figure at $Z = 0.044m$, which aligns with the findings of the spatial correlations section where a jump around the cathode bound was also found.

Apart from this main diagonal propagation, two other information flows seem to be present inside the inner channel. First, at $t = 0s$, as the main diagonal propagation starts to occur, another propagation further down the channel begins simultaneously. In this case, the presence of the cathode bound does not seem to interrupt the propagation of electron temperature, since no jump is appreciated around this point ($Z = 0.044m$) in the Figure.

On the other hand, around a time lag of $t = 3.5x10^{-5}s$, coinciding with the time when the main diagonal propagation experiences its jump, an additional propagation blocks starts appearing from the reference point studied. This propagation presents again a jump right before the cathode bound where the propagation seems to become trapped.

6. CONCLUSION AND FUTURE OUTLOOK

This Thesis focused on the creation of a catalogue of Non-Linear Analysis Techniques that could be used in order to better understand the mechanisms that lead to the appearance of turbulence in plasma thrusters. Different classifications were formulated according to the use of the techniques presented.

First, a quick review of the linear techniques traditionally used in signal analysis was performed, since these techniques serve as the basis of almost all data analysis performed nowadays.

After that, the Time Delay Embedding techniques were introduced. These techniques allow the user to reconstruct the phase space in which the dynamics of the time series take place. The two techniques presented in this category complement each other, from the Time Delayed Mutual Information the optimum time lag for embedding can be recovered, and once that parameter is obtained, the False Nearest Neighbours (FNN) technique can be applied in order to determine the minimum embedding dimension of the system.

In the third place, the Visual techniques were presented. Both the Phase Portraits and the Recurrence Plots allow to visualize patterns that may be hidden in the dynamics of the time series. Phase Portraits make use of the optimum lag in order to find an attractor, or the projection of an attractor, for the signal studied. On the other hand, Recurrence Plots provide information both spatially and temporally by also making use of the embedding dimension found with the FNN, from the embedded space, RPs allow the user to visualize recurrences that occur in dimensions which may be higher than 3 in a simple 2D figure.

Following that, the Model Reduction techniques were introduced. These techniques seek to find a reduced model of the underlying dynamics of the studied system. The Hankel Alternative View of Koopman (HAVOK) combines linear algebra with the embedding of the time-series in order to provide a simple model in which the non-linearities are captured by an external forcing term which can be controlled. On the other hand, the Sparse Identification of Nonlinear Dynamics (SINDy) creates a library of possible families of functions, which can be customized by the user, and seeks to discover the simplest combination to describe the dynamics of the system from that family of functions.

Then, the Chaos Quantification techniques were studied. The Maximal Lyapunov Exponent quantifies the average exponential divergence between the particles of a chaotic system, even though this technique can also be used to study non-chaotic systems. The 0-1 test for Chaos serves as a cheap and direct test that identifies if chaos is present in the studied data set by just providing a 1 or a 0 for the studied system.

Finally, the Statistical techniques, composed by the Intermittency and the Transfer Entropy, close this catalogue that has been built by reviewing state of the art techniques used nowadays. The Intermittency allows the user to identify events in the studied set that normally have really short characteristics time lengths. On the other hand, the Transfer Entropy is a technique that allows to identify between two signals not only statistical causality, but also the direction of said causality, positioning itself over traditional techniques such as the Correlation.

This dissertation serves as a guide to understand and implement the documented techniques, which have been tested against well-known systems, such as Lorenz or the Logistic Map, and data coming from both experiments and simulations.

This catalogue has a great potential to support the analysis of complex systems, not only in plasma propulsion but in other fields too. However, this catalogue is far from complete, additional techniques should be included in the existing categories and in new ones in order to cover all aspects regarding non-linearities inside the plasma physics field. Additionally, in view of how these techniques build on each other, an automatization algorithm could be implemented in which the user simply inputs the time series to be analyzed and selects the desired analyses to be performed, and the program autonomously provides all the parameters and outputs requested.

Regarding the results obtained, some interesting findings have been made. For the HPT case, a statistical causality between the probes employed was identified, indicating an azimuthal transfer of information flow inside the thruster. For the case of the HET simulation, chaos was detected for the electric potential inside the discharge area, but it was mitigated as the plasma advanced through the thruster towards the nozzle, suggesting some dissipation mechanisms may be present. Additionally, a lack of information flow for the electron temperature fluctuations was discovered around the cathode bound area, and some other regions that seem to be aligned with the magnetic field present in the Hall-Effect Thruster.

However, these findings should not be taken as absolute statements, especially those regarding the HPT experiment because of the quality of the data analyzed. Plasma turbulence is a complex field and it would not be sensible to draw definite conclusions from the results obtained in this project. Further analysis and experiments should be performed in order to see if the outputs obtained here are consistent for other simulations or experiments.

On the whole, the proposed objectives for this project have been fulfilled. A large catalogue has been composed with techniques that have been implemented and tested, and a preliminary analysis of both experimental and simulation data has been performed. Certainly there is still a lot of work to do in order to fully understand the mechanisms that govern turbulence in plasma physics, but the creation of these catalogue serves as a first step to bring some light into this complex matter.

7. LEGAL FRAMEWORK

This Bachelor Thesis has been carried out in accordance to all applicable legislation. This project is purely theoretical since the analyzed data has been freely provided by the EP2 group and no participation in the experiments set-up or data collection has been taken part by the author, so no assessment of hazards or risks needs to be made.

There are two main legislations which directly affect this Thesis:

1. Law 1/1996, 12 of April. Article 32 [63].

   All techniques gathered in this project have been collected in accordance to the Intellectual Property Rights Act (Ley de propiedad intelectual. Ley de 1/1996 de 12 de abril, artículo 32), and have been correctly credited and cited.

2. Law 38/2003, 17 of November, General on Subsidies [64].

   Additionally, since this project has been partly sponsored by a Collaboration Scholarship with University Departments from University Carlos III of Madrid by the Spanish Ministry of Education, Law 38/2003 (Ley 38/2003, de 17 de noviembre, General de Subvenciones) and the regulatory framework of these scholarships also affect this Thesis. In conformity with them, it is formally confirmed that the required hours by the scholarship, in addition to the ones corresponding to the 12 ECTS of the Bachelor Thesis, have been fully completed and all responsibilities corresponding to the author of this project have been fulfilled.

Concurrently, all computations and simulations have been carried out in agreement with MATLAB programming standards and practices.

8. SOCIO-ECONOMIC IMPACT

8.1. Socio-Economic significance

This project assesses methods employed to simplify and improve the analysis of the chaotic behaviour of plasma propulsion devices, serving as a first step in the direction of localizing and identifying possible causes and consequences of the non-linear aspects involved in plasma physics. The development of a library of non-linear analysis techniques provides the EP2 (Equipo de Propulsión Espacial y Plasmas) Research Group a very powerful tool that could be employed to optimize the plasma thrusters used in the laboratory, currently the Helicon Plasma Thruster (HPT) and the Hall Effect Thruster (HET) in the prospective future.

As explained in previous sections, when plasma goes through a magnetic field, the plasma becomes unstable and its fluctuations evolve into turbulence. The "anomalous transport" phenomenon represents one of the main challenges that physicist face when trying to confine plasma, resulting in an experimental axial current higher than the one predicted by classical transport theory [65].

As a consequence, the study of plasma turbulence and its evolution has been the target of research for decades [66] [67] [68] [69]. Accurately identifying and controlling plasma parameters that directly impact the appearance and growth of turbulence would allow scientists to better design the magnetic fields used for plasma confinement and therefore improve and maximize the overall life expectancy of these propulsion devices [70].

By improving the magnetic confinement of plasma, NASA scientists were able to design a Hall effect Rocket with a service life of 50 kh at specific impulses up to 3000 s, the Hall effect Rocket with Magnetic Shielding (HERMeS) [71]. This represents a huge improvement in the life expectancy when compared to typical thrusters in this field, which have an average lifespan on the order of 10 kh for that range of specific impulses [69]. It has been estimated that the average cost of a communication satellite mission using a Hall thruster is about 0.8 M € per year [72], which means that the 40 kh lifespan improvement for these type of systems corresponds to a monetary value of more than 3.5 M €.

8.2. Project budget

The estimated budget for this project has been separated in two main subsects, namely: software and equipment and human resources. The total cost according to these calculations adds up to a value of **11803.7 €** (Table 8.3). The Value Added Tax (VAT) of the 21% has been applied to the pertinent categories and an overhead of the 15% has been added to the final cost according to the Governing Council Agreement of UC3M passed on 2019 regarding research projects [73].

Software and equipment

TABLE 8.1. Software and equipment costs.

Software of equipment	Price per unit (€)	Quantity	Cost (€)
Matlab Academic License	69.0*	1	69.0
Personal computer. MacBook Air 8-Core CPU and 8-Core GPU.	1151.0	1	1151.0
Subtotal			1220.0
VAT (21 %)			256.2
TOTAL			**1476.2 €**

* Student Academic License

Human Resources

TABLE 8.2. Human resources costs.

Position	Salary (€/hour)	Total hours	Cost (€)
Project coordinator	35.0	22.5	787.5
Undergraduate researcher	10.0	800.0*	8000.0
TOTAL			**8787.5 €**

* 300 hours corresponding to the Bachelor Thesis (12 ECTS) and 500 hours corresponding to a Collaboration Scholarship in University Departments issued by the Spanish Ministry of Education at University Carlos III de Madrid.

Total project budget

TABLE 8.3. Total project costs.

Cost category	Cost (€)
Software and equipment	1476.2
Human resources	8787.5
Subtotal	10263.7
Overhead (15%/Total)	1539.6
TOTAL	**11803.7 €**

BIBLIOGRAPHY

[1] ITER, "Iter scientific data centre | how to manage 2 petabytes of new data every day," *ITER*, Nov. 23, 2020. [Online]. Available: https://www.iter.org/whatsnew/print/285n (visited on 06/10/2021).

[2] T. N. Parashar, W. H. Matthaeus, and M. A. Shay, "Dependence of kinetic plasma turbulence on plasma β," *The Astrophysical Journal Letters*, vol. 864, no. 1, p. L21, 2018.

[3] T. Yamada *et al.*, "Anatomy of plasma turbulence," *Nature physics*, vol. 4, no. 9, pp. 721–725, 2008.

[4] M. N. Wass, L. Ray, and M. Michaelis, "Understanding of researcher behavior is required to improve data reliability," *GigaScience*, vol. 8, no. 5, giz017, 2019.

[5] R. Nowak, "Problems in clinical trials go far beyond misconduct," *Science*, vol. 264, no. 5165, pp. 1538–1542, 1994.

[6] D. Escobar and E. Ahedo, "Low frequency azimuthal stability of the ionization region of the hall thruster discharge. i. local analysis," *Physics of Plasmas*, vol. 21, no. 4, p. 043 505, 2014.

[7] J. P. L. NASA, *JPL's 6 kW Hall thruster*. Jet Propulsion Laboratory. [Online]. Available: https://sec353ext.jpl.nasa.gov/ep/img/hall_thrusters/6kw_234.jpg.

[8] D. M. Goebel and I. Katz, *Fundamentals of electric propulsion: ion and Hall thrusters*. John Wiley & Sons, 2008, vol. 1.

[9] D. Staack, *Schematic of a Hall effect thruster electric propulsion device. This is a cross-section of a radially-symmetric device.* Wikimedia Commons. [Online]. Available: https://commons.wikimedia.org/wiki/File:Wfm_hall_thruster.png#/media/File:Wfm_hall_thruster.svg.

[10] *The technology*, Apr. 2020. [Online]. Available: https://hipatia.uc3m.es/the-technology/.

[11] "Sener aeroespacial and the uc3m work on a helicon plasma thruster for small space platforms," *SENER Group*, Jan. 2020. [Online]. Available: https://www.group.sener/press-releases/sener-aeroespacial-and-the-uc3m-work-on-a-helicon-plasma-thruster-for-small-space-platforms-.

[12] J. Navarro-Cavallé, M. Wijnen, P. Fajardo, and E. Ahedo, "Experimental characterization of a 1 kw helicon plasma thruster," *Vacuum*, vol. 149, pp. 69–73, 2018.

[13] V. Tsytovich, *Nonlinear effects in plasma*. Springer Science & Business Media, 2012.

[14] S. Brunton. (). "
fourier analysis: Overview." mar. 7, 2020," [Online]. Available: https://youtu.be/jNC0jxb0OxE.

[15] S. H. L., *Understanding correlation*, Jan. 2017. [Online]. [Online]. Available: https://www.allaboutcircuits.com/technical-articles/understanding-correlation/.

[16] K. Taira *et al.*, "Modal analysis of fluid flows: An overview," *Aiaa Journal*, vol. 55, no. 12, pp. 4013–4041, 2017.

[17] B. J. Schwarz and M. H. Richardson, "Experimental modal analysis," *CSI Reliability week*, vol. 35, no. 1, pp. 1–12, 1999.

[18] K. Academy. (). ""introduction to eigenvalues and eigenvectors | linear algebra | khan academy" nov. 14, 2009," [Online]. Available: https://youtu.be/PhfbEr2btGQ.

[19] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Transactions on automatic control*, vol. 25, no. 2, pp. 164–176, 1980.

[20] W. Müller, A. Jung, and H. Ahammer, "Advantages and problems of nonlinear methods applied to analyze physiological time signals: Human balance control as an example," *Scientific reports*, vol. 7, no. 1, pp. 1–11, 2017.

[21] B. Carreras, L. Garcıa, J. Nicolau, B. P. van Milligen, U. Hoefel, M. Hirsch, *et al.*, "Intermittence and turbulence in fusion devices," *Plasma Physics and Controlled Fusion*, vol. 62, no. 2, p. 025 011, 2019.

[22] B. Cannas, A. Fanni, A. Murari, and F. Pisano, "Recurrence plots for dynamic analysis of type-i elms at jet with a carbon wall," *IEEE Transactions on Plasma Science*, vol. 47, no. 4, pp. 1871–1877, 2019. DOI: 10.1109/TPS.2019.2901313.

[23] H. Kantz and T. Schreiber, *Nonlinear time series analysis*. Cambridge university press, 2004, vol. 7.

[24] A. Krakovská, K. Mezeiová, and H. Budáčová, "Use of false nearest neighbours for selecting variables and embedding parameters for state space reconstruction," *Journal of Complex Systems*, vol. 2015, 2015.

[25] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical review A*, vol. 33, no. 2, p. 1134, 1986.

[26] K. Academy. (2014). "Information entropy | journey into information theory | computer science | khan academy," Youtube, [Online]. Available: https://youtu.be/2s3aJfRr9gE.

[27] S. Vajapeyam, "Understanding shannon's entropy metric for information," *arXiv preprint arXiv:1405.2061*, 2014.

[28] T. Schreiber, "Measuring information transfer," *Physical review letters*, vol. 85, no. 2, p. 461, 2000.

[29] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. Cambridge University Press, 2003. DOI: 10.1017/CBO9780511755798.

[30] A. Fernández Dıaz, L. Escot Mangas, and P. Grau Carles, "What
s new and useful about chaos in economic science," 2012.

[31] M. B. Kennel, R. Brown, and H. D. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical review A*, vol. 45, no. 6, p. 3403, 1992.

[32] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1-2, pp. 43–50, 1997.

[33] BTL, *Poincaré graph. complete ecg record in one sight*, 2014. [Online]. [Online]. Available: https://files.btlnet.com/product-document/9792e3d5-3dbf-45d8-9e84-5c964a6a8602/BTL-Cardiopoint_WP_Poincare-graph_EN400_9792e3d5-3dbf-45d8-9e84-5c964a6a8602_original.pdf.

[34] C. Grebogi, E. Ott, and J. A. Yorke, "Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics," *Science*, vol. 238, no. 4827, pp. 632–638, 1987.

[35] G. A. Gottwald and I. Melbourne, "The 0-1 test for chaos: A review," *Chaos detection and predictability*, pp. 221–247, 2016.

[36] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, *et al.*, "Recurrence plots of dynamical systems," *World Scientific Series on Nonlinear Science Series A*, vol. 16, pp. 441–446, 1995.

[37] B. Goswami, "A brief introduction to nonlinear time series analysis and recurrence plots," *Vibration*, vol. 2, no. 4, pp. 332–368, 2019.

[38] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5, pp. 237–329, 2007. DOI: https://doi.org/10.1016/j.physrep.2006.11.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157306004066.

[39] P. Mörters and Y. Peres, *Brownian motion*. Cambridge University Press, 2010, vol. 30.

[40] C. Stan, C. P. Cristescu, and D. G. Dimitriu, "Analysis of the intermittent behavior in a low-temperature discharge plasma by recurrence plot quantification," *Physics of Plasmas*, vol. 17, no. 4, p. 042 115, 2010. DOI: 10.1063/1.3385796. eprint: https://doi.org/10.1063/1.3385796. [Online]. Available: https://doi.org/10.1063/1.3385796.

[41] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, “Chaos as an intermittently forced linear system,” *Nature communications*, vol. 8, no. 1, pp. 1–9, 2017.

[42] D. L. Donoho and M. Gavish, *The optimal hard threshold for singular values is 4/3*, 2013.

[43] H. D. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, “The analysis of observed chaotic data in physical systems,” *Reviews of modern physics*, vol. 65, no. 4, p. 1331, 1993.

[44] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the national academy of sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

[45] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos*. Springer, 1996.

[46] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, “Determining lyapunov exponents from a time series,” *Physica D: Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, 1985.

[47] N. Frolov, V. Maksimenko, K. Ilyenko, A. Koronovskii, and A. Hramov, “Using the spectrum of lyapunov exponents to analyze the dynamics of beam-plasma systems simulated by the large particle method,” *Bulletin of the Russian Academy of Sciences: Physics*, vol. 78, Jan. 2014. DOI: 10.3103/S1062873814020130.

[48] S. Marini and R. Pakter, “Single electron dynamics in a hall thruster electromagnetic field profile,” *Physics of Plasmas*, vol. 24, no. 5, p. 053 507, 2017.

[49] G. A. Gottwald and I. Melbourne, “A new test for chaos in deterministic systems,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2042, pp. 603–611, 2004.

[50] G. K. Batchelor, *The theory of homogeneous turbulence*. Cambridge university press, 1953.

[51] L. T. DeCarlo, “On the meaning and use of kurtosis.,” *Psychological methods*, vol. 2, no. 3, p. 292, 1997.

[52] U. Frisch and A. N. Kolmogorov, *Turbulence: the legacy of AN Kolmogorov*. Cambridge university press, 1995.

[53] C. Meneveau and K. Sreenivasan, “Simple multifractal cascade model for fully developed turbulence,” *Physical review letters*, vol. 59, no. 13, p. 1424, 1987.

[54] C. Meneveau and K. Sreenivasan, “The multifractal nature of turbulent energy dissipation,” *Journal of Fluid Mechanics*, vol. 224, pp. 429–484, 1991.

[55] B. van Milligen, I. Voldiner, U. Losada, B. A. Carreras, and C. Hidalgo, May 2021.

[56] B. P. Van Milligen, C. Hidalgo, and E. Sanchez, “Nonlinear phenomena and intermittency in plasma turbulence,” *Physical review letters*, vol. 74, no. 3, p. 395, 1995.

[57] G. Antar, P. Devynck, X. Garbet, and S. Luckhardt, “Turbulence intermittency and burst properties in tokamak scrape-off layer,” *Physics of Plasmas*, vol. 8, no. 5, pp. 1612–1624, 2001.

[58] M. Hirsch *et al.*, “Ece diagnostic for the initial operation of wendelstein 7-x,” in *EPJ Web of Conferences*, EDP Sciences, vol. 203, 2019, p. 03 007.

[59] B. van Milligen, B. Carreras, L. García, and J. Nicolau, “The radial propagation of heat in strongly driven non-equilibrium fusion plasmas,” *Entropy*, vol. 21, no. 2, 2019. DOI: 10.3390/e21020148. [Online]. Available: https://www.mdpi.com/1099-4300/21/2/148.

[60] *An advanced simulation code for hall effect thrusters*, EPIC Workshop presentation, Madrid, Spain: EP2 Research Group, 2017.

[61] D. Tskhakaya, K. Matyash, R. Schneider, and F. Taccogna, “The particle-in-cell method,” *Contributions to Plasma Physics*, vol. 47, no. 8-9, pp. 563–594, 2007.

[62] J. Gazeau, J. Nesetril, and B. Rovan, “The topology of deterministic chaos: Stretching, squeezing and linking,”

[63] M. de Cultura (España). (Apr. 1996), [Online]. Available: https://www.boe.es/buscar/pdf/1996/BOE-A-1996-8930-consolidado.pdf.

[64] J. del Estado (España). (Nov. 2003), [Online]. Available: https://www.boe.es/eli/es/l/2003/11/17/38/con.

[65] M. Reza, F. Faraji, T. Andreussi, and M. Andrenucci, “A model for turbulence-induced electron transport in hall thrusters,” in *Proc. 34th Int. Electric Propulsion Conf*, 2017.

[66] B. Parent, K. M. Hanquist, and A. Omprakas, “Fully-coupled simulation of plasma discharges, turbulence, and combustion in a scramjet combustor,” in *AIAA AVIATION 2020 FORUM*, 2020, p. 3230.

[67] Z. Brown, E. Dale, and B. Jorns, “Experimental correlation between anomalous electron collision frequency and plasma turbulence in a hall effect thruster,” in *36th International Electric Propulsion Conference, IEPC-2019-843 (University of Vienna, Austria, 2019)*, 2019.

[68] A. Smolyakov *et al.*, “Fluid and kinetic modelling of instabilities and transport in exb plasma discharges,” in *Proc. Int. Electr. Propuls. Conf.(IEPC)*, 2017, p. 411.

[69] N. P. Brown and M. L. Walker, “Review of plasma-induced hall thruster erosion,” *Applied Sciences*, vol. 10, no. 11, p. 3775, 2020.

[70] M. Cappelli, “The hall effect and rocket flight,” *Physics Today*, vol. 62, no. 4, pp. 76–77, 2009.

[71] R. Hofer *et al.*, “Completing the development of the 12.5 kw hall effect rocket with magnetic shielding (hermes),” 2019.

[72] R. Monill González, “Design and performance analysis study of a hall effect thruster,” B.S. thesis, Universitat Politècnica de Catalunya, 2016.

[73] C. de Gobierno. Universidad Carlos III de Madrid. (Jan. 2019), [Online]. Available: http://hdl.handle.net/10016/27961.