New Features of Permutation Entropy in Ordinal Patterns Complexity Plane

Rasika Dilhani

Supervisor: Alejandro C. Frery

Submitted in partial fullfilment of the requirements for
Doctor of Philosophy
in Data Science.

Victoria University of Wellington

2025

Abstract

This proposal provides a brief introduction to ordinal patterns and a literature review related to the Bandt & Pompe research article. After reviewing the literature on ordinal patterns and permutation entropy, we can conclude that ordinal patterns play an important role in conducting research. Bandt & Pompe proposed analyzing time series from the perspective of ordinal patterns to develop fast and automated methods for extracting qualitative information from nonlinear time series. They introduced the concept of permutation entropy to measure the complexity of a system underlying a time series while taking into account the ordering patterns that represent variations in the data. In this report I show how to use data to extract all ordinal patterns. First, I use the sample dataset to calculate their ordinal patterns with embedding dimension 3; I use letters to convert time series data into patterns. Secondly, I get the histogram on the data. And third, calculate the permutation entropy and statistical complexity. Next, determine the confidence interval of H without and with time dependence and finally the confidence interval of C.

Bandt and Pompe's ordinal pattern methodology has been widely used to investigate the latent dynamics of time series through their entropy, also known as permutation entropy. Nevertheless, there are no theoretical findings regarding the distribution of the permutation entropy, which needs to take the correlation effect between patterns into account. Considering that our approximation ignores the series dependence between symbols and that the asymptotic distribution of the permutation entropy is Normal We compare this result with the Multinomial sample entropy with embedding dimension 3, which assumes independence, and find that the expression

of the asymptotic variance becomes more complex as the embedding dimension increases. After that, a hypothesis test is developed and used to differentiate between the bearing fault diagnoses for four distinct rotating machines.

Contents

Chapter 1

Introduction

Time series contain valuable insights about the underlying system that generates the data. Their analysis is typically conducted using two primary approaches: time-domain and transformed-domain methods. In the context of time-domain analysis, Bandt and Pompe [5] introduced a novel methodology that is non-parametric and rooted in information theory descriptors: Ordinal Patterns symbolization.

Bandt and Pompe [5] proposed transforming small subsets of the time series observations into symbols that encode the sorting properties of the values in these subsets. Then, they computed a histogram of those symbols. The resulting distribution is less sensitive to outliers compared to the original data, and the histogram is independent of any specific model. The proposal proceeds by computing descriptors from this histogram, and extracting information about the system from these descriptors. As a result, this approach is versatile and applicable to a wide range of scenarios.

This study explores features derived from Bandt and Pompe symbolization, specifically Shannon entropy, complexity, and represents them graphically in the entropy-complexity plane, which is represented as a point in the $\mathbb{R}^2$ manifold. This proposal focuses on the independent analysis of permutation entropy and ordinal patterns in time series analysis. It involves calculating pattern histograms, entropy, complexity, and confidence intervals to better understand the statistical properties of these tools. Additionally, the role of confidence intervals in entropy and complexity, along with their applications in time-series clustering, will be explored. Future work will expand to include alternative measures, such as Rényi entropy, Tsallis entropy and Fisher information, with a focus on deriving confidence intervals for their entropy and complexity under the Multinomial model.

Time series analysis is widely applied across various fields, including engineering, economics, physical sciences, and more. A time series is defined as a collection of observations x_t, each representing a realized value of a particular random variable X_t, where time can be either discrete or continuous.

Examples of time series applications include finance (e.g., analyzing exchange rate movements or commodity prices), biology (e.g., modeling the growth and decline of bacterial populations), medicine (e.g., tracking the spread of diseases like COVID-19 or influenza), and geoscience (e.g., predicting wet or dry days based on past weather conditions).

The primary goal of time series analysis is to understand the nature of the phenomenon represented by the observed sequence. Time domain and frequency domain methods are the two primary approaches used in time series analysis. The temporal approach relies on concepts such as auto-correlation and regressions, where a time series' present value is analyzed in relation to its own past values or the past values of other series. This method represents time series directly as a function of time. On the other hand, the spectral approach represents time series through spectral expansions, such as wavelets or Fourier modes [49].

However, these methods often require assumptions such as large sample sizes or normally distributed observations that are rarely met in real-world empirical data. For many statistical techniques to be valid, these assumptions must hold, but in practice, they are frequently violated.

For example, traditional approaches to time series analysis, such as time domain and frequency domain methods, rely on assumptions that are not always valid in real-world data. The time domain approach, which uses techniques like auto-correlation and regression, assumes stationary and often struggles with nonlinear or non-stationary data. Similarly, the frequency domain approach, which represents time series through spectral expansions such as wavelets or Fourier modes, may require assumptions about periodicity and may not effectively capture short-term fluctuations.

Many statistical methods in these approaches depend on specific conditions, such as large sample sizes or normally distributed observations. However, these assumptions are often unrealistic, leading to inaccurate or biased results. When such conditions are not met, alternative methods must be considered.

As a result, alternative methods, often referred to as non-parametric techniques, must be considered. These methods rely on the rank R_t of the observations x_t rather than their actual values, making them robust and applicable to a wide range of data sets. Since non-parametric tests do not assume a normal distribution, they are highly reliable. For example, the Kruskal-Wallis H test and the Wilcoxon test are effective tools for comparing two or more population probability distributions from independent random samples. However, these techniques are not always suitable for time series data, which often require specialized methods tailored to their unique characteristics.

To address these challenges, ordinal pattern methods provide a robust alternative. Instead of analyzing the absolute values of a time series, these methods focus on the order relationships among consecutive data points.

This approach effectively captures the underlying dynamics of complex systems and offers several advantages.

The ordinal pattern-based method has become a widely used tool for characterizing complex time series. Since its introduction nearly twenty-three years ago by Bandt and Pompe in their foundational paper [5], it has been successfully applied across various scientific fields, including biomedical signal processing, optical chaos, hydrology, geophysics, econophysics, engineering, and biometrics. It has also been used in the characterization of pseudo-random number generators.

The Bandt and Pompe method successfully analyzes time series by transforming them into ordinal patterns, constructing a histogram, and computing Shannon entropy, making it robust against outliers and independent of predefined models.

Later, Rosso [39] introduced an additional dimension to this analysis Statistical Complexity derived from the same histogram of causal patterns.

Introduction to Ordinal Pattern Analysis

Ordinal patterns are a non-parametric representation of real-valued time series and ordinal patterns are transformations that encode the sorting characteristics of values in $\mathbb{R}^D$ into $D!$ symbols, where D represents for the "Embedding Dimension" and usually ranges between three to six. One of the possible encoding is the set of indexes that sort the D values in non-decreasing order.

To illustrate this idea, let $\boldsymbol{x} = \{x_1, x_2, \dots, x_{(n+D-1)}\}$ be a real valued time series of length $n + D - 1$ without ties. If the $\boldsymbol{x}$ takes infinitely many values, it is common to replace them with a symbol sequence $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ consisting of finitely many symbols and then compute the entropy from this sequence. The corresponding symbol sequence naturally emerges from the time series without requiring any model assumptions. We compute $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ symbols from sub-sequences of embedding dimension D. There are $D!$ possible symbols: $\pi_j \in \boldsymbol{\pi} = (\pi^1, \pi^2, \dots, \pi^{D!})$. The histogram of proportions $h = (h_1, h_2, \dots, h_{D!})$ in which the bin h_ℓ is the proportion of symbols of type π^ℓ of the total number of symbols. For convenience, we will model those symbols as a k dimensional random vector where $k = D!$.

Problem Statement

To illustrate this concept, imagine tracking the mean monthly humidity in Wellington. You want to analyze how humidity changes throughout the year. By examining this data, you can uncover interesting patterns that highlight the variations in humidity across different months.

Month	Mean of relative humidity
January	77.3
February	81.0
March	82.4
April	81.7
May	83.6
June	85.6
July	84.4
August	83.1
September	78.8
October	79.6
November	78.2
December	78.8

Table 1.1: Mean monthly humidity variations in Wellington throughout the year

Mean monthly humidity in Wellington is shown in Figure 1.1

We can convert this actual data into ordinal patterns. To do this, for each month, we determine the order of the humidity values rather than their actual magnitudes. Each three-time-point sequence (which can be adjusted based on preference) is converted into an ordinal pattern. This "embedding dimension" usually varies between 3 and 6, but any dimension is possible. The conversion can be made in any way that uniquely maps the sorting properties of the sub sequence into a symbol. For example, consider the

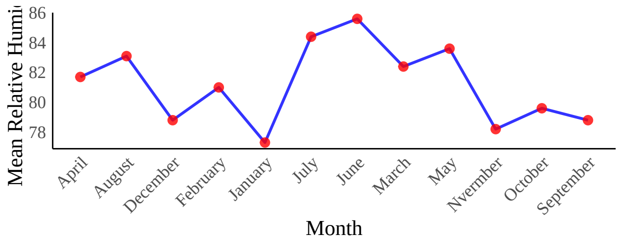

Figure 1.1: Mean Monthly humidity in Wellington

time series presented in Table 1.1. We can transform this series into ordinal patterns as follows. Assume we use patterns of length $D = 3$ with a time lag $\tau = 1$. The first overlapping window $(77.3, 81, 82.4)$ corresponds to the pattern $(1, 2, 3)$, where type is considering with the order of the real time series data. Here 77.3 is the smallest value and is assigned rank 1; 81 is the next highest and is assigned rank 2; and 82.4 is the largest, assigned rank 3. As another example, consider the overlapping window $(83.1, 78.8, 79.6)$. In this case, 78.8 is the smallest value and is assigned rank 1; 79.6 is the next highest, assigned rank 2; and 83.1 is the largest, assigned rank 3. Therefore, the pattern for this window is $(3, 1, 2)$. Table 1.2 has shown this scenario.

t	Mean Humidity sequence	Ordinal Pattern
1	(77.3,81,82.4)	(123) $= \pi^1$
2	(81,82.4,81.7)	(132) $= \pi^2$
3	(82.4,81.7,83.6)	(213) $= \pi^3$
4	(81.7,83.6,85.6)	(123) $= \pi^1$
5	(83.6,85.6,84.4)	(132) $= \pi^2$
6	(85.6,84.4,83.1)	(321) $= \pi^6$
7	(84.4,83.1,78.8)	(321) $= \pi^6$
8	(83.1,78.8,79.6)	(312) $= \pi^5$
9	(78.8,79.6,78.2)	(231) $= \pi^4$
10	(79.6,78.2,78.8)	(312) $= \pi^5$

Table 1.2: Ordinal Patterns

As shown in Table 1.2, we have six mutually exclusive events which we denote as $\{\pi^1, \pi^2, \ldots, \pi^6\} = \{(123), (132), (213), (231), (312), (321)\}$. The probability distribution of the mean humidity is calculated based on ordinal patterns as given below.

$$\hat{p}_i = \frac{\#\{\pi_j \in \boldsymbol{\pi} : \pi_j = \pi^i\}}{n}; 1 \leq i \leq 6, \tag{1.1}$$

where $\hat{\boldsymbol{p}} = (\hat{p}_1, \ldots, \hat{p}_6)$.

Notation	Probability
$p(\pi^1)$	$\frac{2}{10}$
$p(\pi^2)$	$\frac{2}{10}$
$p(\pi^3)$	$\frac{1}{10}$
$p(\pi^4)$	$\frac{1}{10}$
$p(\pi^5)$	$\frac{2}{10}$
$p(\pi^6)$	$\frac{2}{10}$

Table 1.3: Probability function

We construct the histogram of proportions $h = (h_1, h_2, h_3, h_4, h_5, h_6)$, where each bin h_ℓ represents the proportion of symbols of type π^ℓ out of the total six symbols. The histogram graph is shown Figure 1.2.

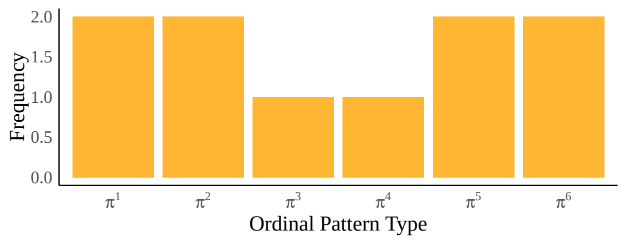

Figure 1.2: Histogram of proportions of the observed patterns according to Table 1.3.

This example explains how time series data can be converted into ordinal patterns and how the probability distribution function can be calculated from these patterns. Chapter Two will review the literature on ordinal pattern analysis, while Chapter Three will expand on this concept by exploring the characterization of time series. It will also cover the computation of two key descriptors entropy and complexity from the resulting histograms. Additionally, Chapter Three will outline the main ideas and objectives of this research project.

Chapter 2

Literature Review

This chapter mainly focus to discuss about the literature review related to this research work. We are analyzing Bandt and Pompe based research work and we found 3

2.1 Preliminaries

sssss

Chapter 3

The Research Project

In this chapter, we outline the main ideas and objectives of this research project. Section 3.1 discusses entropy and complexity analysis in time series, highlighting its advantages and limitations. Section 3.2 examines the advantages and limitations of the Bandt and Pompe method. Section 3.3 provides background knowledge on the entropy-complexity plane. Section 3.4 explores the entropy-complexity plane for a broad class of time series. Section 3.5 provides the asymptotic distribution of the entropy. Finally, the chapter concludes with the objectives of the research project and a case study related to our work.

3.1 Entropy and Complexity Analysis in Time Series: Advantages and Limitations

Entropy and complexity analysis in time series provides powerful tools for measuring the unpredictability and structural richness of dynamical systems. These methods help describe the behavior of a system using mathematical models. Entropy measures, such as Shannon entropy or permutation entropy, quantify the degree of randomness or disorder in the data, while complexity measures assess the balance between order and chaos. Together, they are particularly valuable for detecting nonlinear patterns and distinguishing between deterministic and stochastic behavior. However, their effectiveness depends heavily on parameter choices, such as embedding dimension and time delay. They can be sensitive to noise and data length. Additionally, complexity measures may be computationally intensive (It requires a lot of processing power or time to run) and do not always yield intuitive interpretations such as the results of entropy or complexity are not always easy to understand or explain in simple terms, especially in high-dimensional or multi-scale systems. Despite these limitations, entropy and complexity remain essential in modern time series analysis for uncovering hidden dynamics beyond the reach of traditional linear methods. They also offer a robust framework for enhancing ordinal pattern analysis.

3.2 The Bandt and Pompe Method: A Robust Approach

The concept of ordinal patterns in time series can be effectively studied through real world examples. Traditionally, numerous algorithms, techniques, and heuristics have been employed to estimate complexity measures from real world data. However, these methods often perform well only for low dimensional dynamical systems and struggle when noise is introduced.

The Bandt and Pompe method overcomes this limitation by providing a robust approach that remains reliable even in noisy environments. In time series analysis, key complexity measures such as entropy, fractal dimension, and Lyapunov exponents play a crucial role in comparing neighboring values and uncovering the underlying structure and dynamics of the data.

The advantages of Bandt & Pompe methods:

- Simplicity

- Extremely fast calculation
- Robutness
- Invariance to nonlinear monotonous transformations

This method exhibits low sensitivity to noise and naturally accounts for the causal order of elements in a time series. As a result, it can be applied to various real-world problems, particularly in differentiating between chaotic and stochastic signals.

Despite its limitations, researchers have developed extensions to the original method to address its shortcomings and enhance its applicability to a broader range of complex systems.

3.3 Statistical Complexity measures

After Bandt and Pompe introduced a successful method for analyzing time series, Rosso et al. [39] expanded on this approach by incorporating the statistical complexity derived from the same histogram of causal patterns. This method, which utilizes the entropy-complexity plane, has been successfully applied to various dynamic regimes including, system parameter change [4, 9, 12, 16, 17, 40, 56, 57], optical chaos [24, 46, 48, 52, 55], hydrology [19, 44, 47], geophysics [11, 42, 45], engineering [2, 3, 33, 51], biometrics [41], characterization of pseudo-random number generators [13, 14], biomedical signal analysis [20, 21, 22, 23, 27, 28, 29, 30, 31, 32, 53], econophysics [6, 7, 8, 53, 54, 58, 59].

After computing all symbols as described in Chapter 1, the histogram proportions are used to estimate the probability distribution of ordinal patterns. From this distribution, two key descriptors are calculated to characterize the time series:

1. Entropy (a measure of system disorder)
2. Statistical complexity

The most common metric for the first descriptor is the normalized Shannon entropy, defined as:

$$H(\mathbf{p}) = -\frac{1}{\log k}\sum_{j=1}^{k} p_j \ln p_j. \tag{3.1}$$

Here, $k = D!$ represents the total number of possible permutation patterns. This entropy is bounded within the unit interval:

- It reaches its minimum value $(H = 0)$ when a single pattern dominates for some $p_j = 1$ for some j
- It achieves its maximum $(H = 1)$ under uniform probability $p_j = 1/k$ for all j.

This normalized entropy is often termed permutation entropy in time series analysis.

While normalized Shannon entropy is a powerful tool for quantifying disorder, it fails to fully characterize complex dynamics. To address this limitation, López-Ruiz et al. [25] introduced the disequilibrium Q concept, which quantifies the deviation of a probability distribution $\mathbf{p}$ from a uniform (non-informative) equilibrium state. López-Ruiz and the team employed the Euclidean distance between $\mathbf{p}$ and the uniform distribution, providing a complementary metric to Shannon entropy for assessing structural complexity in systems.

The Jensen-Shannon distance between histogram of proportion $\mathbf{p}$ and the uniform probability function $\mathbf{u} = u_1, u_2, \ldots, u_k$, where $k = D!$ corresponds to the number of possible permutation patterns provides a robust metric for quantifying deviations from uniformity. This distance measure, derived from the symmetric Jensen-Shannon divergence, is particularly suited for analyzing ordinal pattern distributions due to its ability to capture both structural differences and statistical equilibrium in time series data. Here is the formula for Q':

$$Q'(\mathbf{p}, \mathbf{u}) = \sum_{\ell=1}^{k} p_\ell \log \frac{p_\ell}{u_\ell} + u_\ell \log \frac{u_\ell}{p_\ell} \tag{3.2}$$

Lamberti et al. [18] proposed Jensen-Shannon distance as a symmetric metric rooted in the Jensen-Shannon divergence. As the reference model, most works consider the uniform distribution $\mathbf{u} = (1/k, 1/k, \ldots, 1/k)$. The normalized disequilibrium is defined as follows

$$Q = \frac{Q'}{\max(Q')}, \tag{3.3}$$

where $\max(Q')$ is defined as follows

$$\max(Q') = -2\left[\frac{k+1}{k}\log(k+1) - 2\log(2k) + \log k\right]. \tag{3.4}$$

After this concept, Lamberti et al. [18] proposed **Statistical Complexity Measure** which is defined as

$$C = HQ \tag{3.5}$$

where both H and Q are normalized quantities, therefore C is also normalized.

3.4 The Entropy Complexity Plane

The entropy-complexity plane is a two-dimensional representation where time series are mapped based on their permutation entropy and statistical complexity. These metrics are derived from ordinal pattern distributions obtained through time-delay embedding, a technique involving:

- Embedding dimension D: Determines the number of permutation patterns $D!$ used to construct histograms.
- Time delay (τ): Often optimized for specific applications

3.4.1 Key Dynamics in the plane

1. Highly Ordered Systems, where the behavior is very predictable, structured, and often repeats in a regular pattern over time.

Example: Strictly monotonic time series.

- Produces a single ordinal pattern $(H = 0)$.
- Maximal disequilibrium (distance from uniform distribution).
- Maps to $(0, 0)$, indicating minimal complexity.

2. Perfectly Random Systems

   Example: White noise

   - Uniform ordinal pattern distribution $(H = 1)$.
   - Disequilibrium vanishes (distance $= 0$).
   - Maps to $(1, 0)$, reflecting maximal entropy without structural complexity.

The two extreme values are proved by Anteneodo & Plastino [1].Expressions for the boundaries, derived using geometrical arguments within space configurations, were proposed by Martin et al. [26]. These formulations provide a structured approach to understanding and analyzing the spatial behavior of specific systems or models. The lower boundary is characterized by a smooth curve, whereas the upper boundary consists of $D! - 1$ distinct segments. As the embedding dimension D approaches infinity, the upper boundary gradually converges into a smooth curve. Example for the entropy complexity plane is shown in Figure 3.1

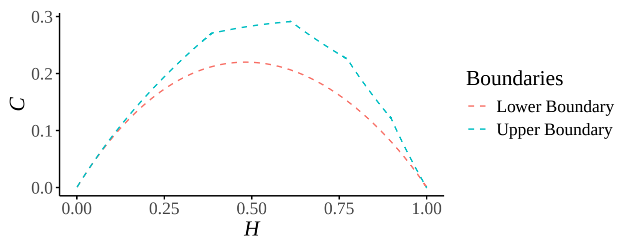

Figure 3.1: Entropy Complexity Plane for Embedding dimension 3

3.5 Asymptotic Distribution of the Shannon Entropy under the Multinomial Distribution

The Multinomial distribution describes how observations fall into categories when an adequate model is available. It is similar to the Multivariate normal distribution, which is one of the continuous Multivariate distributions. Furthermore, it has received considerable attention from researchers, both in theoretical studies and in applications related to discrete Multivariate distributions. The normalized Shannon entropy, often employed in applications like permutation entropy, can be rigorously connected to its asymptotic distribution through the lens of statistical estimation theory. When estimating entropy from finite data, the plug-in estimator (computed directly from observed frequencies) converges to a normal distribution as the sample size $N \longrightarrow \infty$ even for dependent processes such as Markov chains. The Statistical properties of entropy measures under Multinomial distributions are crucial for analyzing complex systems where entropy serves as a key descriptors. Rey at al [37] investigate the asymptotic distributions of various entropy measures specifically, the Rényi and Tsallis entropies of order q, as well as Fisher information when these are computed using maximum likelihood estimators of probabilities from Multinomial random samples. The authors demonstrate that the Tsallis entropy and Fisher information asymptotically follow a normal distribution, whereas the Rényi entropy does not exhibit asymptotic normality. Through simulation studies, the paper validates that these asymptotic models effectively describe a variety of data scenarios. Additionally, the study introduces test statistics for comparing different types of entropies derived from two samples, even when the samples have differing numbers of categories. An application of these tests to social survey data indicates that the results are consistent and offer a more general approach compared to traditional chi-squared tests.

In a subsequent study, Rey et al [35] focus on the statistical complexity measure defined as the product of normalized Shannon entropy and the Normalized Jensen-Shannon divergence between a given probability distribution and the uniform distribution. They derive the asymptotic distribution of this complexity measure under the assumption that the observed data follow a Multinomial distribution. Further the study demonstrates that, as the sample size increases, the distribution of the statistical complexity converges to a normal distribution, with its variance and bias determined by the dynamics of the underlying system. This result provides a theoretical foundation for using statistical complexity as a tool for analyzing systems where the probability distributions are estimated from finite samples. The results are validated with theoretical findings through numerical experiments, showing that the asymptotic normality holds even in scenarios where the Multinomial model is not strictly applicable, such as in applications involving Bandt and Pompe ordinal patterns.

Crucially, the convergence rate and limiting distribution depend on the system's correlation structure deviating from the Multinomial case but remain tractable via spectral analysis of the transition matrix or ordinal pattern distributions [10, 38]. This connection underscores the reliability of normalized entropy measures in large data regimes while highlighting the need to account for dependence structures in finite sample applications.

Imagine a sequence of n independent trials, each resulting in precisely one outcome from a set of k distinct possibilities labeled $\pi^1, \pi^2, \dots, \pi^k$ and so on. These outcomes are mutually exclusive, meaning only one can occur per trial, with respective probabilities $\boldsymbol{p} = \{p_1, p_2, \dots, p_k\}$, such that $p_\ell \geq 0$ and $\sum_{\ell=1}^k p_\ell = 1$. The random vector $\boldsymbol{N} = (N_1, N_2, \dots, N_k)$ counts the number of occurrences of the events $\pi^1, \pi^2, \dots, \pi^k$ in the n trials, with $N_\ell \geq 0$ and $\sum_{\ell=1}^k N_\ell = n$. A $\boldsymbol{n}$ is a sample from $\boldsymbol{N}$ and it has a $k-$variate vector of integer values $\boldsymbol{n} = (n_1, n_2, \dots, n_k)$. Then the joint distribution of $\boldsymbol{N}$ is

$$Pr(\boldsymbol{N} = \boldsymbol{n}) = Pr(N_1 = n_1, N_2 = n_2, \dots, N_k = n_k) = n! \prod_{\ell=1}^{k} \frac{{p_\ell}^{n_\ell}}{n_\ell!} \tag{3.6}$$

This situation is denoted as $\boldsymbol{N} \sim \text{Mult}(n, \boldsymbol{p})$. [37]

In practical applications, the true probability distribution $\boldsymbol{p}$ governing a Multinomial system is typically unknown. Instead, estimators $\widehat{p}_l$, are derived empirically by calculating the observed frequency of each event π^l within the set of k possible outcomes $\boldsymbol{\pi} = \pi^1, \pi^2, \ldots, \pi^k$ across n independent trials. These frequencies approximate the underlying probabilities, enabling inference about the system's behavior. This maximum likelihood estimator (MLE) aligns with the empirical estimator derived from first-moment matching of the distribution. Due to its consistency, asymptotic normality, and computational tractability under regularity conditions, it remains the predominant choice in applied statistical modeling.

Shannon entropy quantifies the level of disorder within a system. When the system's behavior is entirely predictable, the Shannon entropy reaches its minimum, indicating complete knowledge of future observations. Conversely, when the system follows a uniform distribution where all possible outcomes have equal probability—the entropy is maximized, reflecting minimal knowledge about the system's behavior. Chagas et al. [10] have analyzed the asymptotic distribution of Shannon entropy in their study.

Moreover, other types of descriptors, such as Rényi entropy[34], Tsallis entropy[50], and Fisher information [15], have been proposed to extract additional information that is not captured by Shannon entropy. From these entropy measures, Fisher information has garnered more attention due to its unique properties. Fisher information is defined as the average logarithmic derivative of a continuous probability density function.

For discrete probability distributions, Fisher information can be approximated by calculating the differences between probabilities of consecutive distribution elements. A key distinction between Shannon entropy and Fisher information lies in their focus: Shannon entropy quantifies the overall unpredictability of a system, while Fisher information measures the rate of change between consecutive observations, making it more sensitive to small changes and perturbations.

The following equations define Tsallis entropy $(H_T^q(\widehat{\boldsymbol{p}}))$, Rényi entropy $(H_R^q(\widehat{\boldsymbol{p}}))$, and Fisher information measures $(H_F(\widehat{\boldsymbol{p}}))$ [43] :

$$H_T^q(\widehat{\boldsymbol{p}}) = \sum_{\ell=1}^{k} \frac{\widehat{p_\ell} - \widehat{p_\ell}^q}{q-1}, \tag{3.7}$$

where the index $q \in \mathbb{R}\backslash\{1\}$

$$H_R^q(\widehat{\boldsymbol{p}}) = \frac{1}{1-q} \log \sum_{\ell=1}^{k} \widehat{p_\ell}^q, \tag{3.8}$$

where the index $q \in \mathbb{R}^+\backslash\{1\}$

$$H_F(\widehat{\boldsymbol{p}}) = F_0 \sum_{\ell=1}^{k-1} (\sqrt{\widehat{p_{\ell+1}}} - \sqrt{\widehat{p_\ell}})^2 \tag{3.9}$$

where the re-normalization coefficient is $F_0 = 4$ [43]

The main results of the asymptotic distribution of the Shannon, Tsallis, Rényi entropy, and Fisher information are defined as follows [37].

For any $k-$dimensional multivariate normal distribution $\boldsymbol{Z} \sim N(\boldsymbol{\mu}, \sum)$, with $\mu \in \mathbb{R}^k$ and covariance matrix $\sum = (\sigma_{\ell j})$, holds that the distribution of $W = \boldsymbol{a}^T \boldsymbol{Z}$, with $\boldsymbol{a} \in \mathbb{R}^k$, is $N(\boldsymbol{a}^T \boldsymbol{\mu}, \sum_{\ell=1}^{k} a_\ell^2 \sigma_{\ell\ell} + 2 \sum_{j=1}^{k-1} \sum_{\ell=j+1}^{k} a_\ell a_j \sigma_{\ell j})$.

$$H_s(\widehat{\boldsymbol{p}}) = -\sum_{\ell=1}^{k} \widehat{p_\ell} \log \widehat{p_\ell} \tag{3.10}$$

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{\ell=1}^{k} p_\ell (1-p_\ell)(\log p_\ell + 1)^2 - \frac{2}{n} \sum_{j=1}^{k-1} \sum_{\ell=j+1}^{k} p_\ell p_j (\log p_\ell + 1)(\log p_j + 1) \tag{3.11}$$

3.6 Case Study of Asymptotic Distribution of the Complexity

Statistical complexity is defined as the product of two normalized quantities:

- The Shannon entropy
- The Jensen-Shannon distance between the observed probability distribution and the uniform distribution

In this section we discuss three key aspects with real world scenario:

1. **Significance of Asymptotic Complexity Distributions**: Why understanding large-sample behavior matters for statistical inference
2. **Multinomial Model Framework**: Derivation of the asymptotic distribution for statistical complexity under Multinomial assumptions
3. **Practical Formula**: A working equation for calculating the asymptotic distribution of complexity

As a case study for our work, we consider data from the Bearing Data Center and the seeded fault test data from Case Western Reserve University, School of Engineering. The datasets includes ball bearing test data for normal bearings as well as single-point defects on the fan end and drive end. Data were collected at a rate of $48,000(48k$ drive-end) data points per second during bearing tests. Each file contains motor loads $(0, 1, 2,$ and $3)$, drive-end vibration data, and fan-end vibration data. The approximate motor speeds in RPM during testing: $1797, 1772, 1750,$ and 1730. For our case study, we consider two time series (Normal Baseline and $48k$ Drive-End) with a motor load of 0 and an RPM of 1797.

The primary objective of this study is to detect malfunctioning machinery by analyzing two time series using ordinal patterns. We introduce a distance metric based on the ordinal structure of the segments to quantify similarity. This metric facilitates the identification of faulty machines across various embedding dimensions, ranging from 3 to 6. For this case study, we employ an embedding dimension of 3 for convenience; subsequent analyses will extend to the remaining dimensions to compare results. Permutation entropy under asymptotic conditions is computed by considering the probability distribution of ordinal patterns. The results are further analyzed using the complexity–entropy plane, providing insights into the system's dynamics.

Initially, we analyzed the complete datasets from two time series such as one comprising $250,000$ data points representing the normal baseline at motor load 0, and another containing $2,540,000$ data points from the $48k$ drive end under the same motor load. We computed entropy and complexity measures fro these entire datasets, with the results presented in the Table 3.1

Entropy	Complexity	Std Deviation	Semi Length
0.665235	0.226447	0.358893	0.000441
0.772973	0.170954	0.324376	0.001287

Table 3.1: Entropy Complexity Results

Subsequently, we segmented the data into batches of $10,000$ points, categorizing them as either 'Normal' or '48k Drive End'. We then performed a batch wise comparison of entropy and complexity metrics to identify fault data segments. The normal dataset comprises 25 batches, all corresponding to motor load 0, while the $48k$ drive end dataset includes 254 batches. Due to the extensive volume of entropy and complexity data generated, the complete results table is not included in this report. However, the entropy–complexity plane effectively illustrates both batch-wise and full-data analyses. As depicted in Figure 3.2 below, faulty machines form a distinct cluster in the entropy–complexity plane, highlighting their deviation from normal operational patterns. It is clear from the graph that there are both overlapping and non-overlapping confidence intervals. This indicates that some machines differ significantly, while others do not. The main purpose of our experiment is to identify faulty machines. Therefore, we highly recommend extending these results by increasing the embedding dimension to better understand the final outcomes. The general framework of this experiment is also provided in this chapter to clarify the main objective of the research.

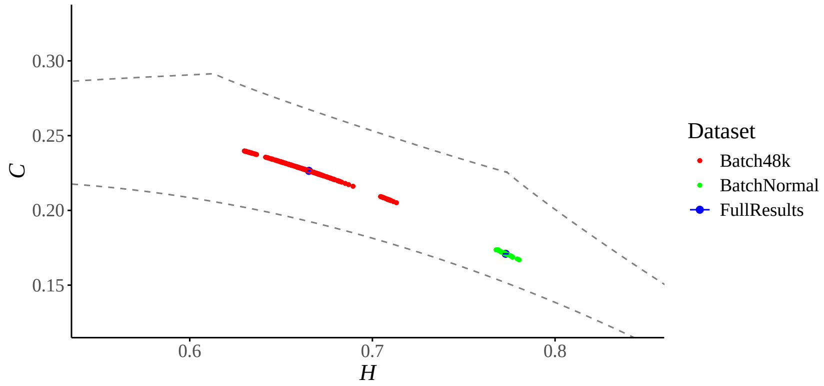

Figure 3.2: Entropy Complexity Plane

The general framework for analyzing entropy-complexity planes with confidence intervals are given as follows.

1. **Calculate Entropy (H) and Complexity (C):** appropriate estimator are Shannon entropy, statistical complexity measures

2. **Compute Confidence Intervals:** Generate multiple resampled datasets to estimate the variance of H and C.

3. **Plot on Entropy-Complexity Plane:**

   - Axes:x-axis: Entropy (H), y-axis: Statistical complexity (C)
   - Data Points: Plot individual or aggregated results.
   - Confidence Regions: Represent uncertainty

4. **Interpretation**

Region of Plane	Interpretation
High H and High C	Complex, structured systems
Low H and Low C	Simple, predictable systems
High H and Low C	Random/noisy systems
Low H and High C	Non-random systems

5. **Statistical testing:**
   - Compare confidence intervals between groups to assess significant differences.
   - Overlapping intervals $\rightarrow$ No significant difference.
   - Non-overlapping intervals $\rightarrow$ Potential significance.

Chapter 4

Future Works

As a short summary, we have completed the necessary preliminaries studies on various topics such as:

- Entropy
- Complexity
- Entropy Complexity Plane
- Confidence interval
- Multinomial distribution.

Also we have examined the relevant research articles by Bandt & Pompe [5], and statistical properties of the entropy from ordinal patterns, asymptotic distribution of certain types of entropy under the Multinomial law, and the asymptotic distribution of the permutation entropy studied by Rey et.al [35, 36, 37].

This proposal has three objectives in order to continue this research work.

- Define a data base of time series for clustering, i.e., finding similar time series.

- Extract all the features we know from their Bandt & Pompe symbolization (Shannon, Tsallis and Renyi entropies, Fisher information measure, complexities, and the available confidence intervals)
- Use those features for time series clustering

Bibliography

[1] ANTENEODO, C., AND PLASTINO, A. R. Some features of the lópez-ruiz-mancini-calbet (lmc) statistical measure of complexity. *Physics Letters A 223*, 5 (1996), 348–354.

[2] AQUINO, A. L., RAMOS, H. S., FRERY, A. C., VIANA, L. P., CAVALCANTE, T. S., AND ROSSO, O. A. Characterization of electric load with information theory quantifiers. *Physica A: Statistical Mechanics and its Applications 465* (2017), 277 – 284. Cited by: 27; All Open Access, Bronze Open Access, Green Open Access.

[3] AQUINO, A. L. L., CAVALCANTE, T. S. G., ALMEIDA, E. S., FRERY, A. C., AND ROSSO, O. A. Characterization of vehicle behavior with information theory. *European Physical Journal B 88*, 10 (2015). Cited by: 21.

[4] BANDT, C. Ordinal time series analysis. *Ecological Modelling 182*, 3-4 (2005), 229 – 238. Cited by: 123.

[5] BANDT, C., AND POMPE, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett. 88* (Apr 2002), 174102.

[6] BARIVIERA, A. F., GUERCIO, M. B., MARTINEZ, L. B., AND ROSSO, O. A. The (in)visible hand in the libor market: an information theory approach. *European Physical Journal B 88*, 8 (2015). Cited by: 24; All Open Access, Green Open Access.

[7] BARIVIERA, A. F., GUERCIO, M. B., MARTINEZ, L. B., AND ROSSO, O. A. A permutation information theory tour through different interest rate maturities: The libor case. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 373*, 2056 (2015). Cited by: 34; All Open Access, Green Open Access.

[8] BARIVIERA, A. F., GUERCIO, M. B., MARTINEZ, L. B., AND ROSSO, O. A. Libor at crossroads: Stochastic switching detection using information theory quantifiers. *Chaos, Solitons and Fractals 88* (2016), 172 – 182. Cited by: 9; All Open Access, Green Open Access.

[9] CAO, Y., TUNG, W.-W., GAO, J., PROTOPOPESCU, V., AND HIVELY, L. Detecting dynamical changes in time series using the permutation entropy. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 70*, 4 2 (2004), 046217–1–046217–7. Cited by: 385.

[10] CHAGAS, E., FRERY, A., GAMBINI, J., LUCINI, M., RAMOS, H., AND REY, A. Statistical properties of the entropy from ordinal patterns. *Chaos 32*, 11 (2022). cited By 4.

[11] CONSOLINI, G., AND DE MICHELIS, P. Permutation entropy analysis of complex magnetospheric dynamics. *Journal of Atmospheric and Solar-Terrestrial Physics 115-116* (2014), 25 – 31. Cited by: 19.

[12] DE MICCO, L., FERNÁNDEZ, J. G., LARRONDO, H. A., PLASTINO, A., AND ROSSO, O. A. Sampling period, statistical complexity, and chaotic attractors. *Physica A: Statistical Mechanics and its Applications 391*, 8 (2012), 2564 – 2575. Cited by: 32.

[13] DE MICCO, L., GONZÁLEZ, C., LARRONDO, H., MARTIN, M., PLASTINO, A., AND ROSSO, O. Randomizing nonlinear maps via symbolic dynamics. *Physica A: Statistical Mechanics and its Applications 387*, 14 (2008), 3373 – 3383. Cited by: 55.

[14] DE MICCO, L., LARRONDO, H., PLASTINO, A., AND ROSSO, O. Quantifiers for randomness of chaotic pseudo-random number generators. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 367*, 1901 (2009), 3281 – 3296. Cited by: 29; All Open Access, Green Open Access.

[15] FRIEDEN, B. R. *Science from Fisher information: a unification*. Cambridge University Press, 2004.

[16] KOWALSKI, A., MARTÍN, M., PLASTINO, A., AND ROSSO, O. Bandt-pompe approach to the classical-quantum transition. *Physica D: Nonlinear Phenomena 233*, 1 (2007), 21 – 31. Cited by: 81.

[17] KOWALSKI, A., MARTÍN, M., PLASTINO, A., AND ROSSO, O. Fisher information description of the classicalquantal transition. *Physica A: Statistical Mechanics and its Applications 390*, 12 (2011), 2435 – 2441. Cited by: 2; All Open Access, Green Open Access.

[18] LAMBERTI, P., MARTIN, M., PLASTINO, A., AND ROSSO, O. Intensive entropic non-triviality measure. *Physica A: Statistical Mechanics and its Applications 334*, 1-2 (2004), 119–131.

[19] LANGE, H., ROSSO, O., AND HAUHS, M. Ordinal pattern and statistical complexity analysis of daily stream flow time series. *European Physical Journal: Special Topics 222*, 2 (2013), 535 – 552. Cited by: 29.

[20] LI, J., YAN, J., LIU, X., AND OUYANG, G. Using permutation entropy to measure the changes in eeg signals during absence seizures. *Entropy 16*, 6 (2014), 3049 – 3061. Cited by: 99; All Open Access, Gold Open Access, Green Open Access.

[21] LI, X., CUI, S., AND VOSS, L. J. Using permutation entropy to measure the electroencephalographic effects of sevoflurane. *Anesthesiology 109*, 3 (2008), 448 – 456. Cited by: 185; All Open Access, Bronze Open Access.

[22] LI, X., OUYANG, G., AND RICHARDS, D. A. Predictability analysis of absence seizures with permutation entropy. *Epilepsy Research 77*, 1 (2007), 70 – 74. Cited by: 254.

[23] LIANG, Z., WANG, Y., SUN, X., LI, D., VOSS, L. J., SLEIGH, J. W., HAGIHIRA, S., AND LI, X. Eeg entropy measures in anesthesia. *Frontiers in Computational Neuroscience 9*, JAN (2015). Cited by: 239; All Open Access, Gold Open Access, Green Open Access.

[24] LIU, H., REN, B., ZHAO, Q., AND LI, N. Characterizing the optical chaos in a special type of small networks of semiconductor lasers using permutation entropy. *Optics Communications 359* (2016), 79 – 84. Cited by: 25.

[25] LOPEZ-RUIZ, R., MANCINI, H. L., AND CALBET, X. A statistical measure of complexity. *Physics letters A 209*, 5-6 (1995), 321–326.

[26] MARTIN, M., PLASTINO, A., AND ROSSO, O. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A: Statistical Mechanics and its Applications 369*, 2 (2006), 439 – 462. Cited by: 293.

[27] MONTANI, F., BARAVALLE, R., MONTANGIE, L., AND ROSSO, O. A. Causal information quantification of prominent dynamical features of biological neurons. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 373*, 2056 (2015). Cited by: 27; All Open Access, Bronze Open Access, Green Open Access.

[28] MONTANI, F., DELEGLISE, E. B., AND ROSSO, O. A. Efficiency characterization of a large neuronal network: A causal information approach. *Physica A: Statistical Mechanics and its Applications 401* (2014), 58 – 70. Cited by: 29; All Open Access, Green Open Access.

[29] MONTANI, F., AND ROSSO, O. A. Entropy-complexity characterization of brain development in chickens. *Entropy 16*, 8 (2014), 4677 –

4692. Cited by: 30; All Open Access, Gold Open Access, Green Open Access.

[30] MONTANI, F., ROSSO, O. A., MATIAS, F. S., BRESSLER, S. L., AND MIRASSO, C. R. A symbolic information approach to determine anticipated and delayed synchronization in neuronal circuit models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 373*, 2056 (2015). Cited by: 25; All Open Access, Green Open Access.

[31] MORABITO, F. C., LABATE, D., LA FORESTA, F., BRAMANTI, A., MORABITO, G., AND PALAMARA, I. Multivariate multi-scale permutation entropy for complexity analysis of alzheimer's disease eeg. *Entropy 14*, 7 (2012), 1186 – 1202. Cited by: 231; All Open Access, Gold Open Access, Green Open Access.

[32] PARLITZ, U., BERG, S., LUTHER, S., SCHIRDEWAN, A., KURTHS, J., AND WESSEL, N. Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics. *Computers in Biology and Medicine 42*, 3 (2012), 319 – 327. Cited by: 160.

[33] REDELICO, F. O., TRAVERSARO, F., OYARZABAL, N., VILABOA, I., AND ROSSO, O. A. Evaluation of the status of rotary machines by time causal information theory quantifiers. *Physica A: Statistical Mechanics and its Applications 470* (2017), 321 – 329. Cited by: 7; All Open Access, Green Open Access.

[34] RÉNYI, A. On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics* (1961), vol. 4, University of California Press, pp. 547–562.

[35] REY, A., FRERY, A. C., AND GAMBINI, J. Asymptotic distribution of the statistical complexity under the multinomial law. *Chaos, Solitons Fractals 193* (2025), 116085.

[36] REY, A. A., FRERY, A. C., GAMBINI, J., AND LUCINI, M. M. The asymptotic distribution of the permutation entropy. *Chaos: An Interdisciplinary Journal of Nonlinear Science 33*, 11 (11 2023), 113108.

[37] REY, A. A., FRERY, A. C., LUCINI, M., GAMBINI, J., CHAGAS, E. T. C., AND RAMOS, H. S. Asymptotic distribution of certain types of entropy under the multinomial law. *Entropy 25*, 5 (2023).

[38] RICCI, L. Asymptotic distribution of sample shannon entropy in the case of an underlying finite, regular markov chain. *Phys. Rev. E 103* (Feb 2021), 022215.

[39] ROSSO, O., LARRONDO, H., MARTIN, M., PLASTINO, A., AND FUENTES, M. Distinguishing noise from chaos. *Physical Review Letters 99*, 15 (2007). Cited by: 537; All Open Access, Green Open Access.

[40] ROSSO, O. A., DE MICCO, L., PLASTINO, A., AND LARRONDO, H. A. Info-quantifiers' map-characterization revisited. *Physica A: Statistical Mechanics and its Applications 389*, 21 (2010), 4604 – 4612. Cited by: 30.

[41] ROSSO, O. A., OSPINA, R., AND FRERY, A. C. Classification and verification of handwritten signatures with time causal information theory quantifiers. *PLoS ONE 11*, 12 (2016). Cited by: 30; All Open Access, Gold Open Access, Green Open Access.

[42] SACO, P. M., CARPI, L. C., FIGLIOLA, A., SERRANO, E., AND ROSSO, O. A. Entropy analysis of the dynamics of el niño/southern oscillation during the holocene. *Physica A: Statistical Mechanics and its Applications 389*, 21 (2010), 5022 – 5027. Cited by: 70.

[43] SÁNCHEZ-MORENO, P., YÁNEZ, R., AND DEHESA, J. Discrete densities and fisher information. In *Proceedings of the 14th International Conference on Difference Equations and Applications. Difference Equations and Applications. Istanbul, Turkey: Bahçesehir University Press* (2009), pp. 291–298.

[44] SERINALDI, F., ZUNINO, L., AND ROSSO, O. A. Complexity–entropy analysis of daily stream flow time series in the continental united states. *Stochastic Environmental Research and Risk Assessment 28*, 7 (2014), 1685 – 1708. Cited by: 53.

[45] SIPPEL, S., LANGE, H., MAHECHA, M. D., HAUHS, M., BODESHEIM, P., KAMINSKI, T., GANS, F., AND ROSSO, O. A. Diagnosing the dynamics of observed and simulated ecosystem gross primary productivity with time causal information theory quantifiers. *PLoS ONE 11*, 10 (2016). Cited by: 19; All Open Access, Gold Open Access, Green Open Access.

[46] SORIANO, M. C., ZUNINO, L., ROSSO, O. A., FISCHER, I., AND MIRASSO, C. R. Time scales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis. *IEEE Journal of Quantum Electronics 47*, 2 (2011), 252 – 261. Cited by: 170; All Open Access, Green Open Access.

[47] STOSIC, T., TELESCA, L., DE SOUZA FERREIRA, D. V., AND STOSIC, B. Investigating anthropically induced effects in streamflow dynamics by using permutation entropy and statistical complexity analysis: A case study. *Journal of Hydrology 540* (2016), 1136 – 1145. Cited by: 43.

[48] TOOMEY, J., AND KANE, D. Mapping the dynamic complexity of a semiconductor laser with optical feedback using permutation entropy. *Optics Express 22*, 2 (2014), 1713 – 1725. Cited by: 94; All Open Access, Gold Open Access.

[49] TREITEL, S. Spectral analysis for physical applications: multitaper and conventional univariate techniques. *American Scientist 83*, 2 (1995), 195–197.

[50] TSALLIS, C. Possible generalization of boltzmann-gibbs statistics. *Journal of statistical physics 52* (1988), 479–487.

[51] YAN, R., LIU, Y., AND GAO, R. X. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines. *Mechanical Systems and Signal Processing 29* (2012), 474 – 484. Cited by: 375.

[52] YANG, L., PAN, W., YAN, L., LUO, B., AND LI, N. Mapping the dynamic complexity and synchronization in unidirectionally coupled external-cavity semiconductor lasers using permutation entropy. *Journal of the Optical Society of America B: Optical Physics 32*, 7 (2015), 1463 – 1470. Cited by: 4.

[53] ZANIN, M., ZUNINO, L., ROSSO, O. A., AND PAPO, D. Permutation entropy and its main biomedical and econophysics applications: A review. *Entropy 14*, 8 (2012), 1553 – 1577. Cited by: 522; All Open Access, Gold Open Access, Green Open Access.

[54] ZUNINO, L., BARIVIERA, A. F., GUERCIO, M. B., MARTINEZ, L. B., AND ROSSO, O. A. Monitoring the informational efficiency of european corporate bond markets with dynamical permutation min-entropy. *Physica A: Statistical Mechanics and its Applications 456* (2016), 1 – 9. Cited by: 24; All Open Access, Green Open Access.

[55] ZUNINO, L., ROSSO, O. A., AND SORIANO, M. C. Characterizing the hyperchaotic dynamics of a semiconductor laser subject to optical feedback via permutation entropy. *IEEE Journal on Selected Topics in Quantum Electronics 17*, 5 (2011), 1250 – 1257. Cited by: 69; All Open Access, Green Open Access.

[56] ZUNINO, L., SORIANO, M., FISCHER, I., ROSSO, O., AND MIRASSO, C. Permutation-information-theory approach to unveil delay dynamics from time-series analysis. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 82*, 4 (2010). Cited by: 198; All Open Access, Green Open Access.

[57] ZUNINO, L., SORIANO, M., AND ROSSO, O. Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 86*, 4 (2012). Cited by: 182; All Open Access, Green Open Access.

[58] ZUNINO, L., ZANIN, M., TABAK, B. M., PÉREZ, D. G., AND ROSSO, O. A. Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A: Statistical Mechanics and its Applications 388*, 14 (2009), 2854 – 2864. Cited by: 199.

[59] ZUNINO, L., ZANIN, M., TABAK, B. M., PÉREZ, D. G., AND ROSSO, O. A. Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A: Statistical Mechanics and its Applications 389*, 9 (2010), 1891 – 1901. Cited by: 183; All Open Access, Green Open Access.