Complexity Science Meets Psychology: Entropy-based Computational Models of Stress Estimation in Humans and Finance

Hongjian Xiao

Supervisor: Prof. Danilo P.Mandic

A thesis submitted for the degree of
Doctor of Philosophy (Ph.D)

Communications and Signal Processing Group
Department of Electrical and Electronic Engineering
Imperial College London

December 2023

Science has explored the micro-cosmos and the macro-cosmos; we have a good sense of the lay of the land. The great unexplored frontier is complexity.

Heinz R. Pagels. 1988

Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line.

Benoit Mandelbrot. 1993

Abstract

The examination of real-world systems through the lens of complexity science has highlighted the need for a transition from qualitative analysis to quantitative evaluation. The well-established "Complexity Loss Theory" serves as a theoretical framework that links changes in structural complexity within dynamic systems with the manifestation of stress at macro levels. This thesis aims to expand upon the concept of complexity loss to compute and model stress levels in two prototypical real-world systems: biological systems in humans and economic systems in finance.

Recent studies in complexity science have underscored the vital role of nonlinear features in discerning between system states. Entropy-based algorithms have garnered significant attention in this context, however, despite substantial progress, practical applications for real-world signal analysis are still underexplored. Given the constraints of traditional entropy measures, this thesis introduces innovative entropy metrics: "ClassA Entropy," "Variational Embedding Multiscale Sample Entropy," and "Multivariate Multiscale Cosines Similarity Entropy". Analysis across several case studies demonstrates that the new measures effectively detect an increase in structural complexity, signifying the occurrence of stress-related stimuli in systems.

The assessment of economic stress through complexity science is a burgeoning field. Various complexity assessment techniques are employed to examine and estimate the non-representational stress in financial indices. This thesis introduces an innovative approach termed "Young's Modulus for Finance" to quantify the robustness of individual stocks/equities. This novel measure can identify the pre-crisis stage, enabling the detection of economic crises and providing a valuable indicator for performance evaluation.

The innovative algorithms presented in this thesis address various limitations of existing entropy measures. The presence of 'open systems' in real-world data sets poses a challenge to many parametric analytic models. Non-parametric and model-free entropy estimation plays unique and crucial roles in the analysis of real-world systems, forming the core focus of this thesis.

Acknowledgements

I extend my deepest gratitude to my supervisor, Prof. Danilo Mandic, for providing me with the opportunity to pursue this Ph.D. and for his unwavering support throughout my entire doctoral research journey. Over the past four years, his invaluable guidance, encouragement, and numerous opportunities have played a pivotal role in shaping the trajectory of my research and fostering my growth as a scholar.

I am thankful to all my colleagues and team members for their assistance in writing publications, engaging in academic discussions, providing technical support for experiments, and their outstanding collaboration on our projects.

Special thanks go to Dr. Tricia Adjei for introducing me to the concept of the Entropy approach and encouraging me to pursue a Ph.D. title. Her guidance has been instrumental in shaping the focus of my research.

Last but not least, I express heartfelt appreciation to my family for their unwavering encouragement, understanding, patience, and financial support throughout this demanding journey. I am particularly grateful to Dr. Michalis Lazarou for his steadfast emotional support in the last year, which has served as a source of strength during both the triumphs and tribulations of the doctoral process.

Hongjian Xiao

December 2023

Statement of Originality

I declare that the work presented in this thesis is the outcome of my research under the supervision of Prof. Danilo P. Mandic. Ideas and results from the work of others are appropriately acknowledged with the standard referencing practices employed in the discipline. the material in the thesis has not been submitted for any degree at any other academic or professional institution.

Copyright Declaration

Contents

List of Tables

List of Figures

List of Abbreviations

AA-PermEn Amplitude-aware Permutation Entropy

ApEn Approximate Entropy

AR autoregressive

AttEn Attention Entropy

BubbEn Bubble Entropy

CCE corrected Condition Entropy

CDV composite delay vector

CE Conditional Entropy

CGP Coase-Graining Proces

ClassA Classidication Angle

CLT Complexity Loss Theory

CSEn Cosine Similarity Entropy

DET Degree of Determinism

DFA Detrended Fluctuation Analysis

DispEn Dispersion Entropy

DistEn Distribution Entropy

DivEn Diversity Entropy

DVV Delay Vector Variance

ECG Electrocardiograph

EEG Electroencephalogram

EMD Empirical Mode Decomposition

EMH Efficient Market Hypothesis

EnoEn Entropy of Entropy

E-PermEn Edge Permutation Entropy

E-R Eckmann-Ruelle

F-DispEn Fluctuation-based Dispersion Entropy

FG-PermEn Fine-grained Permutation Entropy

FS-DispEn Fine-sorted Dispersion Entropy

FuzzEn Fuzzy Entropy

GDE Gridded Distribution Entropy

GFC Global Financial Crisis

HP Hierarchical Process

HRV Heart Rate Variability

iAAFT Iterative amplitude-adjusted Fourier transform

iA-ALIS instantaneous amplitude Assessment of Latent Index of Stress

LAM Laminarity

IAUQ integrated approach based on uniform quantization

IBB Internet Bubble Burst

IBI interbreath interval

ICU intensive care unit

I-PermEn Improved Permutation Entropy

IMEn intrinsic mode entropy

IMFs intrinsic mode function

IncrEn Increment Entropy

I-SODP improved Second Order Difference Plot

K-S Kolmogorov-Sinai

MA Moving Average

MA-MSE Moving Average Multiscale Sample Entropy

MCSE Multiscale cosine similarity entropy

MEP Maximum Entropy Partitioning

MFE multiscale fuzzy entropy

MMCSE multivariate multiscale cosine similarity entropy

MMDispEn Multivariate Multiscale Dispersion Entropy

MMDistEn Multivariate Multiscale Distribution Entropy

MMFE Multivariate Multiscale Fuzzy Entropy

MMPE Multivariate Multiscale Permutation Entropy

MMSE Multivariate Multiscale Sample Entropy

Mod-MSE Modified Multiscale Entropy

M-PermEn Modified Permutation Entropy

MSE Multiscale Sample Entropy

NCDF Normal Cumulative Distribution Function

PermEn Permutation Entropy

pdf probability density function

PhaseEn Phase Entropy

PNS parasympathetic nervous system

RAS Real Angle Sum

R-DispEn Reverse Dispersion Entropy

RP Recurrence Plot

RQA Recurrence Quantification Analysis

RRI RR interval

SampEn Sample Entropy

SE Shannon Entropy

SlopEn Slope Entropy

SNS sympathetic nervous system

SODP Second Order Difference Plot

SyDyEn Symbolic Dynamic Entropy

veMDE Variational Embedding Diversity Entropy

veMSE variational embedding Multiscale Sample Entropy

WGN White Gaussian Noise

W-PermEn Weighted Permutation Entropy

YM Young's Modulus

YMF Young's Modulus for Finance

Chapter 1

Introduction

The term 'complexity' has been widely used in the literature, and complexity analysis has found application in various interdisciplinary fields [1–3]. However, complexity science has been recognized as a crucial complementary tool in real-world signal processing, distinct from analyses based on frequency and time domains. The lack of a unified definition for key concepts has impeded the in-depth development of related studies, and the use of various terminologies has introduced additional challenges in practical applications. To this end, this thesis will commence by providing clear definitions of the keywords used throughout the entire report: Complex system, Entropy, and Stress.

1.1 Complex System & Entropy

While a singular definition of 'complexity' may be elusive, the exploration of complexity science begins with the examination of the properties inherent in complex systems via complexity analysis. In contrast to systems in physics, the defining characteristic of complex systems is their 'open' nature with continuous interactions from the environment that cannot be modelled [4–6]. While models in physics, such as the principle of entropy increase in thermodynamics, are confined to closed and isolated systems [7], real-world systems are inherently complex due to continuous interactions with the environment [5, 8–10]. Another crucial feature of complex systems is adaptation or self-organization [11,12]. The behaviours of complex systems emerge from interactions among components, leading to self-organization towards observed phenomena without the need for a central controller [11, 12]. In light of the understanding of complex systems, the target systems for which complexity analysis is crucial can be summarized as those with a large number of interacting objects that cannot be separately analysed [5, 8].

In the realm of signal processing, complexity analysis serves as a prominent tool for quantifying nonlinear properties of complex systems, encompassing aspects like randomness, regularity, time reversibility, non-linearity, self-correlation, determinism, and other nonlinear features [13–17]. In the context of complexity analysis, the majority of literature measures the degree of complexity by assessing the degree of randomness, which represents the proportion of stochastic components in the system [18, 19]. Nevertheless, measuring randomness is just one facet of the features analysed in complex systems. Consequently, complexity analysis encompasses a variety of methodologies capable of quantifying different nonlinear features of complex systems [15, 20]. Prominent technologies in complexity analysis comprise dimension algorithms [21], Recurrence Plot strategy [22], Detrended Fluctuation Analysis [23], Delay Vector Variance [24], and Entropy [25, 26]. A common dimension algorithm, fractal dimension, has been utilized in systems with chaotic behaviours as an indicator for describing the irregularity of a time series [27]. However, it has been deemed unsuitable for arbitrary series of observations [28]. Recurrence Plot (RP) can identify hidden correlations in multi-dimensional space without the limitations of data stationarity and size restrictions [22]. Detrended Fluctuation Analysis (DFA), another popular dynamical analysis modified based on the random-walk strategy [23], reduces the influence of trends by removing local non-stationarities during the analysis. It is designed to evaluate the degree of irregularity in non-stationary data [23, 29]. Delay Vector Variance (DVV), based on the Iterative amplitude-adjusted Fourier transform (iAAFT) method, was introduced to detect the presence of determinism and nonlinearity.

The output, known as DVV plots, illustrates the differences in properties between the original and surrogate time series [24,30]. Finally, the primary focus of my PhD research will be on entropy-based algorithms as the highlighted dynamical analysis. However, entropy is not a fixed quantity, nor is it an independently analysed methodology. Complexity analysis based on entropy typically incorporates the concepts of fractal dimension and knowledge of statistics [25]. Regardless of prior knowledge, entropy analysis is a model-free method with easy implementation and settings [31–33].

The consensus in the scientific community acknowledges the absence of a direct correspondence between entropy and complexity, primarily stemming from the ambiguous nature of the term 'complexity.' As mentioned earlier, the common misconception of complexity being synonymous with randomness often leads to the perception that uncorrelated random signals, such as White Gaussian Noise (WGN), exhibit the highest entropy and complexity. However, as emphasized in much of the literature, the example of WGN provides a simplistic description with limited relevance to complex systems. It is crucial to note that an increase in the entropy of a system is generally, but not always, synonymous with an increase in complexity [15,34].

The relationship between entropy and complexity is illustrated in Figure 1.1[1]. Two key terminologies require clarification. As depicted in Figure 1.1, traditional entropy refers to the probability distribution of signal magnitudes, indicating the degree of randomness. In other words, as the component of uncorrelated stochastic noise within a signal increases, the entropy of the signal correspondingly increases, resulting in more disordered/random behaviour. As indicated by the monotonically increasing curve, higher entropy implies a greater degree of irregularity and an increased presence of stochastic components in the signal. However, recent literature has increasingly regarded self-correlation/determinism as a more physically meaningful property of complex systems. This aspect denotes the structural richness associated with the amount of information in a signal, referred to as 'structural complexity' [15,35,36]. For instance, the well-known White Gaussian Noise (WGN) exhibits the highest randomness, yet its entirely uncorrelated structure implies the minimal information it could contain. Consequently, the structural complexity of WGN is low despite its high irregularity. The U-shaped curve indicates that structural complexity is expected to be low for both completely ordered and disordered systems. Conversely, all other systems exhibit higher structural richness, indicating a greater amount of information contained in the signal. This encompasses almost all systems in the real world, as they fall within this area.

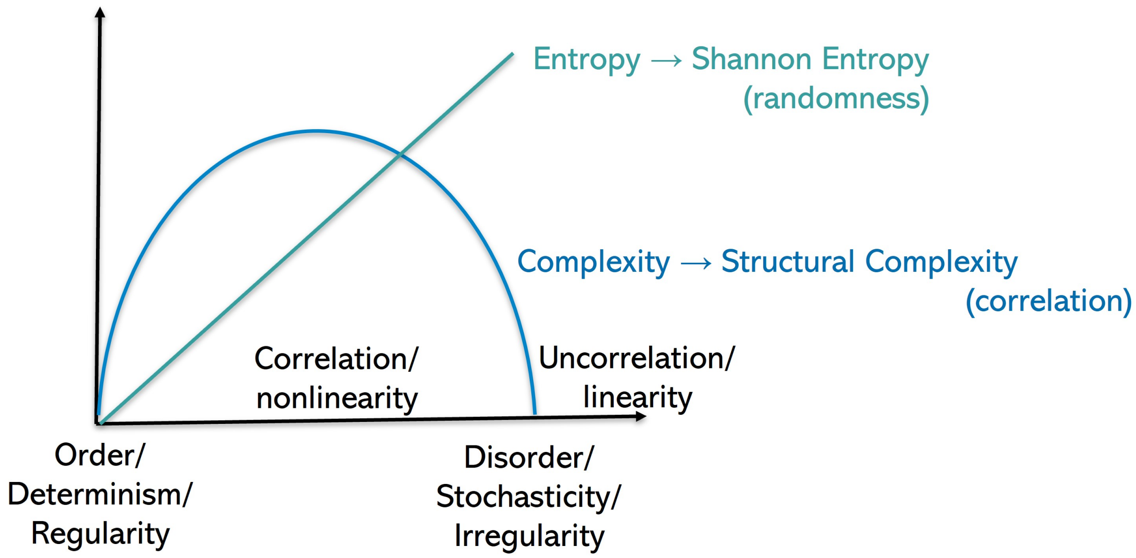


Figure 1.1: Complexity vs Entropy curve, modified based on Kate Jeffery's material. Due to its popularity, the word 'entropy' equally refers to Shannon Entropy in most contexts outside of complexity science, which increases as the system becomes more disordered. However, the level of structure and correlation within the system is inconsistent with the traditional definition of Shannon Entropy. To distinguish the two terminologies, the word 'structural complexity' is used to give the correlation of the system, which is considered as a more physically meaningful feature related to the complexity of the dynamic systems. The level of structural complexity first increases and then decreases as the system approaches complete disorder. The properties of 'Entropy' and 'Complexity' in relation to several nonlinear features are given in the diagram.

Therefore, the curves of entropy and complexity indicate that when a system is in a simple structure and has low determinism, the structural complexity of the system is in line with its entropy. When it comes to highly correlated and irregular systems, the degree of entropy is inconsistent with the structure richness.

[1]The figure is modified based on Kate Jeffery's podcast slides, and the original source is: https://www.preposterousuniverse.com/podcast/2019/10/07/67-kate-jeffery-on-entropy-complexity-and-evolution/

1.2 Stress in Humans

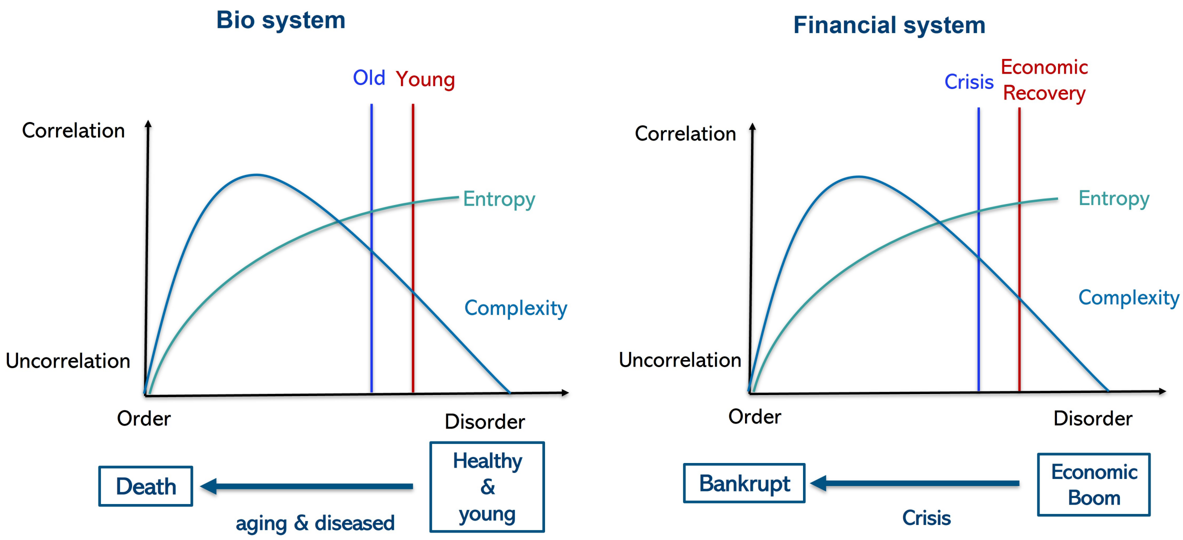
The primary strength of complexity analysis in physiological and physical systems lies in its demonstrated ability to consistently change with pathological states, as indicated by numerous studies [37–44]. As proposed by the randomness-based Complexity Loss Theory (CLT) [45] and a novel correlation-based theory [15], pathological conditions and ageing contribute to a decrease in entropy and an increase in structural complexity. This results in a system that is less random and more correlated.

Illustrated in Figure 1.2(a), a completely disordered state corresponds to young and healthy biosystems. With the emergence of changes in the body due to ageing and disease, organisms may exhibit functional degradation, attributed to limited adaptive responses to the ever-changing environmental stimuli in daily life [38,46,47]. Minor changes in emotions and feelings, manifesting as mental stress and triggered by the external environment, can initiate brain-body interactions, leading to modulations of complex neural bases [48]. These alterations of mind-body states are particularly evident in the cardiovascular system, central nervous system, and motor control system [49]. Given that the human body functions as a holistic system, the relationships between each subsystem are closely interconnected.

For instance, considering the cardiovascular system—one of the most commonly studied subsystems—the Electrocardiograph (ECG) reveals variations in heart rate that can effectively index subtle findings within the inner human body system [13, 50]. An increase in heart rate is primarily linked to the regulation of the sympathetic nervous system (SNS), stimulated by feelings of fear and stress. Conversely, a decrease in heart rate is largely influenced by the control of the parasympathetic nervous system (PNS) branch induced by states similar to relaxation [51]. Nonetheless, the influence of SNS and PNS could be nonreciprocal and coherent [52]. Consequently, the identification of physical states based on the separate analysis of SNS and PNS proves challenging.

Serving as the 'central processing unit' of the human body system, the healthy human brain exhibits remarkable complexity associated with cognitive functions. Furthermore, the sensitivity and specificity of the motor control system are interconnected within the functional modulation network of the central nervous system [47]. Thus, the dynamics of the entire body system offer valuable insights into monitoring physical conditions from an external perspective.

In general, we consider the analysis of complexity as a means to estimate the stress level of biological systems. Ageing and disease lead to a notable increase in stress on human bodies, with even minor changes in stress resulting from cognitive loads generated by physical or mental pathology. Figure 1.2(a) illustrates the crucial relationship between pathological conditions and the complexity of the system. Thus, the life journey of beings represents a process of decreasing entropy. However, despite the high degree of complexity observed between birth and death, the journey begins with minimal complexity and returns to a minimal level at the end.



(**a**) Biological system (**b**) Financial system

Figure 1.2: Complexity VS Entropy in applications on biological systems and financial systems

1.3 Stress in Finance

In addition to biological systems, there is a growing focus on the study of nonlinear features in financial systems [53]. Drawing inspiration from the sympatho-vagal balance of the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) in the human system, the author in [53] applied an analogy of this balance to the 'acceleration-stabilization' dynamic in the financial system. Market expansion can be likened to the activity of the sympathetic nervous system (SNS), while recession is comparable to that of the parasympathetic nervous system (PNS). Analogous to the Complexity Loss Theory, the Efficient Market Hypothesis (EMH) suggests that in the financial system, the 'normal' situation exhibits high irregularity/randomness, whereas the 'abnormal' situation displays more structure.

Regarding the application of complexity science to financial systems, this is a relatively nascent field. Financial markets' behaviours are influenced by multiple interacting factors, including the price histories of assets, traded volumes, dividend yields, recent news, company reports, global climates, and more [4]. Similar to other complex systems, the number of factors is incalculable, and the impact of each factor is inseparable. Therefore, established on the analogy between the financial system and biological systems and aligned with the concept of stress estimation in biological systems, economic systems can also demonstrate varying stress levels determined by the dynamics of price fluctuations [53–55].

In the same complexity-entropy diagram in Figure 1.2(b), the low-stress scenario in a financial system is represented by the period of economic boom, during which the influences of all factors that interactively affect the system's behaviour are dynamically balanced. Conversely, high stress indicates a situation when a financial system is under attack from a crisis. In summary, the occurrence of a crisis or an increase in stress level will result in a decrease in randomness and an increase in correlation, aligning with the outcomes of disease and age in biological systems as depicted in Figure 1.2. Dynamical analysis offers new insights into the estimation of financial stress and crisis events, potentially playing a critical role in predicting future developments in financial markets [2, 53, 56, 57].

1.4 Contributions and Objectives

While the qualitative relationship between entropy and complexity has been proposed, the mathematical quantification of these two curves remains under-explored. Accordingly, the goal of my PhD research is to develop computational approaches for quantifying structural complexity using entropy-based algorithms. Bio-systems and financial systems, as typical complex systems, are the primary fields of application for examining the newly developed complexity analysis methods.

The objectives of this PhD are as follows:

1. Explore and assess the relationship between structural complexity and entropy.

2. Propose a novel multi-channel entropy algorithm that surpasses existing methodologies in stress estimation performance.

3. Verify the proposed algorithm's performance on both biological and financial systems.

It is hypothesised that: i) The nonlinear features of biological systems will be captured by entropy-based complexity analysis. ii) The change of signatures in the financial system driven by crises can be detected by complexity-based methodologies, which can be used as an efficient indicator for stress estimation of the financial market.

The contributions of my PhD research are twofold, summarized and outlined in Figure 1.3: One aspect involves the development of entropy-based algorithms, encompassing both univariate entropy methods and multivariate entropy approaches. The second part centres on the applications of entropy in complex systems, with a primary focus on bio-systems and financial systems. The analyses of bio-systems encompass various physiological and physical signals. In financial systems, the applications involve various nonlinear analysis methods.

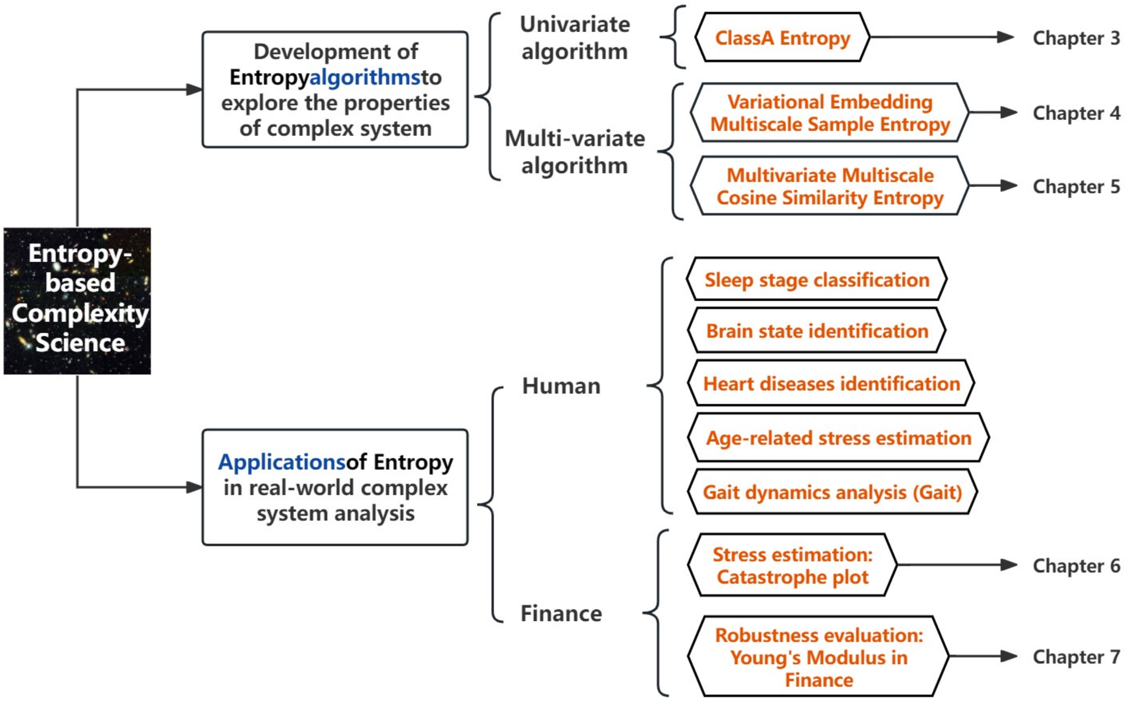


Figure 1.3: PhD objectives & contributions.

1.5 Publications

The following publications support the material given in this thesis:

- **H. Xiao**, and D. P. Mandic, "Variational Embedding Multiscale Sample Entropy: A Tool for Complexity Analysis of Multichannel Systems," *Entropy*, vol. 24, no. 1(26), pp. 1-19, 2022.

- **H. Xiao**, T. Chanwimalueang and D. P. Mandic, "Multivariate Multiscale Cosine Similarity Entropy," *In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal (ICASSP 2022),* pp. 5997-6001, 2022.

- **H. Xiao**, T. Chanwimalueang, and D. P. Mandic, "Multivariate Multiscale Cosine Similarity Entropy and Its Application to Examine Circularity Properties in Division Algebras," *Entropy*, vol. 24, no. 9(1287), pp. 1-19, 2022.

- **H. Xiao**, L. Li and D. P. Mandic, "ClassA Entropy for the Analysis of Structural

Complexity of Physiological Signals," *In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal (ICASSP 2023)*, pp. 1-5, 2023

1.6 Outline of Thesis

- **Chapter 1: Introduction.** This chapter provides an overview of the background of complexity science and introduces the scope of this PhD thesis.
- **Chapter 2: Literature Review: History of Entropy.** This chapter presents a comprehensive literature review of entropy algorithms, ranging from single-scale univariate to multiscale multivariate entropy. Beginning with Shannon Entropy, the chapter lists and demonstrates state-of-the-art entropy algorithms for observed time series using a logistic map.
- **Chapter 3: Univariate Tool: ClassA Entropy.** This chapter introduces a novel univariate tool called ClassA Entropy. The new algorithm is demonstrated and compared with four existing entropy methods. The performance of both the proposed and existing entropy algorithms is assessed using EEG signals from subjects in two states: brain death and coma. This work was published in an IEEE conference.
- **Chapter 4: Variational Embedding Multiscale Sample Entropy.** This chapter presents an innovative multi-channel fusion methodology based on Sample Entropy. It includes a review of the standard MMSE, an evaluation of the new veMSE using real-world non-bio wind data and HRT-based bio signals, and provides details of the experiment and results. This work was published in the Entropy journal.
- **Chapter 5: Multivariate Multiscale Cosine Similarity Entropy.** This chapter introduces a new multivariate entropy algorithm based on Cosine Similarity Entropy, aiming to quantify the structural complexity of multi-channel systems. The new algorithm is compared with standard MMSE, demonstrating its strength

in evaluating the circularity and structure of the system. The work was initially published in an IEEE conference, and an extended journal paper was later published in the Entropy journal.

- **Chapter 6: Financial Stress Evaluation: A Complexity Science Approach.** This chapter provides a comprehensive analysis of financial data using complexity science, incorporating various nonlinear technologies. Additionally, the catastrophe plot is modified and employed to present a novel and comprehensive analysis of the financial market.
- **Chapter 7: Financial Robustness Quantification: A Young's Modulus Inspired Approach.** This chapter introduces a new robustness evaluation index in financial systems, established on entropy-based stress estimation, named Young's Modulus in Finance. It provides details of the analysis and results, which are subsequently revealed and discussed.
- **Chapter 8: Conclusion.** This chapter provides conclusions and summaries of the PhD project, highlighting key remarks and outlines.

Chapter 2

Literature Review: History of Entropy

The concept of entropy originated in thermodynamics to describe the probability distribution of molecules [20]. As a widely applied non-linear analysis, Claude Shannon revolutionized Information Theory in 1948, utilizing entropy to measure the complexity of long sequences of trials by quantifying *the amount of information* [25], as shown in Equation 2.1. Shannon defined the information of a macro-state as a function of the number of possible micro-states that could be sent by the source. The entropy of a source is thus defined in terms of message probabilities without concern for its meaning [12, 25]. Equation 2.1 illustrates that, given the probability distribution of micro-states t for $t \to w$, Shannon Entropy, with the unit in bits, explains how the behaviour of macroscopic entities arises from the statistics of large ensembles of microscopic entities. In the equation, k denotes the Boltzmann constant from Ludwig Boltzmann's entropy definition, and p_t refers to the probability of state t.

$$Shannon\,Entropy : S = -k \sum_{t=1}^{w} p_t \ln p_t \tag{2.1}$$

In 1961, inspired by the principle of fractal dimension, Alfréd Rényi introduced an arbitrary real number q into Equation 2.1 based on the original formula of SE. This led to the proposal of Rényi Entropy, a general form of complexity measure based on information dimension [58] in Equation 2.2. As the parameter q approaches 1, Rényi Entropy is increasingly determined by the events of highest probability p_t, where Rényi Entropy reduces to the original format of SE, serving as the lower bound [58]. It undergoes changes by an additive constant during the re-scaling of variables, while remaining unchanged for different density functions. However, it is noteworthy that Rényi Entropy lacks sub-additivity and recursiveness without possessing the branching and sum properties [59].

$$\textit{Rényi Entropy} : S^{(R)} = \frac{1}{q-1} \log_2(\sum_{t=1}^{w} p_t^q), \quad p_t \neq 0 \tag{2.2}$$

Despite its irreplaceable role, the logarithmic operation in Rényi Entropy makes it unstable when $q \neq 1$, referring to instability under small perturbations, typically noise in physical situations [60]. Therefore, the converted form of Rényi Entropy, known as Tsallis Entropy or q-Entropy in Equation 2.3, removes the logarithmic computation to enhance stability [61]. The two entropy algorithms can be converted according to Equation 2.4. Similar to Rényi Entropy, Tsallis Entropy maintains the parameter q and coincides with SE as the limit $q \to 1$ [62].

$$Tsallis\, Entropy : S^{(T)} = \frac{1}{q-1}(1 - \sum_{t=1}^{w} p_t^q) \tag{2.3}$$

$$S^{(T)} = \frac{1}{q-1}(1 - exp[(q-1)S^{(R)}]) \tag{2.4}$$

Shannon Entropy is considered the cornerstone of entropy in information theory, forming a vital branch of extended entropy algorithms that measure the amount of information. Another key branch of entropy-based methodology is derived from Kolmogorov-Sinai

(K-S) Entropy, which provides *the mean rate of information generation* based on the entropy rate of ordinal patterns [63]. The main concepts of K-S Entropy involve the partition of the phase space of the system and the joint probability of the state in the partitioned space. The K-S Entropy is defined as Equation (2.5), with details of the deduction available in [64]. However, the presentation of K-S Entropy illustrating the infinite condition of order n is numerically infeasible to compute. Consequently, K-S Entropy lacks a direct route or a realistic way for statistical applications [19].

$$\textit{K–S Entropy}: S^{(KS)} = -\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n\delta}\sum_{i=1}^{n} p(k_1,\ldots,k_n)\log p(k_1,\ldots,k_n). \tag{2.5}$$

In 1985, motivated by the above equation, Eckmann and Ruelle defined the E-R Entropy formula to estimate K-S Entropy [65], as shown in Equation (2.6) & (2.7), combined with Taken's Theorem [66]. Taken's Theorem, also known as the delay embedding theorem, reconstructs a sequence of observations of the state of a dynamical system. This reconstruction preserves the geometric shape of structures in phase space, a technique commonly applied in many entropy algorithms [66, 67]. It is important to note that E-R Entropy serves as a lower bound of K-S Entropy. While it is effective in identifying chaotic systems in low dimensions, the presence of superimposed noise in the process can result in E-R Entropy approaching an infinite value.

$$\textit{E–R Entropy}: S^{(ER)} = \lim_{r\to 0}\lim_{m\to 0}\lim_{N\to\infty}[\Phi^{m}(r) - \Phi^{(m+1)}(r)]. \tag{2.6}$$

$$\Phi^{(m)}(r) = \frac{\sum_{i=1}^{N-m+1}\log C_i^m(r)}{N-m+1} \tag{2.7}$$

Recognizing the limitations of K-S Entropy and E-R Entropy, Conditional Entropy was introduced in 1984 [68]. Aligned with Taken's Theorem, Conditional Entropy in Equation

2.8 offers a reliable and practical estimation of K-S Entropy [69]. In the equation, the operator $S(X_m)$ denotes the Shannon Entropy formulation concerning the reconstructed phase space, with the embedding dimension, m, manually set. Therefore, Conditional Entropy serves as the discrete state analogue of K-S Entropy. However, it is important to note that there is no fundamental physical interpretation of Conditional Entropy in a continuous state [70].

$$Conditional\, Entropy : S^{(CE)} = S(X_{m+1}||X_m) = S(X_m, X_{m+1}) - S(X_m) \tag{2.8}$$

In the contemporary landscape of increasing interest in assessing entropy in information theory within observed time series, Shannon Entropy and Conditional Entropy are widely adopted as fundamental measures, representing the amount of information and the rate of information generation, respectively [71]. Broadly speaking, Shannon Entropy-based (SE-based) methods involve calculations within a single reconstructed phase space, while Conditional Entropy-based (CE-based) algorithms entail comparisons across multiple phase spaces. Consequently, the computational load of CE-based methods exceeds that of SE-based ones in most cases. Figure 2.1 illustrates a diagram of entropy algorithms in information theory commonly employed in practical experiments, commencing with two foundational ones: Shannon Entropy (SE) and Conditional Entropy (CE).

The subsequent sections of this chapter will be divided into two parts, focusing on SE-based and CE-based enhanced entropy methods in the analysis of logistic maps. As expressed in Equation 2.9 [72],

$$x_{t+1} = R\, x_t(1 - x_t) \tag{2.9}$$

the behaviour of the logistic map, x_t, presents deterministic chaos and is contingent on the change of parameter, R. The bifurcation diagram in Figure 2.2 summarizes the behaviour of x_t for parameter R in the range of 2-4 with iterations $t = 500$. Starting with an initial

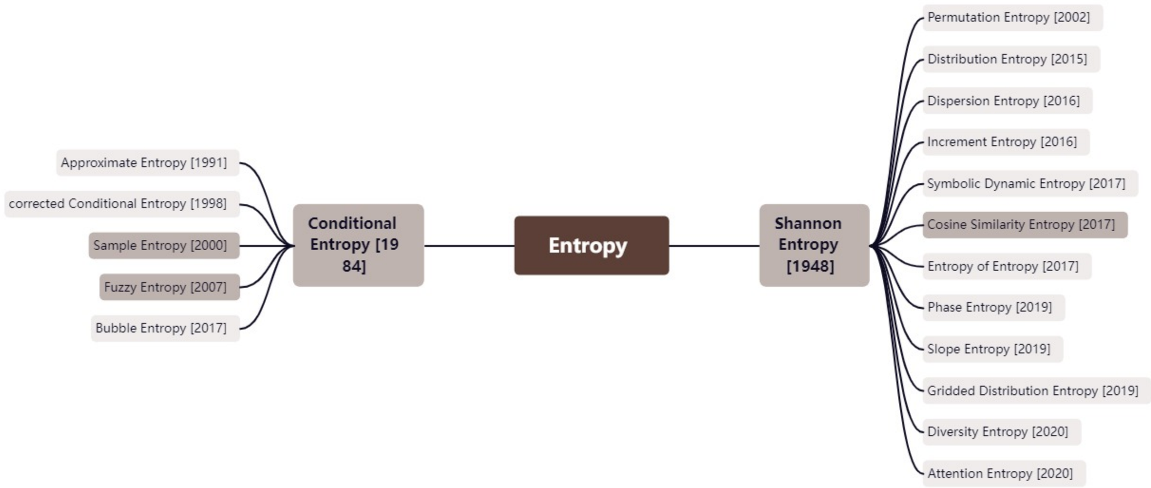


Figure 2.1: Summary of Entropy in Information Theory

value x_0 (ranging from 0-1), when R is set within the range of 2 to 3, x_t will converge to a fixed point after fluctuating around that fixed value for some time. When R is set within the range of 3 to 3.45 (approximately), x_t will approach permanent oscillations between two values. With increasing R, the period of x_t doubles to 4, 8, 16, etc until R reaches 3.5699 (approximately), marking the onset of chaos, where the period of x_t is too large to observe the oscillations.

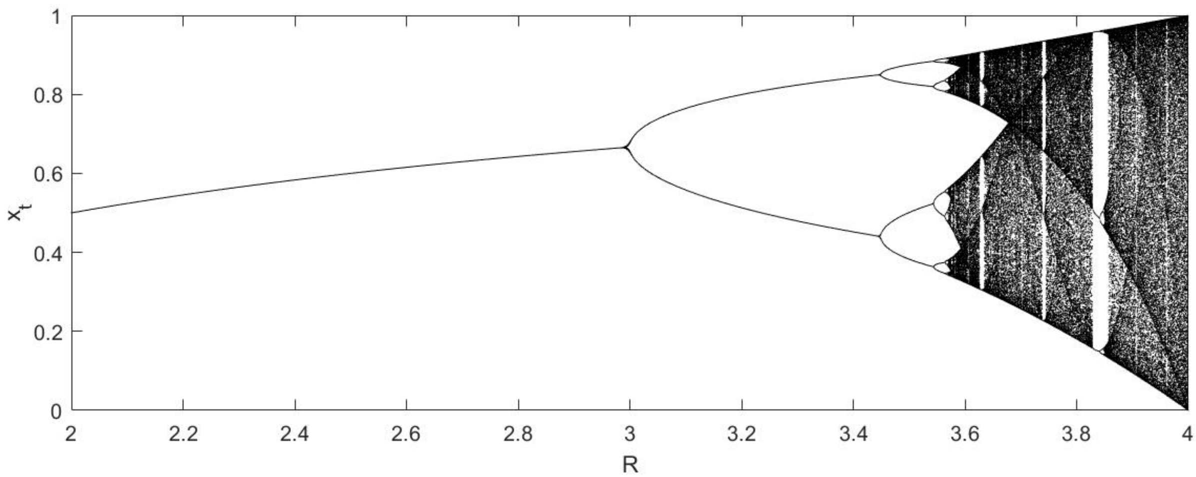


Figure 2.2: Bifurcation diagram of logistic map

The logistic map is a common tool in complexity analysis. Therefore, to showcase the characteristics of each entropy algorithm in the subsequent subsections, the logistic map is employed for univariate single-scale entropy analysis. Before the onset of chaos, the randomness of the signal is expected to be lower than that after. While periodic signals are supposed to show a higher correlation than chaotic signals. To visualize the features given by each entropy method, the bifurcation diagram is presented in line with each figure in the rest of the subsections, depicting R values ranging from 2.5 to 4.

2.1 Shannon Entropy-based algorithms

The standard formulation of Shannon Entropy is presented in Equation 2.1. In real-world signal processing, especially with finite time series and leveraging Taken's Theory [66], practical applications of SE involve characterizing the diversity of embedding vectors based on amplitude. This process includes a manually set parameter known as the embedding dimension, denoted as m. Algorithm 1, provided below, outlines the specifics of SE, noting that it does not consider the comparison of structures in different phase spaces.

Since Shannon's groundbreaking work, entropy has served as a pivotal tool in information theory, aiming to quantify the level of regularity within a system or time series based on the probability distribution of its states [73, 74]. It is crucial that the expectation of information quantity follows a continuous and monotonically increasing function [59]. However, classical SE demands a relatively lengthy time series, making it computationally expensive [75]. Despite its significance in complexity analysis, it has been advised against using SE for real-valued data [76].

2.1.1 Permutation Entropy and its modified versions

In real-world signal analysis, Permutation Entropy (PermEn) stands out as one of the most widely used SE-based enhanced algorithms, finding successful applications in various studies and research endeavours [77–79]. Similar to SE, PermEn relies on the probability

Algorithm 1. Shannon Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the involved parameters are the embedding dimension, m, time delay, l, and the number of classes, c:

1. Normalize the original data sets by subtracting the mean and dividing by the standard deviation.
2. Digitize the normalized data into c classes and the transformed data is denoted as $\{u(i)\}_{i=1}^{N}$.
3. Form the embedding vectors $\mathbf{u}_{(m)}(i)$ derived from the digitized signal $\{u(i)\}_{i=1}^{N}$ by

$$\mathbf{u}_{(m)}(i) = (u(i), u(i+l), \ldots, u(i+(m-1)l)).$$

4. Compute the probability, p_t, of each pattern, t, with respect to the digitized embedding vectors, $\mathbf{u}_{(m)}(i)$. And according to the definition, Shannon Entropy is calculated as

$$Shannon\,Entropy : SE = -k \sum_{t=1}^{w} p_t \ln p_t,\; 1 \leq t \leq w$$

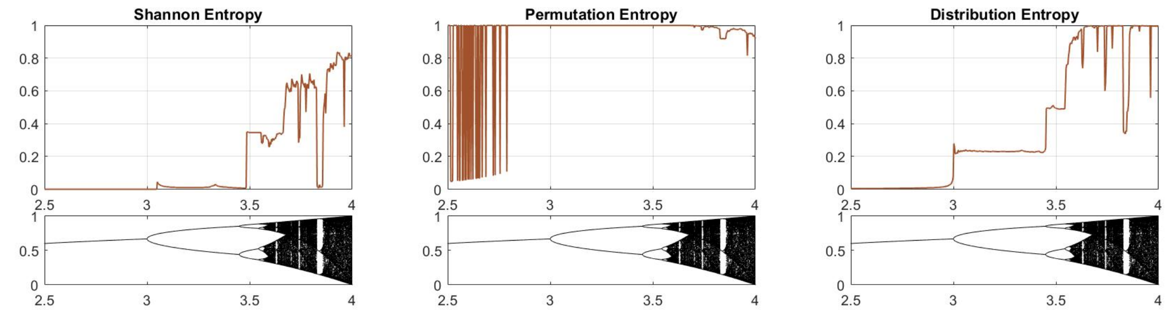
where w refers to the total number of unique patterns in the embedding matrix, $\mathbf{u}_{(m)}$.

distribution linked to the information of ordinal patterns ('motifs') rather than amplitude [28]. The concept behind PermEn involves measuring the relative frequencies of different motifs, granting it structural robustness against noise and artefacts, invariance concerning nonlinear monotonous transformations, and applicability to diverse time series types [59, 80]. Furthermore, it eliminates the need for the customary normalization pre-processing step [37]. PermEn retains the conceptual simplicity of SE, ensuring low computational costs and swift processes without involving the calculation of distance [28]. The algorithm's steps are outlined in Algorithm 2, with the requirement that the data length, N, must exceed the factorial of the embedding dimension, $m!$, a condition typically met. The recommended range for m selection is between 3-7 [28]. In general, a smaller PermEn value indicates a more regular time series. Fig. 2.3(a) showcases the performance of traditional Shannon Entropy, keeping the data length fixed at $N = 5000$ for each calculation. The performance of PermEn on the logistic map, varying with the parameter R, is illustrated in Fig. 2.3(b). As mentioned, SE and PermEn serve as a method to quantify the irregularity, disorder degree, and unpredictability of time series.

Algorithm 2. Permutation Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the involved parameters are the embedding dimension, m and time delay, l:

1. Form the embedding vectors $\mathbf{x}_{(m)}(i)$ derived from the original signal $\{x(i)\}_{i=1}^{N}$ by
$$\mathbf{x}_{(m)}(i) = (x(i), x(i+l), \ldots, x(i+(m-1)l)).$$
2. Find the ordinal patterns ('motifs') $\pi_t = \{\pi_1, \pi_2, ..., \pi_{m!}\}$ for each embedding vector $(\mathbf{x}_{(m)}(i))$ based on amplitude order.
3. Calculate the relative frequency (number of occurrence) $p_t(\pi)$ for each permutation π_t as
$$p_t(\pi_t) = \frac{\#\{\pi_t\}}{N-n},\ n = (m-1)l.$$
4. Permutation Entropy is defined and normalized as
$$PermEn = -\sum_{t=1}^{m!} p_t \log_2 p_t$$
$$PermEn_{norm} = \frac{1}{\log_2 m!} PermEn,\ \ (0 \leq PermEn_{norm} \leq 1)$$



(**a**) SE (**b**) PermEn (**c**) DistEn

Figure 2.3: Behaviour of SE, PermEn and DistEn on the performance of logistic map as a function of the parameter, R. The increase in SE and DistEn signifies higher randomness in chaotic systems, particularly when oscillations cease after R = 3.5699. PermEn in the middle graph demonstrates its capability to detect the reduction in randomness when oscillations appear, as indicated by the presence of the white band in the bifurcation diagram.

The underlying concept of PermEn lies in the idea that the probability of patterns may reveal pertinent knowledge about the underlying dynamical system. Despite the fast computation and simplicity of the original PermEn, defined by Bandt & Pompe in Algorithm 2, two drawbacks in the symbolization process emerge: 1) the disregard of neighbouring amplitude discrepancies, where different neighbouring values were not considered when symbolizing the original time series into patterns. 2) the original symbolization process overlooked the property of equal values, with Bandt & Pompe assuming that the time series consists only of unequal data points [81]. Therefore, considering these drawbacks of the original PermEn, enhanced versions were explored and proposed.

The first algorithm modification is the Fine-grained Permutation Entropy (FG-PermEn), introduced in 2009. This modification involves the introduction of another coefficient, q, which incorporates the amplitude information of the signal and is added to the symbolized embedding matrix [82]. This adjustment rectifies the error where vectors with distinct values are incorrectly mapped onto the same order pattern, rendering FG-PermEn more sensitive to changes in the amplitude of time series compared to PermEn. However, due to the additional processing required for FG-PermEn, it is noted that the time consumption is slightly higher than that needed for calculating conventional PermEn [82].

Subsequently, Bian *et al.* introduced the Modified Permutation Entropy (M-PermEn) method to enhance the symbolization of equal values [83]. This modification involves mapping equal data points onto the same symbol, thereby eliminating bias introduced by equalities and enabling a more accurate characterization of system states. Significant improvements have been demonstrated in the discriminatory ability of Heart Rate Variability (HRV) signals under various physiological and pathological conditions [83].

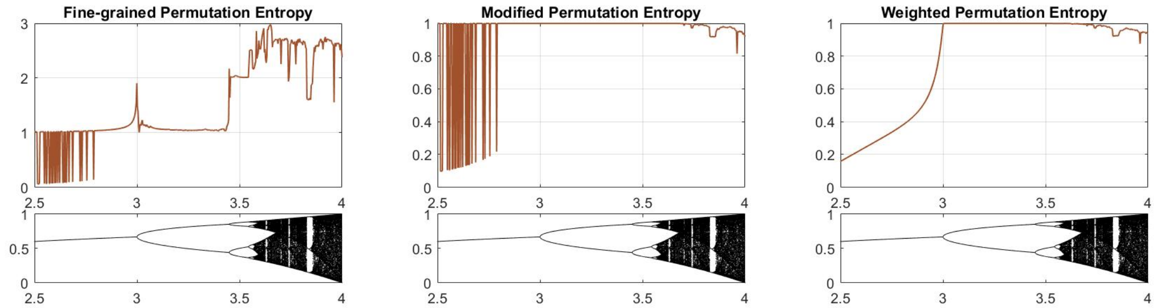
In contrast to FG-PermEn and M-PermEn, Weighted Permutation Entropy (W-PermEn) introduces a weight for each embedding vector, incorporating amplitude information [84]. This allows embedding vectors that symbolize identical patterns in the original PermEn to be distinguished with distinct contributions by different weights. Previous studies have demonstrated that W-PermEn can detect abrupt changes in the signal and effectively distinguish amplitude differences by assigning less complexity to segments that exhibit regularity or are subject to noise effects. W-PermEn is particularly advantageous when applied to signals containing amplitude-coded information, as it demonstrates immunity to degradation by noise and linear distortion [84].

Similar to W-PermEn, Amplitude-aware Permutation Entropy (AA-PermEn) is another PermEn-based algorithm [85]. However, in AA-PermEn, the weights are more carefully chosen based on the joint effects of magnitude and the difference of the embedding vectors, considering an additional coefficient, A. Studies have shown that AA-PermEn is effective in detecting amplitude changes in time series and offers greater flexibility than the original PermEn in pattern quantification. Due to the introduction of the parameter A, it is recommended to choose $A = 5$ for signal segmentation applications, while for spike detection, a value of $A < 0.5$ is suggested [85].

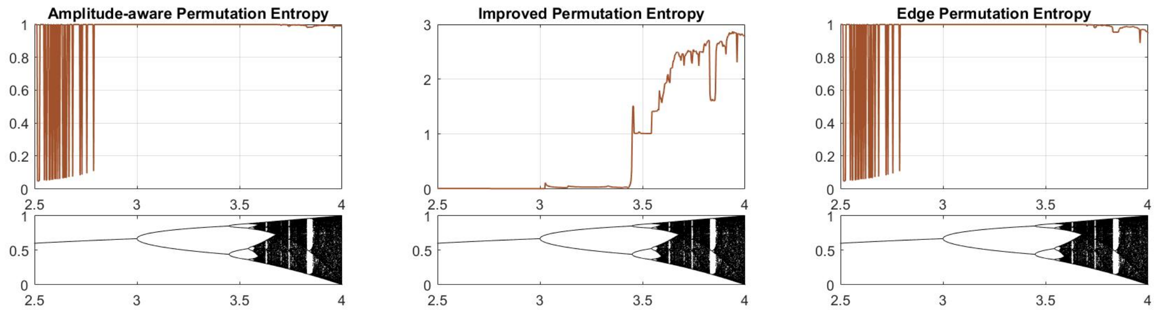
The next modified algorithm is Improved Permutation Entropy (I-PermEn) [86], where symbolization takes into account the difference of neighbouring values. Unlike M-PermEn, which overlooks amplitude information, and W-PermEn and AA-PermEn, which neglect signal fluctuations in their integrated approach based on uniform quantization (IAUQ), I-PermEn addresses these limitations. The proposed I-PermEn tackles the issue of equal values and jointly considers amplitude information and signal fluctuations [86]. As demonstrated by [86], I-PermEn is capable of detecting sharp changes in time series, distinguishing Heart Rate Variability (HRV) signals, and outperforming PermEn under noisy conditions [86].

The last PermEn-based algorithm discussed is Edge Permutation Entropy (E-PermEn), designed to address the loss of amplitude information in classical PermEn using graph theory [87]. In E-PermEn, patterns are symbolized twice, with the second symbolization based on the 'edge' distance between neighbouring elements in undirected graphs. The relationship between two nodes is such that a larger variance between two neighbouring points results in a longer distance between them. In cases of equal values, the difference between two neighbouring nodes is set to 'one', representing the minimum value that edge distance can hold [87].

The performance of modified and enhanced PermEn applied to the logistic map is shown in Fig. 2.4 (a)-(f). FG-PermEn, M-PermEn, AA-PermEn, and E-PermEn exhibit a similar instability in the analysis of short-period signals at the beginning of the map, with FG-PermEn providing the most accurate estimation. Both FG-PermEn and I-PermEn in Fig. 2.4 (a) & (e) demonstrate a moderately improved performance compared to the original PermEn. They effectively capture that chaotic signals exhibit a higher level of entropy and lower predictability than periodic signals in the logistic map analysis [87].



(**a**) FG-PermEn (**b**) M-PermEn (**c**) W-PermEn



(**d**) AA-PermEn (**e**) I-PermEn (**f**) E-PermEn

Figure 2.4: Behaviour of modified versions of PermEn on the performance of the logistic map as a function of parameter, R. M-PermEn, AA-PermEn, and E-PermEn exhibit performance similar to the original PermEn. W-PermEn erroneously assigns high randomness to periodic signals. While FG-PermEn and I-PermEn demonstrate enhancements over the original PermEn, distinguishing among different levels of randomness for each state.

2.1.2 Distribution Entropy

Combining SE with the calculation of the widely used Chebyshev distance, Li *et al.* introduced a novel approach for estimating the irregularity/randomness of time series called Distribution Entropy (DistEn) [88]. DistEn has been applied in the study of Heart Rate Variability (HRV) [89] and is known for its lower sensitivity to predetermined parameter modifications, providing stability in quantifying extremely short time series [88]. Algorithm 3 outlines the details of DistEn, and its performance on the logistic map is depicted in Fig. 2.3(c). In comparison to SE (Fig. 2.3(a)), DistEn exhibits a saturated level when the chaos of the signal reaches maximal resolution based on parameter settings. Additionally, DistEn shows a greater sensitivity to the signal's periodicity compared to SE.

Algorithm 3. Distribution Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the involved parameters are the embedding dimension, m, time delay, l, and the number of classes, T:

1. Form the embedding vectors $\mathbf{x}_{(m)}(i)$ derived from the original signal $\{x(i)\}_{i=1}^{N}$ by

$$\mathbf{x}_{(m)}(i) = (x(i), x(i+l), \ldots, x(i+(m-1)l)).$$

2. Compute the Chebyshev distance between each neighbouring embedding vectors, the obtained distance is donated as $\{D_m(j)\}_{j=1}^{N-n}$, where $n = (m-1) * l$.
3. Digitize the distance, D_m, into T classes by equally spaced classification and compute the probability, p_t, of each digitized distance.
4. Thus, according to the definition of Shannon Entropy in Eq. 2.1, Distribution Entropy is calculated as

$$Distribution\,Entropy : DistEn = -\sum_{t=1}^{T} p_t \log_2 p_t$$

2.1.3 Dispersion Entropy and its modified versions

Recently, Dispersion Entropy (DispEn) has been proposed by Rostaghi & Azami, where DispEn introduces a general approach and a family of DispEn-based algorithms with various mapping functions to symbolize the time series on top of PermEn [73]. Addressing the aforementioned drawback of PermEn, the loss of information regarding amplitude needs to be considered. The improved versions of PermEn, as listed in the previous sub-section, are developed based on the extra column added to the embedding matrix or weights on top of relative frequency, which contain amplitude information. However, limited merits can be observed with an extra computational load. More importantly, a single modification of PermEn has limited improvements in exploring complexity estimation.

Regarding Dispersion Entropy, it applies 'weights' of varying amplitude via mapping functions before the phase space reconstruction and the calculation of probability distribution. Thus, the name Dispersion Entropy is given due to the spread of data in advance. Additionally, the information on amplitude is magnified by the mapping process [73].

The details of Dispersion Entropy are provided in Algorithm 4. It is noteworthy that the original Dispersion Entropy offers a broad range of choices for mapping functions, including both linear and nonlinear techniques. In simulations, mapping functions such as the fast equally spaced function, linear function, and common normal cumulative distribution function (NCDF) were applied, and the corresponding graphs are presented in Fig. 2.5 (a)-(c). Despite requiring less time, Dispersion Entropy with an equally spaced mapping function exhibits instability and erroneously high entropy for highly regular signals. In comparison to PermEn, DispEn with either linear or nonlinear mapping functions demonstrates clear advantages in estimating chaotic signals. The recommended choice for the class number, T, in practical DispEn applications is between 4 and 8 [73].

Building upon classical DispEn, several modified algorithms were proposed. First, inspired by Fine-grained Permutation Entropy [82], Fine-sorted Dispersion Entropy (FS-DispEn) was introduced as an enhanced algorithm, incorporating a factor, f, to measure the difference between neighbouring elements in the vector, added to the mapped Embedding Matrix [90]. The second algorithm is a modified version of DispEn, named Reverse Dispersion Entropy (R-DispEn), defined as the distance to White Noise in the opposite trend compared to DispEn, displaying high stability for time series with varying lengths [91]. Another DispEn-based algorithm is Fluctuation-based Dispersion Entropy (F-DispEn). In contrast to Amplitude-based DispEn, F-DispEn relies solely on the fluctuations of time series, overcoming sensitivity to noise associated with DispEn and providing a better representation of the dynamics within the time series [92]. Fig. 2.5 (d)-(f) illustrates the performance of FS-DispEn, R-DispEn, and F-DispEn, where the

Algorithm 4. Dispersion Entropy

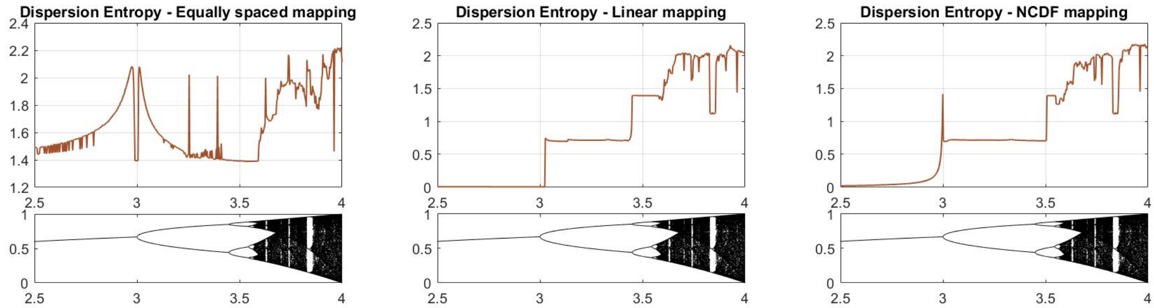
Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m, time delay, l and number of classes, T:

1. Transfer the original time series, $\{x(i)\}_{i=1}^{N}$, into $\{u(i)\}_{i=1}^{N}$, via different mapping functions, e.g., standard normal cumulative distribution function.
2. Form the embedding vectors $\mathbf{u}_{(m)}(i)$ derived from the transferred signal $\{u(i)\}_{i=1}^{N}$ by
$$\mathbf{u}_{(m)}(i) = (u(i), u(i+l), \ldots, u(i+(m-1)l)).$$
3. Equally map $\mathbf{u}_{(m)}(i)$ into T classes donated as $\mathbf{z}_{(m)}(i)$ and find the patterns ('motifs') $\pi_t = \{\pi_1, \pi_2, ..., \pi_{m^T}\}$ for each mapped embedding vector, $z_{(m)}(i)$, based on the transferred amplitude.
4. Calculate the relative frequency (number of occurrence) $p_t(\pi)$ for each permutation π_t as
$$p_t(\pi_t) = \frac{\#\{\pi_t\}}{N-n}, \, n = (m-1)l.$$
5. Dispersion Entropy is defined and normalized as
$$DispEn = -\sum_{t=1}^{m!} p_t \log_2 p_t, \quad DispEn_{norm} = \frac{1}{\log_2 m^T} DispEn$$

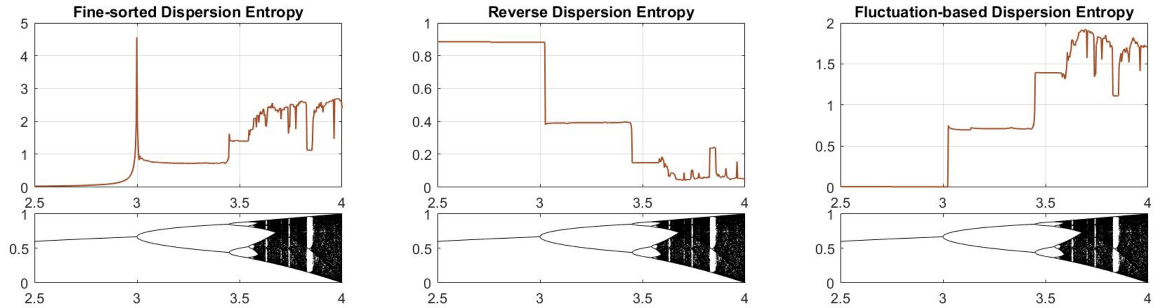
curve of R-DispEn is reversed as expected, and R-DispEn demonstrates outstandingly stable performance within a limited range.

2.1.4 Increment Entropy

To better assess the dynamical property, the difference has been considered a more informative variable for exploring chaotic changes within the signal compared to the original magnitude. In alignment with this perspective, Increment Entropy (IncrEn), as shown in Algorithm 5, was introduced to measure complexity, where "increment" refers to the difference [93]. IncrEn maps the difference of the signal onto a word of two letters, containing information about sign and magnitude, respectively. Then, IncrEn applies SE to the coded words. Therefore, IncrEn has the ability to detect either energetic changes or structural changes in real-world data. Besides, IncrEn exhibits conceptual simplicity,



(**a**) Equally spaced mapping (**b**) Linear mapping (**c**) NCDF mapping



(**d**) FS-DispEn (**e**) R-DispEn (**f**) F-DispEn

Figure 2.5: Behaviour of DispEn with different mapping functions and its modified versions on the performance of logistic map as a function of parameter, R. The performance of DispEn with an equal mapping function is suboptimal compared to the other two methods. Both FS-DispEn and F-DispEn provide similar estimations, except for an unstable transition occurring at $R = 3$ in FS-DispEn. Conversely, R-DispEn exhibits a reversed trend compared to the other graphs, as per its defined characteristics.

with a boundary between 0 and $m \cdot \log(E + 1)$, where E is the resolution suggested to be set lower than 4.

2.1.5 Symbolic Dynamic Entropy

The principle of entropy in time series is the quantification of change. In addition to the difference utilized in IncreEn, the dynamic of a time series can also be probed by the transition of the next state. To address this, Symbolic Dynamic Entropy (SyDyEn) was proposed based on the probability of state transitions [94]. The symbolization method used in SyDyEn is Maximum Entropy Partitioning (MEP) [95], with the number of partitions denoted as c. An improved procedure in SyDyEn involves the inclusion of the state transition probability, p_2, where the transition to the next state for each symbolized

Algorithm 5. Increment Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m, time delay, l, and resolution, E:

1. Obtain the difference of the original signal, $\{x(i)\}_{i=1}^{N}$, denoted as, $\{v(i)\}_{i=1}^{N-1}$.
2. Form the embedding matrix $\mathbf{v}_{(m)}(i)$ derived from the difference signal $\{v(i)\}_{i=1}^{N-1}$ by
$$\mathbf{v}_{(m)}(i) = (v(i), v(i+l), \ldots, x(i+(m-1)l)).$$
3. Obtain the sign and the standard deviation for each embedding vector, $v_{(m)}(i)$, donated as $\mathbf{S}$ and $\mathbf{Sd}$ respectively.
4. Quantify the embedding vector according to E and its $\mathbf{Sd}$ by
$$Q(i) = min(E, \lfloor \frac{||v_i|| \times E}{Sd(i)} \rfloor).$$
where $Q(i)$ is an integer ranging from 0 to E. If the $std(v_i)$ is 0, then, $Q(i)$ equals to 0.
5. Give another matrix denoted as $\mathbf{W}$ containing the information of sign and amplitude of each embedding vectors.
$$\mathbf{W} = \mathbf{S}. * \mathbf{Q}$$
6. Compute the probability, p_t, of each pattern with respect to each vector in W.
7. Increment Entropy is defined and normalized as
$$IncEn = -\sum_{t=1} p_t \log_2 p_t, \quad IncEn_{norm} = \frac{1}{m-1} IncEn$$

embedding vector is considered. Thus, SyDyEn has been proven to better evaluate the amplitude difference of time series as well as the degree of difference. Furthermore, it exhibits noise robustness and resistance to fluctuations [94]. The details of SyDyEn are provided in Algorithm 6.

2.1.6 Cosine Similarity Entropy

All the aforementioned entropy algorithms are based on amplitude information of the time series. However, amplitude-based calculations face challenges, as they are more prone to undefined estimations, lower stability, and weak noise robustness. To address these issues, Cosine Similarity Entropy (CSEn) was introduced in 2017 [15] as shown in

Algorithm 6. Symbolic Dynamic Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m, time delay, l and the number of class, c:

1. Sort the amplitude of the signal and partition them into c cells according to Maximum Entropy Partitioning (MEP) [95].
2. Convert the original time series, $\{x(i)\}_{i=1}^{N}$, into symbolized series, $\{u(i)\}_{i=1}^{N}$, based on the cell index each sample point belonging to.
3. Form the embedding vectors $\mathbf{u}_{(m)}(i)$ derived from the symbolized signal $\{u(i)\}_{i=1}^{N}$ by
$$\mathbf{u}_{(m)}(i) = (u(i), u(i+l), \ldots, u(i+(m-1)l)).$$
4. Compute the probability, $p_1(t)$, of each pattern, with respect to the symbolized data, $\mathbf{u}_{(m)}(i)$.
5. Construct the state transitions, i.e., the next sample point of each pattern, and compute the probability, $p_2(t)$, of the state transitions.
6. Based on the concept of Shannon Entropy, Symbolic Dynamic Entropy is defined and normalized as
$$SyDyEn = -\sum_{t=1}^{\epsilon^m} p_1(t)\log_2 p_1(t) - p_1(t)\sum_{t=1}^{c} log_2(p_1(t)p_2(t))$$
$$SyDyEn_{norm} = \frac{1}{log_2(c^{(m+1)})} SyDyEn$$

Algorithm 7. Instead of amplitude, CSEn involves angular distance, and the limited range of angles, $[0, 2\pi]$, contributes to a more stable and well-defined estimation, especially with short data lengths. Importantly, the angle difference in space allows the assessment of the signal's structure. Therefore, CSEn quantifies the self-correlation of the signal and the structural richness of the system.

IncreEn in Figure. 2.6(a) shows unexpected changes over highly regular signals before the onset of chaotic behaviours. SyDyEn in Figure. 2.6(b) fails to identify the presence of the band of periodic signals in this case. In contrast to amplitude-based entropy measures like IncreEn and SyDyEn, the angle-based CSEn in Figure. 2.6(c) reveals an opposite trend where chaotic signals exhibit lower levels of self-correlation compared to period signals, highlighting the highly correlated structure in period signals.

Algorithm 7. Cosine Similarity Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m, tolerance, r and time delay, l:

1. (Optional) Remove the offset and form the embedding vectors $\mathbf{x}_m(i)$ derived from the zero-median signal $\{u(i)\}_{i=1}^{N}$ by

$$u(i) = x(i) - median(\{x(i)\}_{i=1}^{N}),\ \mathbf{x}_m(i) = (u(i), u(i+l), \ldots, u(i+(m-1)l)).$$

2. Calculate the angular distance for all pairwise embedding vectors $(x_{(m)}(i)\,\&\,x_{(m)}(j))$ based on Cosine Similarity as

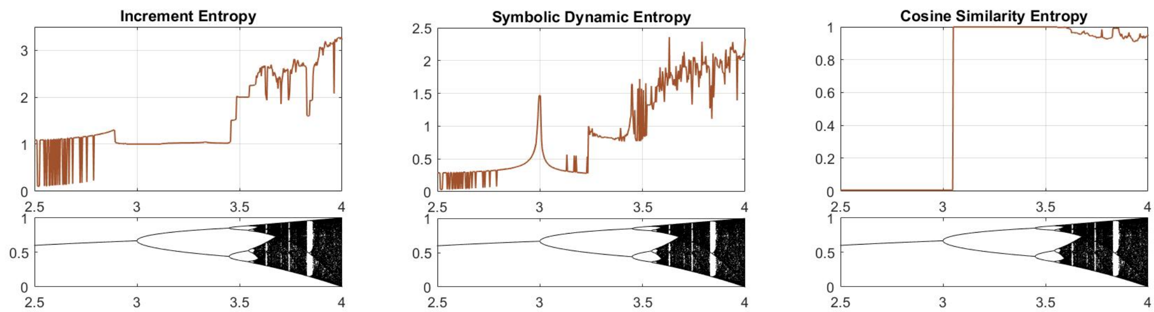
$$d_{(m)}(i,j) = \frac{1}{\pi}\cos^{-1}\left(\frac{x_m(i)\cdot x_m(j)}{|x_m(i)||x_m(j)|}\right), \quad i \neq j.$$

3. Compute the number of similar patterns defined as similar pair $B^r_{(m)}(i)$ that satisfied the criterion $d_{(m)}(i,j) \leq r$.
4. Compute the local probability and global probability of $B^r_{(m)}(i)$ with $n = (m-1)l$ by

$$C^r_m(i) = \frac{B^r_{(m)}(i)}{N-n-1}, \quad \Phi^r_{(m)} = \frac{\sum_{i=1}^{N-n} C^r_{(m)}(i)}{N-n}.$$

5. Cosine Similarity Entropy is defined as

$$CSE = -[\Phi^r_{(m)}\log_2\Phi^r_{(m)} + (1-\Phi^r_{(m)})\log_2(1-\Phi^r_{(m)})].$$



(**a**) IncrEn (**b**) SyDyEn (**c**) CSEn

Figure 2.6: Behaviour of IncrEn, SyDyEn and CSEn on the performance of the logistic map as a function of the parameter, R. IncreEn demonstrates an increase as the system becomes more random. However, SyDyEn fails to accurately identify the location of the oscillation band. On the other hand, CSEn exhibits the highest correlation when the parameter R ranges between 3 and 3.5699, followed by a decrease after the onset of chaos. Subsequently, it increases again during the oscillations in the region represented by the white band.

2.1.7 Entropy of Entropy

While all the previously mentioned algorithms utilize SE once in the final step, there is a method known as Entropy of Entropy (EnoEn) [96] that applies SE twice in the calculation. As outlined in Algorithm 8 and demonstrated in its performance on the logistic map in Figure 2.7(a), this double exploration of distribution probability allows for the extraction of both information and variation hidden within the time series. These aspects are then interpreted as indicative of the system's adaptive capacity.

Algorithm 8. Entropy of Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the mapping scale, τ, matrix of sector number, S, which includes two numbers:

1. Reshape the original time series into matrix, $\{u_\tau(i)\}_{i=1}^{\frac{N}{\tau}}$, with size of $\tau \times \frac{N}{\tau}$, given by the non-overlapping window scale, τ.
2. Calculate the Shannon Entropy of each vector, $Y_\tau(j)$ in $\{u_\tau(j)\}_{j=1}^{\frac{N}{\tau}}$ according to the first given number of slices, $S1$.
3. Compute the Shannon Entropy again on $Y_\tau(j)$ according to the second given number of slices, $S2$. The Entropy of Entropy is defined as

$$EnoEn = -\sum_{t=1} p_t ln(p_t)$$

2.1.8 Phase Entropy & Gridded Distribution Entropy

When the embedding dimension is set as $m = 2$, the original signal is mapped onto a new phase space and can be presented as a 2D Poincaré plot of $x(i+1)$ against $x(i)$ [97]. Poincaré plot, also called a return map, is a commonly used graphical tool for visualizing dynamical signals [97]. Although numerous entropy algorithms have been established based on the embedding space, the utilization of entropy is mostly considered in conjunction with the distance among embedding vectors, which requires a high computational load.

Straightforward applications of Poincaré plots in entropy analysis are limited to a small number of studies; one popular algorithm built on Poincaré plot is the Gridded Distribution Entropy (GDE), which describes entropy over the distribution of data on a gridded 2-D Poincaré plot [98]. This allows for the investigation of the degree of variability in data. However, limited information can be given by Poincaré plot about the states of consecutive acceleration or deceleration of the signal dynamics [20].

To this end, Phase Entropy (PhaseEn) was proposed as a descriptor of the Second Order Difference Plot (SODP) and the angle distribution of data points in the SODP space [20]. Phase Entropy employs the original SODP introduced by Cohen *et al.*, as $(x(i+2)-x(i+1))$ against $(x(i+1)-x(i))$ and exhibits enhanced robustness compared to the Poincaré plot [99]

Algorithms 9 and 10 delineate the details of PhaseEn and GDE, respectively. Their performance is illustrated in Figure 2.7(b) & (c). The results indicate that PhaseEn provides an estimation of structural complexity, whereas GDE offers a measure of randomness complexity.

Algorithm 9. Phase Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameter is the number of group, T:

1. Map the original signal onto Second Order Difference Plot.
2. Calculate the angle of each data point in SODP phase space, $\alpha(i)$.
3. Equally partition the space given by T slices from the origin, and calculate the sum of angles of data point in each sector, $\{S_t\}_{t=1}^{T}$.
4. Normalize $\{S_t\}_{t=1}^{T}$.
5. Phase Entropy is defined as

$$PhaseEn = -\frac{1}{log(T)} \sum_{t=1}^{T} S_t log(S_t).$$

2.1.9 Other SE-based Entropy algorithms

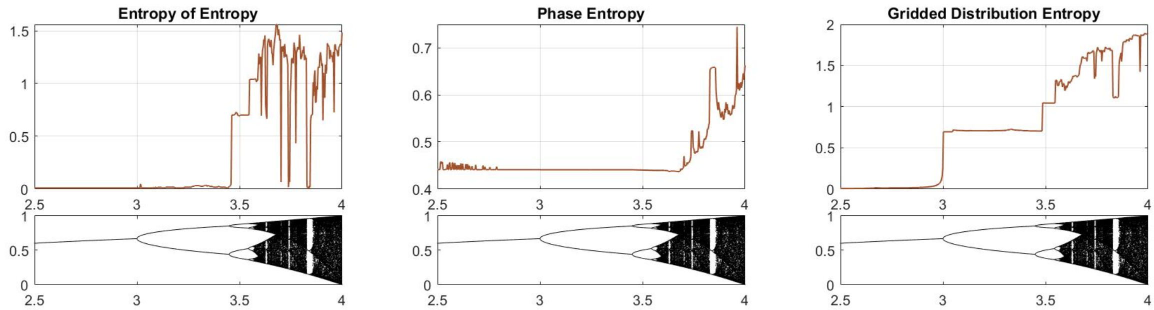
Apart from the previously mentioned entropy methods, this study introduces three novel SE-based algorithms: Slope Entropy (SlopeEn) [100], Diversity Entropy (DivEn) [31], and Attention Entropy (AttEn) [101]. Notably, SlopeEn and AttEn eliminate the need for distance calculations, enhancing computational efficiency. In addition, SlopeEn is

Algorithm 10. Gridded Distribution Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameter is the grid coarse-grain, T:

1. Normalize the original data set, $\{x(i)\}_{i=1}^{N}$ and map the normalized data onto Poincaré plot.
2. Equally partition the phase space into $T \times T$ grid and calculate the probability of data points in each grid, p_t.
3. Gridded Distribution Entropy is defined as

$$GDE = -\sum_{t=1}^{T \times T} p_t log(p_t).$$

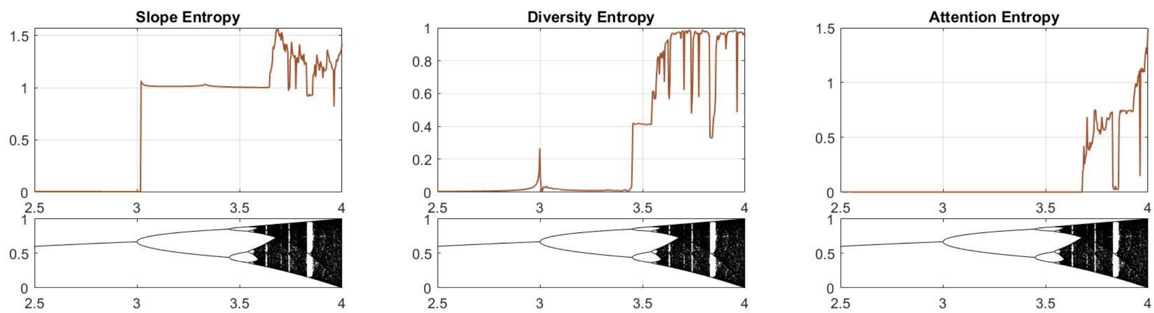


(**a**) EnoEn (**b**) PhaseEn (**c**) GDE

Figure 2.7: Behaviour of EnoEn, PhaseEn and GDE on the performance of logistic map as a function of the parameter, R. EnoEn and GDE exhibit a decrease in randomness during the oscillation band, whereas PhaseEn demonstrates an increase in correlation in the region denoted by the white band in the bifurcation diagram.

crafted based on the angular distribution of data points mapped in the Poincaré Plot.

Details of SlopeEn, DivEn and AttEn are depicted in Algorithm 11-13 and Figure 2.8.



(**a**) SlopeEn (**b**) DivEn (**c**) AttEn

Figure 2.8: Behaviour of SlopeEn, DivEn and AttEn on the performance of logistic map as a function of parameter, R. All three algorithms depict the system's randomness, showing an increase as chaos emerges.

Algorithm 11. Slope Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are angular threshold $[\delta,\gamma]$ and embedding dimension, m:

1. Map the normalized data onto Poincaré plot.
2. Calculate the angle of each data point in the new phase space, $\alpha(i)$.
3. Symbolize the angle signal as follows:

$$u(i) = \begin{cases} 1, & \alpha(i) \in (\delta, \gamma] \\ -1, & \alpha(i) \in (-\gamma, -\delta] \\ 2, & \alpha(i) \in (\gamma, \infty) \\ -1, & \alpha(i) \in (-\infty, -\gamma] \\ 0, & \alpha(i) \in [-\gamma, \gamma) \end{cases}$$

4. Form the embedding vectors based on the symbolized angular signal, $u(i)$, referred to $\mathbf{U}_m(i)$.
5. Calculate the relative frequency (number of occurrences), p_t.
6. Slope Entropy is defined as

$$SlopEn = -\sum p_t \log_2 p_t$$

Algorithm 12. Diversity Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m and the number of intervals, T:

1. Form the embedding vectors $\mathbf{x}_m(i)$ according to the given parameter, m.
2. Calculate the angular distance between the adjacent embedding vectors based on Cosine Similarity as the equation given in Algorithm 7.
3. Partition the scope $[-1, 1]$ into T intervals and calculate the frequency probability of distance series, $d_{(m)}(i)$, falling into each interval, denoted as, p_t, where $\sum_{t=1}^{T} p_t = 1$
4. Diversity Entropy is defined as

$$DivEn = -\frac{1}{\ln T} \sum_{t=1}^{T} p_t \ln p_t$$

2.2 Conditional Entropy-based algorithms

Despite being a variant of Shannon Entropy, Conditional Entropy (CE) assesses complexity by exploring information production or entropy rates through the comparison of varied phase spaces [68, 70]. Numerous studies and applications have been developed

Algorithm 13. Attention Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N:

1. Find all peaks including local maximums and local minimums.
2. Count the number of data points between adjacent maximums, denoted as Y_{xx}. The number of data points between adjacent minimums is referred to as Y_{nn}. The number of data points between adjacent maximums to minimums is marked as Y_{xn} and vice versa, as Y_{nx}.
3. Calculate Shannon Entropy for interval series, Y_{xx} as $H_{xx} = -\sum_{t=1}^{w} p_t \ln p_t$, where w refers to the total number of different patterns in Y_{xx} and p_t denotes the frequency probability.
4. Calculate the Shannon Entropy for interval series, Y_{nn}, Y_{xn}, and Y_{nx}, named as H_{nn}, H_{xn} and H_{nx}.
5. Attention Entropy is defined as the mean value of the four Shannon Entropies

$$AttEn = \frac{H_{xx} + H_{nn} + H_{xn} + H_{nx}}{4}$$

based on CE.

Algorithm 14. Conditional Entropy & corrected Conditional Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ with length of N and the parameters are the embedding dimension, m, time delay, l, and the number of classes, c:

1. Calculate the Shannon Entropy according to the steps given in Algorithm 1 based on embedding dimension, m, referred to SE_m
2. Repeat step 1 and calculate the Shannon Entropy for increased embedding dimension, $m+1$, referred to SE_{m+1}.
3. Calculate the percentage of single points in the time series, and the corrective term is computed as $E_c = p_s \times SE_1$, where SE_1 refers to the Shannon Entropy of original time series as $m = 1$.
4. Conditional Entropy (CE) and corrected Conditional Entropy (CCE) are defined as

$$CE = SE_m - SE_{m+1}, \quad CCE = CE + E_c.$$

In comparison to SE-based Permutation Entropy, Conditional Entropy assesses the diversity of successors of given ordinal patterns π, while Permutation Entropy characterizes the diversity of the ordinal patterns π themselves [102]. Embedding dynamics into a phase space with excessively high dimensions yields entropy rate estimates of zero [103]. To address this limitation, a refined entropy rate estimation named corrected Conditional

Entropy (CCE) was introduced for processing short data sequences. CCE is computed as the sum of Conditional Entropy and a corrective term [103]. Both Conditional Entropy and corrected Conditional Entropy are presented in Algorithm 14. CCE has found application in the analysis of short-term RR intervals associated with systolic arterial pressure oscillations [103, 104].

2.2.1 Approximate Entropy & Sample Entropy

In 1991, Pincus introduced Approximate Entropy (ApEn) for analysing short and noisy data sets [105]. ApEn serves as the continuous state space extension of the entropy rate, recovering the rate of entropy in the discrete state. More broadly, ApEn is associated with Conditional Entropy [70]. The crucial steps of ApEn are defined by Equation (2.10) and estimated by the statistics in Equation (2.11). Analogous to Shannon Entropy, research has demonstrated that a lower ApEn value indicates more regularity, while a higher ApEn value signifies greater randomness [106, 107].

$$S^{(A)} = \lim_{N \to \infty} [\Phi^{(m)}(r) - \Phi^{(m+1)}(r)] \tag{2.10}$$

$$Approximate\, Entropy : ApEn = \Phi^{(m)}(r) - \Phi^{(m+1)}(r) \tag{2.11}$$

Following Approximate Entropy, Richman and Moorman introduced a modified algorithm in 2000 termed Sample Entropy [26], defined in Equation (2.12) with its statistically estimated form in Equation (2.13). A crucial enhancement of Sample Entropy is the exclusion of self-matching, reducing bias in estimation [26]. Additionally, the modification of SampEn from ApEn is reflected by Rényi Entropy, where ApEn can be seen as an approximation of Rényi Entropy with order $q = 1$, while SampEn corresponds to order $q = 2$ [64]. In the realm of short-time series processing, SampEn has demonstrated advantages over ApEn and has been applied in various studies [108, 109]. For instance, research utilizing SampEn of the electrocardiogram (ECG) has estimated different degrees of atrial fibrillation [110], and in the analysis of stress during public speaking [111]. Details of ApEn and SampEn calculation are provided in Algorithm 15 & 16, separately.

$$S^{(S)} = \lim_{N \to \infty} - \ln \frac{\Phi^{(m+1)}(r)}{\Phi^{(m)}(r)} \tag{2.12}$$

$$Sample\, Entropy : SampEn = - \ln \frac{\Phi^{(m+1)}(r)}{\Phi^{(m)}(r)} \tag{2.13}$$

Algorithm 15. Approximate Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N, the parameters involved are the embedding dimension, m, tolerance, r_s, and time delay, l.

1. Construct the embedding matrix, $\mathbf{X}_m$, derived from the original signal, $\{x(i)\}_{i=1}^{N}$, in the form $x_m(i) = (x(i), x(i+l), \ldots, x(i+(m-1)l))$.
2. Compute the distance between all the pairwise embedding vectors, $x_m(i)$ & $x_m(j)$, based on the Chebyshev distance, as $d_m(i,j) = max\{x_m(i+k) - x_m(j+k)|| \, 0 \leq k \leq m-1\}$.
3. Compute the number of matching patterns, $B_m^{r_s}(i)$, defined as the similar pairs that satisfy the criterion $d_m(i,j) \leq r_s$.
4. Compute an estimate of the local probability of $B_m^{r_s}(i)$ as $C_m^{r_s}(i) = \frac{B_m^{r_s}(i)}{N-n}$, $n = (m-1) * l$.
5. Compute an estimate of the global probability of $B_m^{r_s}(i)$ as $\Phi_m^{r_s} = \frac{\sum_{i=1}^{N-n} C_m^{r_s}(i)}{N-n}$.
6. Repeat steps 1–5 with an increased embedding dimension, $m+1$, and obtain the updated global probability, denoted as $\Phi_{m+1}^{r_s} = \frac{\sum_{i=1}^{N-n} C_{m+1}^{r_s}(i)}{N-n}$, $n = m * l$.
7. Approximate entropy is defined as

$$ApEn(m, l, r_s, N) = \Phi_m^{r_s} - \Phi_{m+1}^{r_s}.$$

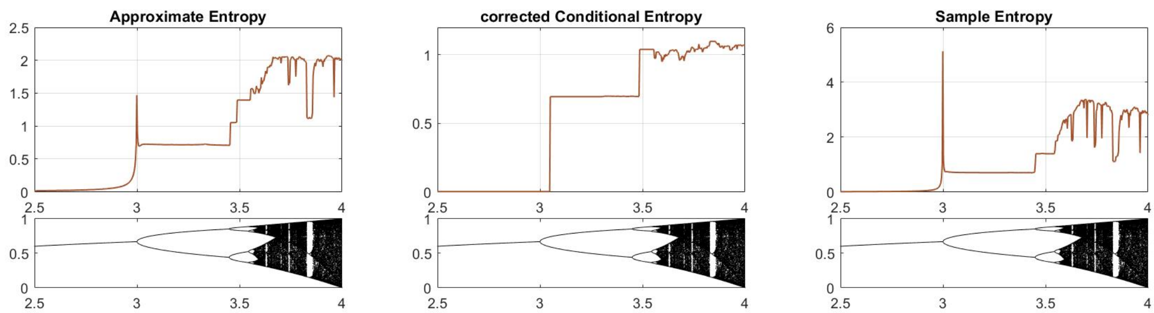
The performance of ApEn, CCE, and SampEn on the logistic map was implemented and is depicted in Figure 2.9. It is evident that both ApEn and SampEn exhibit a superior ability to detect the band of periodic signals compared to CCE, particularly in the middle of the figure.

Algorithm 16. Sample Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N, the parameters involved are the embedding dimension, m, tolerance, r_s, and time delay, l.

1. Construct the embedding matrix, $\mathbf{X}_m$, derived from the original signal, $\{x(i)\}_{i=1}^{N}$, in the form $x_m(i) = (x(i), x(i+l), \ldots, x(i+(m-1)l))$.
2. Compute the distance between all the pairwise embedding vectors, $x_m(i)$ & $x_m(j)$, based on the Chebyshev distance, as $d_m(i,j) = max\{x_m(i+k) - x_m(j+k)|| 0 \leq k \leq m-1, i \neq j\}$.
3. Compute the number of matching patterns, $B_m^{r_s}(i)$, defined as the similar pairs that satisfy the criterion $d_m(i,j) \leq r_s$.
4. Compute an estimate of the local probability of $B_m^{r_s}(i)$ as $C_m^{r_s}(i) = \frac{B_m^{r_s}(i)}{N-n-1}$, $n = (m-1) * l$.
5. Compute an estimate of the global probability of $B_m^{r_s}(i)$ as $\Phi_m^{r_s} = \frac{\sum_{i=1}^{N-n} C_m^{r_s}(i)}{N-n}$.
6. Repeat steps 1–5 with an increased embedding dimension, $m+1$, and obtain the updated global probability, denoted as $\Phi_{m+1}^{r_s} = \frac{\sum_{i=1}^{N-n} C_{m+1}^{r_s}(i)}{N-n}$, $n = m * l$.
7. Sample entropy is defined as

$$SampEn(m, l, r_s, N) = -\ln\left[\frac{\Phi_{m+1}^{r_s}}{\Phi_m^{r_s}}\right].$$



(**a**) ApEn (**b**) CCE (**c**) SampEn

Figure 2.9: Behaviour of ApEn, CCE, and SampEn on the performance of logistic map as a function of parameter, R. All three algorithms demonstrate an increase in chaos, with ApEn and SampEn providing a clearer delineation of the oscillation band.

2.2.2 Fuzzy Entropy

A further modification of SampEn, known as Fuzzy Fntropy (FuzzyEn), employs fuzzy functions as a replacement for the hard thresholding within SampEn in defining similar patterns [112]. With a continuous relation between similarity and obtained distance, FuzzyEn provides a more reliable and well-defined estimation, requiring a shorter data length compared to SampEn [112]. FuzzyEn has found application in the study of gaze complexity dynamics during virtual sailing and the research of brain dysfunction [113, 114]. Algorithm 17 outlines the steps of FuzzyEn, and Figure 2.10(a) & (b) present the results of FuzzyEn with different fuzzy functions, albeit inaccurately assigning lower levels to chaotic signals than periodic signals.

Algorithm 17. Fuzzy Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N, the parameters involved are the embedding dimension, m, tolerance, r_f, and time delay, l. The fuzzy function applied here is the Gaussian function with a chosen order, η.

1. Form the centered embedding matrix, $\mathbf{U}_m$, derived from the original signal, $\{x(i)\}_{i=1}^{N}$, by $u_m(i) = x_m(i) - \mu_m(i)$, where $\mu_m(i) = \frac{1}{m}\sum_{k=1}^{m} x_m(i+k)$.
2. Compute the distance between all the pairwise embedding vectors, $u_m(i)$ & $u_m(j)$, based on the Chebyshev distance, as $d_m(i,j) = max\{u_m(i+k) - u_m(j+k)|| 0 \leq k \leq m-1, i \neq j\}$.
3. Convert the distance matrix, $\mathbf{D}_m$, to the similarity matrix, $\mathbf{S}_m^{r_f}$, calculated through the Gaussian function as $s_m^{r_f}(i,j) = e^{(\frac{d_m(i,j)^\eta}{r_f})}$.
4. Compute the estimated local probability of $S_m^{r_f}(i)$ by $C_m^{r_f}(i) = \frac{\sum_{j=1,j\neq i}^{N-n-1} S_m^{r_f}(i,j)}{N-n-1}$, $n = (m-1)*l$.
5. Compute the estimated global probability of $S_m^{r_f}(i)$, by $\Phi_m^{r_f} = \frac{\sum_{i=1}^{N-n} C_m^{r_f}(i)}{N-n}$.
6. Repeat step 1–5 with an increased embedding dimension, $m+1$, and obtain the updated global probability as $\Phi_{m+1}^{r_f} = \frac{\sum_{i=1}^{N-n} C_{m+1}^{r_f}(i)}{N-n}$, with $n = m*l$.
7. Fuzzy Entropy is defined as

$$FuzzEn(m,l,r_f,N) = -\ln\left[\frac{\Phi_{m+1}^{r_f}}{\Phi_m^{r_f}}\right].$$

2.2.3 Bubble Entropy

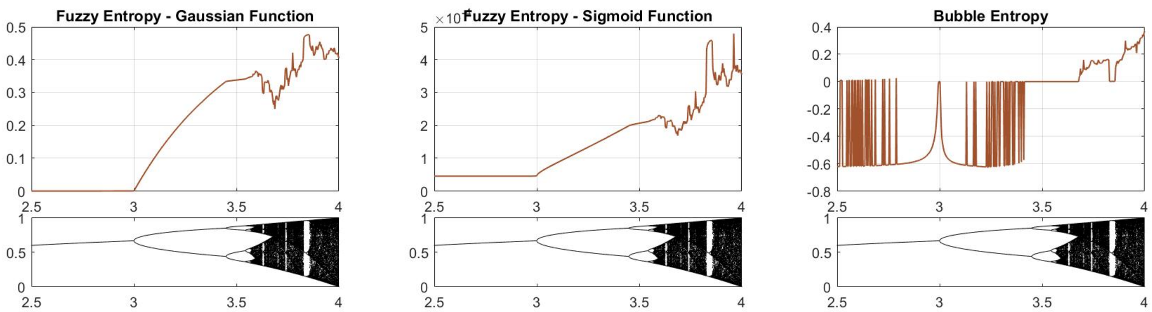
Considering the need for and the comparison of reconstruction in two-phase spaces (m & $m+1$), CE-based algorithms generally exhibit higher time complexity than SE-based methods. In response, Bubble Entropy, shown in Algorithm 18, was introduced with a reduced time load and fewer tunable parameters compared to traditional CE-based algorithms like SampEn and FuzzyEn [115]. Bubble Entropy (BubbEn) is developed based on phase reconstruction with a manually set embedding dimension, m. However, without distance calculation, regularity is quantified based on the randomness of embedding vectors, as detailed in Algorithm 18, and its performance is illustrated in Figure 2.10(c).

Algorithm 18. Bubble Entropy

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N, the parameters involved are the embedding dimension, m.

1. Construct the embedding matrix, $\mathbf{X}_m$, derived from the original signal, $\{x(i)\}_{i=1}^{N}$, in the form $x_m(i) = (x(i), x(i+1), \ldots, x(i+m-1))$.
2. For each embedding vector, $\mathbf{X}_m$, re-range the elements in $\mathbf{X}_m$ in ascending order and calculate the time of swaps during the organization, denoted as $s(i)$.
3. Calculate the relative frequency based on the swap time series, $s(i)$, as $E_m = -log(\sum P_t^2)$.
4. Repeat step 1–3 with an increased embedding dimension, $m+1$, and obtain the updated relative frequency, referred as E_{m+1}.
5. Bubble Entropy is defined as

$$BubbEn(m, N) = \frac{E_{m+1} - E_m}{log\frac{m+1}{m-1}}.$$



(**a**) FuzzyEn (Gaussian) (**b**) FuzzyEn (Sigmoid) (**c**) BubbEn

Figure 2.10: Behaviour of the FuzzyEn and BubbEn on the performance of logistic map as a function of the parameter, R. In this case, FuzzyEn erroneously indicates higher randomness during the oscillation band. Conversely, BubbEn demonstrates improved performance when the signal exhibits more structures.

2.3 Multiscale Entropy

Before the proposal of the concept of structural complexity and the dominance of Sample Entropy-based analysis in entropy-related research, studies revealed that single-scale

SampEn assigns a higher entropy level—interpreted as higher complexity—to healthy and young systems. This contradicts the loosely correlated and highly adaptive behaviours resulting from the intrinsic multiscale property of physiological signals, which single-scale analysis fails to illustrate [116]. It is well understood that complexity in biosignals is inherently multiscale, given that physical systems operate across various spatial and temporal scales [64].

The concept of 'scale' plays a unique and crucial role in Information Theory studies. To illustrate, an electroencephalogram (EEG) signal in a resting state and its shuffled version were plotted in the Time domain, Frequency domain, and Information Domain in Figure 2.11. The Time domain involves analysing signal changes over time. Since randomization shuffled the order of amplitudes, the behaviour of the randomized EEG resembles White Gaussian Noise with a zero mean. The Frequency domain studies signals changing in the frequency spectrum. As depicted in the middle panel, the main frequency of the original EEG is 13 Hz, with the second harmonic at 26 Hz and an artefact at 50 Hz, while the shuffled EEG displays a flat spectrum.

Aligned with the analyses in the Time domain and Frequency domain, the Information domain investigates signals changing in scale. Analysing under different scales provides varying perspectives or information levels. For instance, analysing an essay at the alphabet level lacks meaningful information about the content. Reading it by word offers a more feasible understanding, and studying it by sentence or paragraph aids in analysing the essay's structure.

The figure illustrates Sample Entropy-based information analysis, showing that the shuffled EEG demonstrates higher entropy levels with a greater degree of randomness at low scales, while the signal's structure fades away with scaling. This quantifies the richness of structure, with the randomized signal holding the lowest level. Regarding the EEG signal, its complexity maintains a certain level as the scale increases. The human brain is considered one of the most complex systems with high structural richness at all

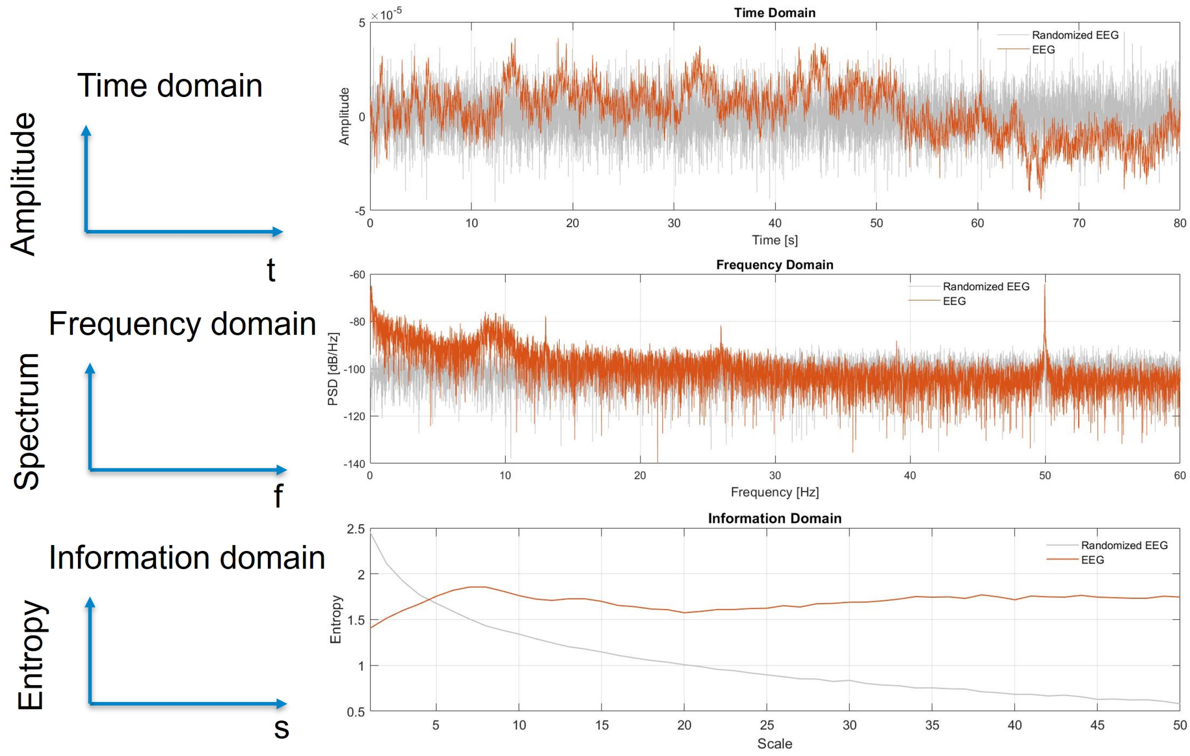


Figure 2.11: Comparison among time domain, frequency domain and information domain. The Coarse Graining Process (CGP) is applied to scale the EEG signal in the entropy calculation up to 50 with original data length $N = 10{,}000$. Details of CGP can be found in the following sub-section.

scales [62, 117, 118], as shown in the information domain.

2.3.1 Coarse-graining Process (CGP)

In the early stages of research on multiple scales, Zhang incorporated a sum of scale-independent entropies into complexity measures, albeit with limited physical applications [119]. In 2002, Costa *et al.* introduced the 'Coarse-Graining' Process (CGP) into Sample Entropy, naming it Multiscale Sample Entropy (MSE) [120]. This introduction marked the beginning of multiscale analysis playing an overwhelmingly vital role in subsequent research on entropy analysis. The CGP involves a non-overlapping average over the original signal based on the scale factor set, enabling the reflection of long-range correlations [120, 121].

In alignment with the concept of multiple-scale analysis, Costa and Goldberger extended this idea to propose Generalized Multiscale Entropy, which generalizes the CGP to a family of statistics [122, 123]. Instead of employing the mean as in CGP, Generalized Multiscale Entropy uses the variance of the embedding vector to better quantify the volatility of the signal. However, the generalized process using variance faces an inherent problem—the discrepancy of dimensions between the original and scaled signals, resulting in the distortion of the embedding vector due to quadratic behaviour. To address this, standard deviation was employed instead of variance in another entropy method [124]. To distinguish the three multiscale transformations, subscripts are added, denoted as MSE_μ, MSE_σ, and MSE_{σ^2}, referring to Multiscale Entropy based on the mean, standard deviation, and variance, respectively. The equations for MSE_μ, MSE_σ, and MSE_{σ^2} are provided below.

$$\mathrm{MSE}_\mu : {}^{\mu}y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq \frac{N}{\tau} \tag{2.14}$$

$$\mathrm{MSE}_\sigma : y_j^{(\tau)} = \sqrt{\frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} (x_i -{}^{\mu} y_j^{(\tau)})^2}, \quad 1 \leq j \leq \frac{N}{\tau} \tag{2.15}$$

$$\mathrm{MSE}_{\sigma^2} : y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} (x_i -{}^{\mu} y_j^{(\tau)})^2, \quad 1 \leq j \leq \frac{N}{\tau} \tag{2.16}$$

Despite the improvements in MSE_σ and MSE_{σ^2}, MSE_μ continues to play a critical role in the further development of entropy algorithms due to its high calculation efficiency and easy implementation. Therefore, in the subsequent abbreviations, MSE without an additional subscript refers to a multiscale process based on the mean by default (i.e., the traditional coarse-graining process). However, MSE has two general drawbacks. One is the lack of data length after non-overlapping averaging [125], and the second is the information loss of high-frequency components resulting from the down-sampling behaviour of the Coarse-Graining Process (CGP) [124]. Moreover, the averaging process is equivalent to a low-pass finite-impulse response (FIR) filter in the frequency domain. Applying CGP at a high scale might introduce aliasing [126].

2.3.2 Moving-averaging Method

To address the first problem of MSE, a modified Multiscale Entropy [127] was developed using the Moving Average method (MA-MSE). The moving-average process in Eq. (2.17) retains the same sample points as the original data, significantly reducing the occurrence of undefined and imprecise estimations, leading to a more reliable entropy analysis. Despite the improvement introduced by the moving-average process, the efficiency of MSE is limited. The high computational load of the moving-average method makes it challenging to apply in large sample sizes.

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=j-\tau/2-1}^{j+\tau/2-1} x_i, \quad 1 \leq j \leq N \tag{2.17}$$

2.3.3 Other Multiple Scale Process

In addition to the direct operation on the original data sets as in CGP, it has been demonstrated that decomposition methods also have the potential to contribute to the extraction of 'scale.' Among various decomposition algorithms, Empirical Mode Decomposition (EMD) [128] and Hierarchical Process (HP) [129] have been widely used in entropy analysis as substitutes for CGP. Both EMD and HP are analogous to wavelet analysis, aiming to decompose the original signal into a series of sub-signals through a linear combination model. The components resulting from EMD are named intrinsic mode functions (IMFs) with narrow-band scales. EMD operates on extrema with iterative subtraction of IMF from the original signal and has been applied to the analysis of nonlinear and nonstationary signals [130, 131]. Further details about EMD can be found in [129]. The multiscale data set by EMD is formed by separate or cumulative intrinsic mode functions (IMFs) in the decomposition. The improvement of EMD-MSE compared to MSE can be observed in the same frequency band as before 'scaled' [132].

Concerning the Hierarchical Process, it was recently introduced in Hierarchical Entropy [129]. The hierarchical components are generated into low-frequency elements and high-frequency elements through averaging and difference operators, separately. The scaled signal corresponds to a hierarchical component in each layer. Yet, as a new method, it has been integrated into various entropy algorithms such as Hierarchical Permutation Entropy [133], Hierarchical Symbol Dynamic Entropy [134], Hierarchical Fuzzy Entropy [135], etc.

The multiscale processes based on decomposition, such as EMD and HP, have the capability to alleviate the second problem of MSE, retaining information in both low and high frequencies. However, the drawback becomes apparent in the form of low computational efficiency, as observed in MA-MSE [124].

2.4 Multivariate Multiscale Entropy

With the evolution of hardware and advanced recording equipment, the use of multiple channels has become prevalent in the majority of research studies. Consequently, to fully leverage all variables and overcome the limitation of data length, multivariate entropy analysis has emerged as a significant topic, especially in physical and physiological systems. Various fusion methods have been proposed and investigated, aiming not only to reveal the structure within individual signals but also to explore the relations and correlations among channels.

Multivariate Multiscale Sample Entropy (MMSE), in Algorithm 19, was proposed as an extension to SampEn to simultaneously process information from multiple channels and alleviate the requirement for long data lengths in univariate entropy analysis [136]. Despite the certain expense of the loss of local information, in the context of multichannel data, MMSE offers a more robust and accurate estimation, even with limited data lengths [136].

Subsequently, Multivariate Multiscale Fuzzy Entropy (MMFE) was introduced, maintaining the merits of FuzzyEn in contrast to SampEn [137], in the multivariate setting. It's important to note that the scale of interest in multiscale entropy, based on amplitude

Algorithm 19. Multivariate Multiscale Sample Entropy

The steps of standard Multivariate Multiscale Sample Entropy are given below, for a multivariate data set $\{x_{c,i}\}_{i=1}^{N}$, $1 \leq c \leq P$ of length N and number of channel P. The manually selected parameters are the embedding dimension $M = [m_1, m_2, \ldots, m_P]$, tolerance r, time delay $L = [l_1, l_2, \ldots, l_P]$, and scale factor τ:

1. Normalize the original multivariate data sets by subtracting the mean and dividing by the standard deviation.
2. Perform Coarse Graining Process to obtain the scaled multi-channel time series $\{y_{c,i}^{(\tau)}\}_{j=1}^{N/\tau}$, according to $y_{c,i}^{(\tau)}(j) = \frac{1}{\tau}\sum_{i=j-\tau/2-1}^{j+\tau/2-1} x_c(i)$, $1 \leq j \leq \frac{N}{\tau}$, $c = 1, 2, \ldots, P$.
3. Form the Composite Delay Vectors (CDV) $\mathbf{Y}_M(i)$ according to M and L in the form

$$\begin{aligned}\mathbf{Y}_M(i) = [&y_{1,i}, y_{1,i+l_1}, \ldots, y_{1,i+(m_1-1)l_1}, \\ &y_{2,i}, y_{2,i+l_2}, \ldots, y_{2,i+(m_2-1)l_2}, \\ &\vdots \\ &y_{c,i}, y_{c,i+l_c}, \ldots, y_{c,i+(m_c-1)l_c},]\end{aligned}$$

4. Compute the similarity for all pairwise CDVs, $Y_M(i)$ & $Y_M(j)$, based on the Chebyshev distance as $D(i,j) = max\{Y_M(i+k) - Y_M(j+k)|| \, 0 \leq k \leq ((\sum_{c=1}^{P} m_c) - 1), i \neq j\}$.
5. Calculate the number of matching patterns, defined as similar pairs $B(i)$ that satisfy the criterion $D(i,j) \leq r$.
6. Compute the local probability, $C(i)$, and global probability, $\Phi(i)$, of $B(i)$ as $C(i) = \frac{B(i)}{N-n-1}$, $\Phi = \frac{\sum_{i=1}^{N-n} C(i)}{N-n}$, $n = max(M) * max(L)$.
7. Repeat Steps 3-6 with an increased embedding dimension, m_c+1, and obtain the updated global probability as $\Phi^* = \frac{\sum_{i=1}^{N-n} C^*(i)}{N-n}$, $n = max(M^*) * max(L)$. Recall that there are P ways to increase the embedding dimension and the modified global probability, Φ^*, is the averaged result.
8. Multivariate Multiscale Sample Entropy is defined as $MMSE = -\ln\left[\frac{\Phi^*}{\Phi}\right]$.

distance, lacks unified measure indicators as criteria when the scale is set to exhibit the long-range correlation of the signal. The introduction of MMSE has further catalyzed research on practical entropy algorithms. Existing multivariate entropies to date include:

- Multivariate Multiscale Sample Entropy (MMSE) [136]: A method that conducts joint multivariate analysis of physiological signals associated with multiple channels.
- Multivariate Multiscale Fuzzy Entropy (MMFE) [137]: This method combines composite delay vectors and fuzzy entropy [112], presenting smoother and more stable

estimates compared to MMSE.

- Multivariate Multiscale Permutation Entropy (MMPE) [37]: An extension of standard Permutation Entropy [28], inheriting desirable properties such as fast computation and simple implementation from PE.

- Multivariate Multiscale Distribution Entropy (MMDistEn) [138]: Recently introduced, this entropy method is developed on the basis of SE, incorporating Euclidean distance. It exhibits high calculation efficiency in quantifying the randomness of a system.

- Multivariate Multiscale Dispersion Entropy (MMDispEn) [139]: An extension of standard Dispersion Entropy [73] and an improvement of Permutation Entropy with different mapping techniques. It demonstrates more reliable and robust calculations.

- Variational Embedding Diversity Entropy (veMDE) [140]: Developed on the basis of Diversity Entropy [31], this method combines angular distance and relative probability, offering a low computational load with performance similar to MMPE.

Among the existing multivariate entropy algorithms listed above, the last four methods were constructed based on SE, providing the average uncertainty of a system. In contrast, the first two algorithms are founded on Conditional Entropy, quantifying the generation rate of new information. Both categories are established based on the amplitude of the original signal.

Chapter 3

Univariate tool: ClassA Entropy

3.1 Chapter Summary

Despite the recent theoretical boom in Sample Entropy based algorithms for the analysis of physiological and pathological systems, the major issue which prevents their more widespread use remains that of large computational load, particularly in the studies of quantification of structural richness in data. This issue becomes even more prohibitive when it comes to large data sizes and real-time processing. To this end, a new Classification Angle (ClassA) Entropy is introduced for the quantification of structural complexity of real world signals, based on an improved Second Order Difference Plot and Shannon Entropy. In comparison with existing nonlinear techniques, including Real Sum Angle index, Phase Entropy, Gridded Distribution Entropy and Cosine Similarity Entropy, the proposed method offers the advantages of a minimal number of tunable parameters, lower requirement of data size, wider application field and large relaxation of computational load. Simulations on real world physiological data support the approach.

3.2 Introduction

Physiological signals have been widely analysed in the context of nonlinear dynamics, and many established methods exist for examining their inherent non-linearity and complexity. Apart from time-domain methods, a number of studies, particularly in medical research, have shown the benefits of analysis of structural complexity in diagnosis of diseases and transition of pathological conditions [30, 47, 132, 141, 142]. However, despite its popularity and significance, the understanding of metrics of structural complexity in applications on physical systems has been under-explored, due to the practical limitations such as the requirement of long data length, large computational load and sub-optimal settings for the manually selected parameters. Among the existing methods for the estimation of structural complexity, entropy-related algorithms have occupied a major place, particularly in physiological and pathological analyses [15, 35, 111, 143–145].

Entropy in complexity science was first introduced by Shannon in 1948 as a way to quantify self-similarity of time series based on information theory [25]. In most existing entropy algorithms, the main procedures are built based on the mapping of the original signal onto an embedded phase space; algorithms in this class include Sample Entropy, Fuzzy Entropy and Cosine Similarity Entropy [15, 26, 112]. In a special case, when the embedding dimension is set as $m = 2$, the original signal is mapped onto a new phase space and can be presented as a 2D Poincaré plot of $x(i + 1)$ against $x(i)$ [97].

Poincaré plots, also called return maps, are a commonly used graphical tool for visualizing dynamical signals [97], and have been widely applied in characterizing the complexity features of physiological signals, such as heart rate variability (HRV) [97,146–151]. Although numerous entropy algorithms have been established based on the underlying embedding space, the utilization of entropy is mostly considered in conjunction with the distance among embedding vectors, which requires high computational load. Straightforward applications of Poincaré plots in entropy analysis are limited to a small number of studies; one popular algorithm which rests upon Poincaré plots is the Gridded Distribution Entropy (GDE), which describes entropy over the distribution of data on a gridded 2-D Poincaré plot [98], this allows for the investigation of the degree of variability in data, while limited information is provided by a Poincaré plot about the states of consecutive acceleration or deceleration of the signal dynamics [20]. To this end, Phase Entropy was proposed as a descriptor of the Second Order Difference Plot (SODP) and the angle distribution of data point in the SODP space [20]. Unlike the Poincaré plots which represent a scatter plot of a signal, $x(i)$, against its delayed version, $x(i-1)$, Phase Entropy employs the original SODP introduced by Cohen *et al.* [99], as $(x(i+2)-x(i+1))$ against $(x(i+1)-x(i))$. In other words, Phase Entropy investigates a scatter plot of two consecutive gradient estimates, and exhibits enhanced robustness compared to the standard Poincaré plot.

Recently, a method named the Classification Angle (ClassA) framework was introduced based on an improved SODP, using a three-point forward approximation of the derivative, that is, $[(4x(i+2)-3x(i+1)-x(i+3)))/2]$ against $[(4x(i+1)-3x(i)-x(i+2)))/2]$ [152]. The output of ClassA Framework is given as the graph of the proportion of points in quadrants of the improved SODP (I-SODP) descriptor and the Real Angle Sum (RAS). Therefore, limited improvement can be achieved by further numerical analysis. To this end, we introduce a novel algorithm based on the application of Shannon Entropy within the ClassA framework, termed ClassA Entropy.

The ClassA Framework has been proven to be particularly useful for HRV analysis [152]. We here address a more challenging problem of real-world Electroencephalogram (EEG) signals, to demonstrate that ClassA Entropy represents a general methodology to quantify the structural complexity for all dynamic systems. ClassA Entropy is also demonstrated to exhibit enhanced performance in complexity estimation compared to other graph descriptor-based entropy algorithms, such as Phase Entropy and Gridded Distribution Entropy. In contrast to the existing entropy algorithms, aiming at quantification of structural complexity such as Cosine Similarity Entropy [15], ClassA Entropy offers a tremendous relaxation on the computational load without the involvement of distance calculation, and with a minimal number of manually set parameters.

To establish the properties of ClassA Entropy, the EEG signals of 20 subjects from two pathological states were investigated. The performance of ClassA Entropy is examined over a range of parameter settings and data lengths, and benchmarked against the existing ClassA Framework, Gridded Distribution Entropy, Phase Entropy and Cosine Similarity Entropy algorithms.

3.3 Graphical Descriptors of structural complexity

EEG data sets from 10 coma subjects and 10 brain death subjects were utilised in the analysis, with maximal segment length of $N = 30{,}000$. The data were collected using a standardized 10-20 system via a portable EEG recording instrument with a sampling frequency of 1000 Hz in the intensive care unit (ICU) in Shanghai Huashan Hospital affiliated to Fudan University China, and was extensively analysed for more than a decade [143, 153, 154]. Examples of the phase plots of EEG signals of the coma and brain death patients are shown in Fig. 3.1. For the ease of understanding, the left column represents the time series from one of the coma subjects and the right column the signal recorded from one brain death subject. The same data sets were simultaneously reconstructed by the traditional Poincaré plot and improved Second Order Difference Plot (I-SODP), and shown respectively in the upper row and lower rows in Fig. 3.1.

In Poincaré plots, the data points above the diagonal line refer to an accelerating state in a time series, while the points below the diagonal line designate a decelerating state of the signal. Observe that in the first row in Fig. 3.1, both EEG data sets mostly cluster around the diagonal line, without significant difference between the two conditions. Also the I-SODP, in the bottom row, exhibits the data dynamics in a different view. Unlike the Poincaré plot which uses raw data, it employs the scatter plot of the first derivative of the signal, whereby data points in the first quadrant represent deceleration of the signal

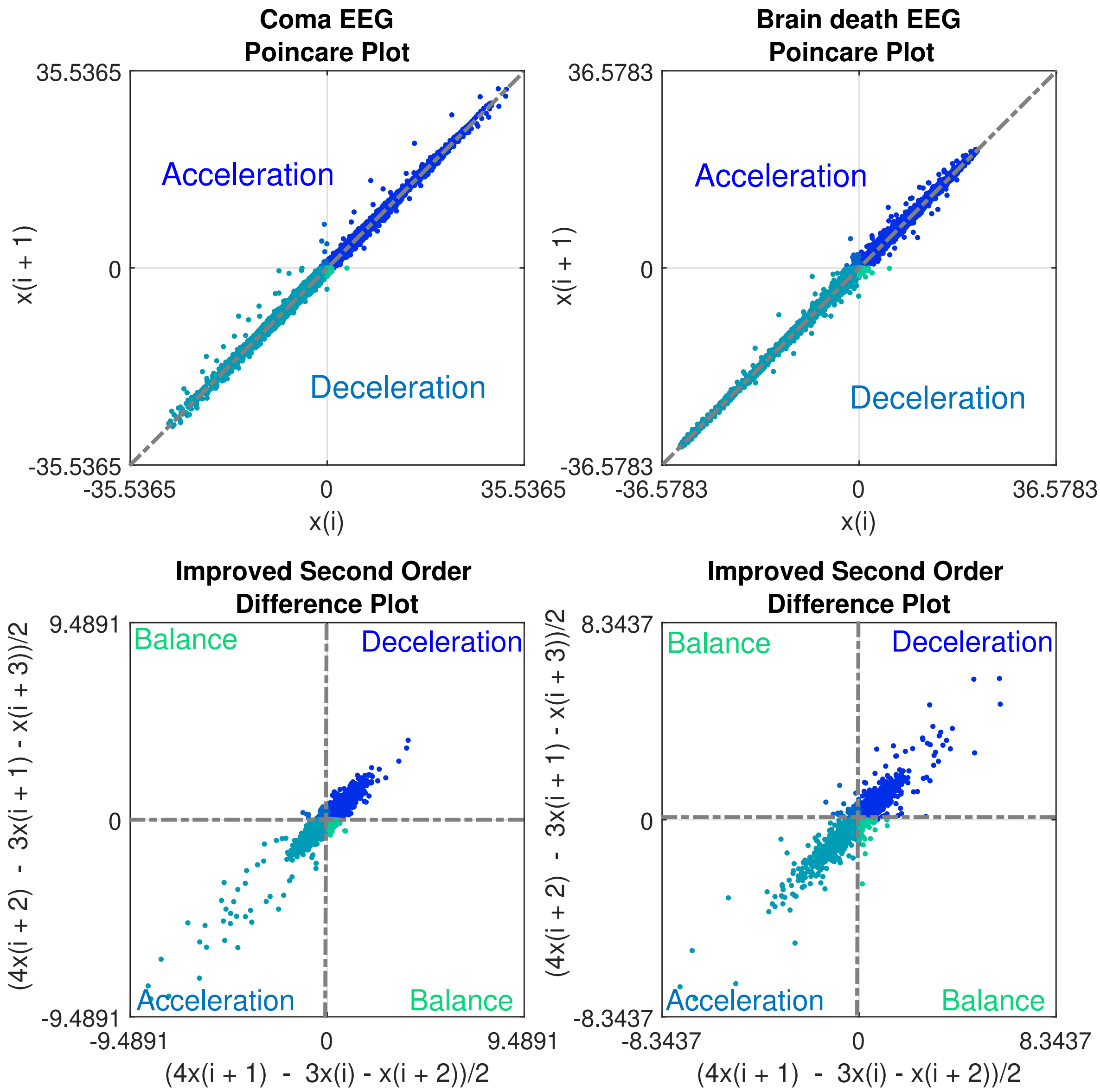


Figure 3.1: Standard Poincaré plots (upper panels) and I-SODP phase spaces (lower panels) of exemplar EEG signals in the states of coma (left panels) and brain death (right panels). Observe the much improved discrimination between the coma and brain death states in the improved SODP plots in the bottom panels.

and those in the third quadrant correspond to the acceleration [152]. The points in the second and fourth quadrants represent the balanced states of a time series. As shown in the lower panels in Fig. 3.1, the coma signal on the bottom left graph suggests more acceleration states in the third quadrant, while the brain death signal is considered more balanced. Therefore, the I-SODP method is observed to be an information-richer and more descriptive measure.

3.4 ClassA Entropy

Prior to introducing the properties of Classification Angle (ClassA) Entropy, details of Classification Angle (ClassA) Framework are given in Algorithm 20. The output of the ClassA Framework are the Real Angle Sum (RAS) and a three-dimensional plot of P^{Q24} (the proportion of data points in the second and fourth quadrant), P^{Q3} (the proportion of data points in the third quadrant) and P^{Q1} (the proportion of data points in the first quadrant). The cornerstone of ClassA Framework is the discrepant physical meanings of a time series in each quadrant in I-SODP space, as discussed in [152].

Algorithm 20. ClassA Framework

Given a data set $\{x(i)\}_{i=1}^{N}$ of length, N.

1. Map the original data into improved Second Order Difference Plot as $[(4x(i+2)-3x(i+1)-x(i+3)))/2]$ against $[(4x(i+1)-3x(i)-x(i+2)))/2]$.
2. Compute the angle of the data points and the abscissa in the I-SODP phase space, $\alpha(i)$.
3. Compute the Real Angle Sum (RAS) based on $\alpha(i)$, denoted as $RAS = \frac{\sum \alpha}{N}$.
4. Count the proportion of the points in the first and third quadrants separately, denoted by P^{Q1} and P^{Q3}.
5. Repeat Step 4 to give the proportion of the points in the second and fourth quadrants jointly, denoted by P^{Q24}.
6. The ClassA metrics are defined as $[RAS, P^{Q1}, P^{Q24}, P^{Q3}]$.

Inspired by the ClassA Framework, ClassA Entropy is proposed by building upon the I-SODP space, to give the probability distribution of the converted angle time series, $\alpha_n(i)$, by Shannon Entropy, with the detailed steps given in Algorithm 21. Note that ClassA Entropy is not restricted to the analysis of data distribution for each quadrant, but the main ideas of the algorithm are the utilization of the I-SODP space and the classification methodology. In Algorithm 21, the simplest linear grouping, through equal segmentation of the space, is applied. Before grouping the data, mapping functions can be employed to introduce the weighting on the original time series, such as the typical Normalized Cumulative Distribution function (NCDF). In addition, a clustering method could be implemented by other linear or non-linear classification tools, such as the popular K-means. For simplicity, we only showed the ClassA Entropy with equally spaced grouping strategy.

Algorithm 21. ClassA Entropy

Input: A data set $\{x(i)\}_{i=1}^{N}$ of length, N, and the number of classes, K.

1. Map the data onto an improved Second Order Difference Plot by plotting $[(4x(i+2) - 3x(i+1) - x(i+3)))/2]$ against $[(4x(i+1) - 3x(i) - x(i+2)))/2]$.
2. Compute the angle of data points and the abscissa in the I-SODP space, $\alpha(i)$.
3. (optional) Map the angle data, $\alpha(i)$, by the Normalized Cumulative Distribution function (NCDF) or other mapping functions (e.g., fuzzy functions).
4. Group the angle time series, $\alpha(i)$, into equally spaced sections and count the number of data points in each sector. Note that this step can be replaced by other linear or non-linear classification methods (e.g., K-means).
5. Calculate the Classification Angle (ClassA) Entropy according to the probability distribution of the data points in the I-SODP space, as $-\sum_{t=1}^{K} p(t) * log(p(t))$.

3.4.1 Selection of Parameter, K

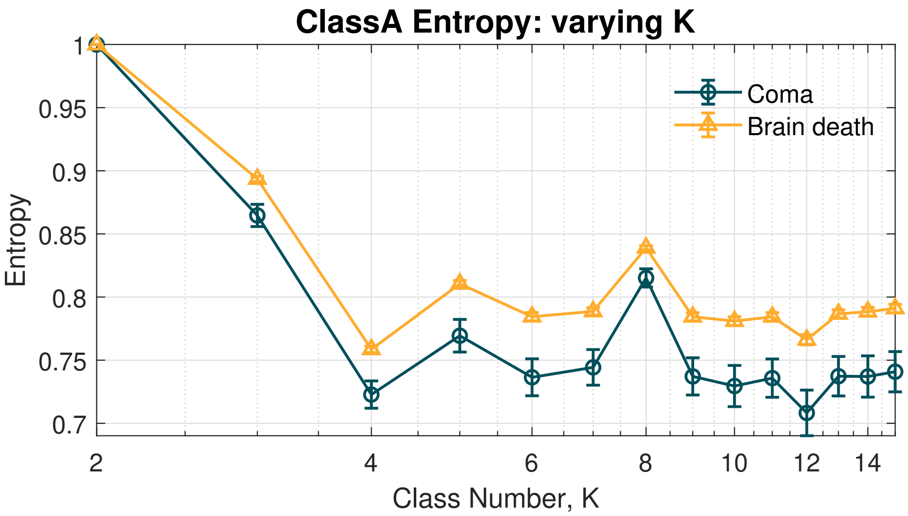


Figure 3.2: ClassA Entropy for the complexity estimation of EEG signals, as a function of the class number, K. For each curve, the error-bars designate the average and standard error over 10 subjects.

As shown in Algorithm. 21, the only parameter that needs to be manually set is the number of classes to be grouped together. To investigate the influence of the values of K, the performance of ClassA Entropy as a function of group number is shown in Fig. 3.2.

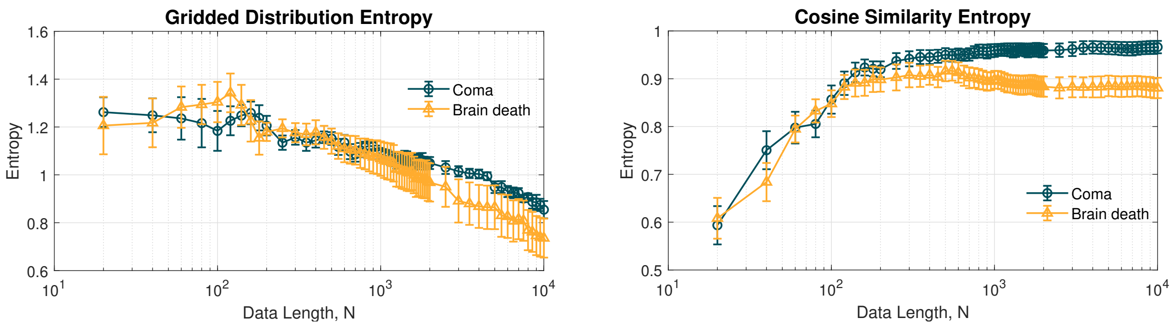
By analysing the EEG signals recorded from coma and brain death patients with the data length fixed as $N = 5000$, the class number, K, was set to vary from 2 to 16. Observe that, when the number of classes was larger than 2, it yielded no significant impact on the ability of ClassA Entropy to separate the two pathological conditions, without overlapping areas between them. When K was larger than 3, the ClassA Entropy has reached a stable level, with only minor fluctuations as K increases. To this end, K is generally recommended to be set as 4, as minimal class number.

Remark 1: The only tunable parameter in ClassA Entropy is the cluster number, K, which is suggested to be set as $K = 4$.

3.4.2 The ClassA Entropy versus existing entropy algorithms

To explore the advantages of the proposed ClassA Entropy, the performance of the existing entropy algorithms were first analysed over EEG signals of 10 subjects. Fig. 3.3(a) and Fig. 3.3(b) illustrate the complexity estimation measured by Gridded Distribution Entropy (GDE) and the standard errors of Cosine Similarity Entropy (CSE). The performance of RAS given by the ClassA Framework and Phase Entropy (PhaseEn) as a function of data length are shown in Fig. 3.4(a) and Fig. 3.4(b). Entropy-based complexity estimation on physiological and physical systems can mainly be categorised into two groups based on the measured properties: *i*) quantification of randomness or irregularity of signals according to the Complexity Loss Theory [45]; *ii*) the description of long-term correlation within signals based on a new theory [15]. Given by these two theories and prior literature, the brain death EEG signals are considered to have a lower degree of randomness and higher long-term correlation [15, 45].

The independent variable, data length N, ranging from 20 to 30,000 was applied in all the graphs, where the x-axis of the plots was plotted in a logarithmic scale. Observe that the RAS and PhaseEn demonstrated a degree of correlation, whereby brain death curves were higher than those for coma signals. While GDE and (single-scale) CSE quantify the randomness of the data that coma signals presented higher irregularity than brain death signals owing to the single-scale (original data). Among the four methods, PhaseEn and CSE in Fig. 3.4 imply that shorter data lengths ($N = 500$ and $N = 700$ respectively) were needed to separate the two pathological conditions and yeild stable estimations owing to the employment of SODP phase space and angular distance, respectively. In contrast, the RAS index required data size of over $N = 2000$ to correctly estimate the complexity without the overlapping area, while GDE was inadequate for the current data set to characterize the complexity property since the magnitude diversity of mapped data on the phase space is less significant than the angle distribution.



(**a**) Gridded Distribution Entropy (GDE) (**b**) Single-scale Cosine Similarity Entropy

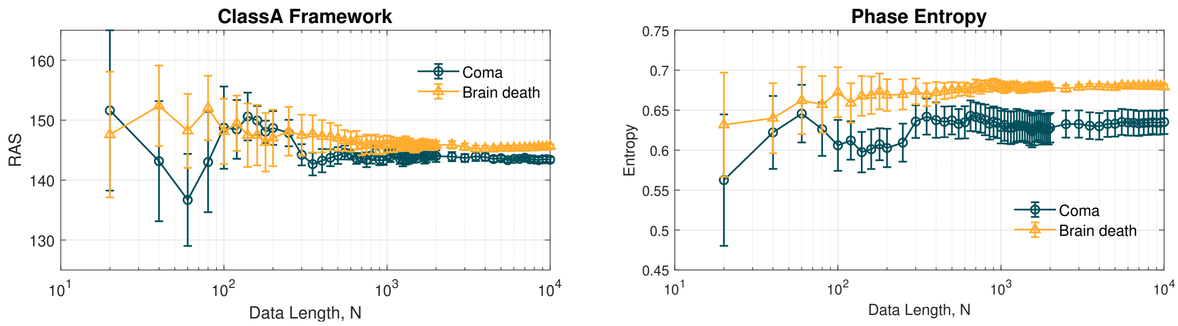
Figure 3.3: Complexity estimation of EEG signals, as a function of the data length, N. The grid setting for GDE is [3×3] and the embedding dimension for CSE is set to $m = 3$.

With the class number, $K = 4$, and the simplest (equally spaced) grouping method, the proposed ClassA Entropy as a function of varying data length is next examined in Fig. 3.5. Observe that, in line with RAS, the ClassA Entropy based on I-SODP indicates the degree of correlation of the signal, with brain death signals having a higher degree of correlation. Note that there was a significant relaxation on the requirement of data length, with $N = 100$ being sufficient to separate the two states in a clear and stable manner.

Remark 2: Less data length is required for the proposed ClassA Entropy without the sacrifice of the estimation stability.

Remark 3: Given the success of the analysis based on EEG data sets, the application of

ClassA Entropy can be extended to all dynamical systems, and not only limited to HRV analysis, as introduced in the ClassA Framework.



(**a**) Real Angle Sum (RAS) (**b**) Phase Entropy

Figure 3.4: Structural complexity analyses of EEG signals, as a function of the data length, N. The group number for both ClassA Framework and PhasEn is set to 4.

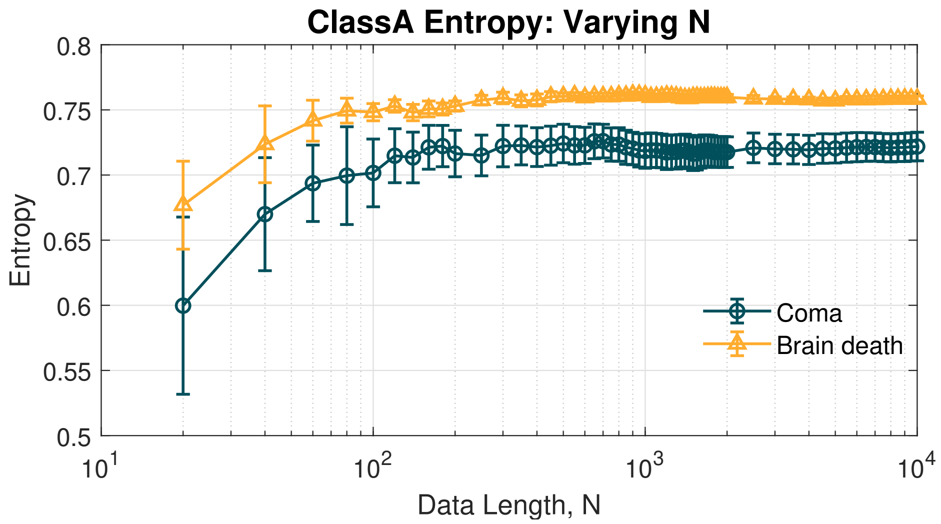


Figure 3.5: Behaviour of ClassA Entropy on the estimation of EEG signals, as a function of the data length, N. The error-bars designate the average value and standard error over 10 subjects in each group.

3.4.3 Time complexity

Considering the considerable potential for practical use of entropy-based algorithms in real-world data processing, the computation load is the factor that needs to be examined and balanced. To this end, we next examined the time complexity of the four existing methods, along with the proposed ClassA Entropy. This was achieved by averaging processing time in MATLAB, conducted on a laptop with a CPU base frequency of

1.60GHz, over 10 realizations of White Gaussian Noise (WGN) with the same data size settings.

As shown in Fig. 3.6, compared to the widely used correlation estimation methods, ClassA Entropy exhibited a significant advantage over CSE in terms of the computational efficiency, owing to no requirement for distance calculation. The right panel in Fig. 3.6 shows the elapsed time for the calculation of PhaseEn and GDE, along with the ClassA Entropy. As discussed above, the three entropy algorithms were built on the different graphical descriptors and aimed to measure complexity over the distribution probability of the data points in the reconstructed phase space. Therefore, considering similar principles and simple mathematical processes of the three algorithms, the computational loads of all the three entropy algorithms were low. The GDE exhibited the least processing time, while ClassA Entropy required slightly higher computational load, but of the same order of magnitude.

Remark 4: The computational efficiency of the ClassA Entropy is significantly higher than the traditional CSE algorithm for structural complexity measurement.

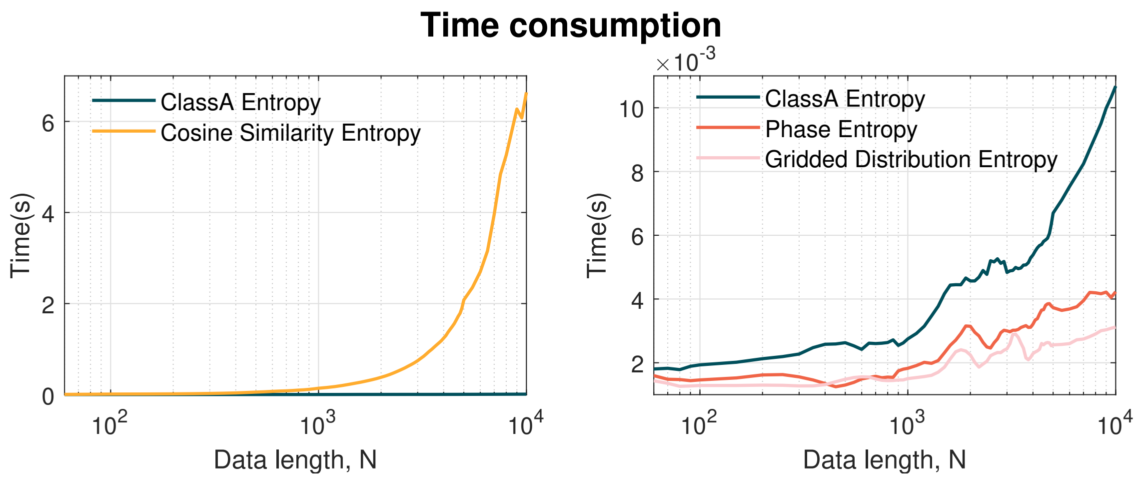


Figure 3.6: Computation time for the existing CSE, PhaseEn, GDE and the proposed ClassA Entropy algorithms, as a function of data length for 10 realizations of White Gaussian Noise (WGN) inputs.

3.5 Conclusion

We have proposed the ClassA Entropy framework for the quantification of structural complexity of real-world data. This has been achieved based on improved Second Order Difference Plot phase reconstruction. The performance of the new ClassA Entropy has been compared against the existing non-linear methods including RAS index, Phase Entropy, Gridded Distribution Entropy and Cosine Similarity Entropy in applications on EEG signals. As the most efficient measurement associated to long-term correlation, Cosine Similarity Entropy has failed to give the correct rank for single-scale estimation, while the proposed ClassA Entropy achieved perfect separation. By virtue of employing the improved Second Order Difference Plot, the ClassA Entropy has been demonstrated to provide a more stable quantification than the Poincaré plot-based (Gridded Distribution Entropy) and SODP-based (Phase Entropy) entropy methodologies. With the utilization of the same graphical representation, the ClassA Entropy has largely relaxed the requirement of data length in contrast to the RAS index given by the ClassA Framework.

The proposed method represents an innovative alternative to the quantification of structural complexity and has outperformed the existing dynamical methodologies on the analysis of an EEG data set. Future research will involve different dynamical systems, and other clustering methods and mapping functions, as extensions of the basic ClassA Entropy.

Chapter 4

Variational Embedding Multiscale Sample Entropy

4.1 Chapter Summary

Entropy-based methods have received considerable attention in the quantification of structural complexity of real-world systems. Among numerous empirical entropy algorithms, conditional entropy-based methods such as sample entropy, which are associated with amplitude distance calculation, are quite intuitive to interpret but require excessive data lengths for meaningful evaluation at large scales. To address this issue, we propose the variational embedding Multiscale Sample Entropy (veMSE) method and conclusively demonstrate its ability to operate robustly, even with several times shorter data than the existing conditional entropy-based methods. The analysis reveals that veMSE also exhibits other desirable properties, such as the robustness to the variation in embedding dimension and noise resilience. For rigor, unlike the existing multivariate methods, the proposed veMSE assigns a different embedding dimension to every data channel, which makes its operation independent of channel permutation. The veMSE is tested on both simulated and real world signals, and its performance is evaluated against the existing

Multivariate Multiscale Sample Entropy methods. The proposed veMSE is also shown to exhibit computational advantages over the existing amplitude distance-based entropy methods.

4.2 Introduction

We are witnessing a surge in the investigation of structural complexity of real world signals, as complexity science is now recognized to have the same importance as the properties in the time and frequency domains. Indeed, the structural complexity of a data set is a unique feature that can be utilized as a feature to understand subtle changes in the signal generating mechanism via nonlinear analytical tools or through machine learning [1]. Studies employing structural complexity as a feature have covered a wide spectrum, from fault diagnosis of rotating machines [31, 130, 155] to the early detection of disease and sickness in humans [18, 37, 38, 131]. It is important to note that bio-signals tend to exhibit high degrees of irregularities and complex dynamical behaviours [27], resulting from interactions between the human body (organisms) and peripheral environment, together with continuous fluctuations in time [22]. The complexity loss theory (CLT) has established the potential relationship between the complexity of physical signals and health of an individual, whereby the higher degree of complexity indicates a healthier condition of the individual [49]. However, new developments have declared that pathology may also manifest itself through an increase in complexity, based on the underlying signal structure; that is, a decrease of self-correlated complexity will also be observed in a healthy body [15].

Although the definition of structural complexity is inconsistent in the literature [13], there are several commonly used methods for the quantification of the "degree of dynamics", with entropy-based methodologies being the most popular ones. Compared to other methods, the estimation of complexity of nonlinear systems, through the fractal dimension [27], recurrence plots [22], and entropy analyses holds the advantage of simplicity and noise robustness [156]. More importantly, complexity based methods do not suffer from any restrictions related to the probability distribution [84]. The features of the loss of complexity (LOC) manifest themselves through, for example, an increase in randomness, reduction in regularity, breakdown of long-term correlations, multiscale variability, and time irreversibility [13]. To this end, a large number of entropy-based algorithms have been proposed to quantify the different facets of complexity or, more precisely, the degree of complexity based on different definitions.

Among numerous existing entropy algorithms, Shannon Entropy (SE) and Conditional Entropy (CE) are the two fundamental methodologies that quantify the amount of information and rate of information generation, respectively [92]. Typical Shannon Entropy-based methods that have been commonly implemented in practical scenarios are Permutation Entropy [28] and the recently introduced Dispersion Entropy [73], while the two early widely used Conditional Entropy-based algorithms are the Approximate Entropy (ApEn) [105] and Sample Entropy (SampEn) [26], proposed, respectively, in 1991 and 2000. A modification of ApEn, the SampEn has been shown to reduce the bias experienced by ApEn by removing self-matching delay vectors; SampEn also exhibits less dependency on the data length, thus giving relatively higher consistency [26]. Both ApEn and SampEn were developed to quantify the randomness and irregularity of a signal generating system. Generally speaking, the lower the value of SampEn, the less complex the system.

However, truly complex signals exhibit varying structures, across multiple time scales, while long-range correlations fail to be observed by single-scale Sample Entropy analysis. To this end, Costa *et al.* introduced a 'coarse-graining' procedure into the Sample Entropy methodology, to verify the structural complexity hidden in at high scales, referred to as the Multiscale Sample Entropy (MSE) [120]. Despite its broad use, the down-sampling procedure, given by the coarse-graining process, will include artefact components in high frequencies thus generating biased scaled signals [157]. As an improvement of the multiscale entropy strategy, refined multiscale entropy was proposed, which uses a low-pass

Butterworth filter to generate scaled signals [157]. Although the drawbacks of the coarse-graining process make it impossible to behave as an optimal filter, its simplicity and fast implementation make it valuable in the development of entropy-based applications. This, in turn, further spurred the development of MSE algorithms, including composite Multiscale Sample Entropy [158] and refined composite Multiscale Sample Entropy [159].

However, due to the 'coarse-graining' procedure, the requirement of long data length remains even more pronounced and is hard to satisfy in most practical situations. In 2011, Multivariate Multiscale Sample Entropy (MMSE) was introduced, which successfully combines data from multiple channels, to estimate the dynamics of the system more accurately and with shorter data lengths [136]. The key improvement of MMSE is the form of composite delay vector, which involves and reconstructs data segments from multiple channels, whereby the inner correlations among diverse signals are preserved [136]. The introduction of MMSE further spurred the research on practical entropy algorithms. The existing multivariate entropies to date include:

- Multivariate Multiscale Sample Entropy (MMSE) [136], a method which performs joint multivariate analysis of physiological signals associated with multiple channels.
- Multivariate Multiscale Fuzzy Entropy (MMFE) [137], which combines composite delay vectors and Fuzzy Entropy [112] and exhibits smoother and more stable estimates than MMSE.
- Multivariate Multiscale Permutation Entropy (MMPE) [37], an extension of standard permutation Entropy [28], which inherits the desirable properties of PE, such as fast computation and simple implementation.
- Multivariate Multiscale Distribution Entropy (MMDistEn) [138], a recently introduced entropy method, developed on the basis of Shannon Entropy with the inclusion of Euclidean distance, which exhibits high calculation efficiency in the quantification of the randomness of system.
- Multivariate Multiscale Dispersion Entropy (MMDispEn) [139], an extension of standard Dispersion Entropy [73] which is an improvement of Permutation Entropy with different mapping techniques, which exhibits a more reliable and robust calculation.

- Variational embedding Diversity Entropy (veMDE) [140], a method developed on the basis of Diversity Entropy [31] that combines angular distance and relative probability and exhibits a low computational load, with similar performance to MMPE.

Among the existing multivariate entropy algorithms listed above, the last four methods were built based on Shannon Entropy, which gives the average uncertainly of a system, while the first two algorithms are based on Conditional Entropy, which quantifies the generation rate of new information; both were established based on the amplitude of the original signal. Despite success, the inherent shortcomings of amplitude-based CE-developed entropy calculations still remain a major obstacle towards their more widespread use. Other issues with current CE-based multivariate entropy methods include:

1. The rule of thumb is that the requirement for data length is around 10^m to 30^m, where m refers to the embedding dimension [110] for CE-based methods, such as Approximate Entropy [105], Sample Entropy [26], and Fuzzy Entropy [112]. Hence, the choice of the embedding dimension is limited by the available sample size.
2. The 'coarse-graining' process further emphasizes the drawback of limited data size, which causes inaccurate and undefined estimation for high scale analysis.
3. Amplitude-based distance between delay vectors is sensitive to outliers, such as noise and artefacts.
4. Poor quality of any single channel has a large impact on the performance in a multivariate setting.
5. Excessive computational burden is required when implementing multi-channel analyses based on CE-developed and amplitude distance calculations.

Recently, Wang *et al.* [140] introduced a new way to combine data sets from multiple channels into one entropy estimation algorithm, termed variational embedding Multiscale Diversity Entropy. Nonuniform embedding space strategy has existed for a long time in the computation of complexity from different perspectives, such as the estimation of nonlinear causality by corrected Conditional Entropy [10, 104]. When it comes to the multivariate case, the key question is the optimization of the embedding dimension [160].

There is no general answer to this problem, since the optimal embedding dimension for a multi-channel system cannot be unique in practical scenarios [161]. Porta *et al.* have provided a comparison and stated that model-free approaches are less efficient when applied to nonlinear systems under a high embedding dimension; for higher scales, these also exhibit lower reliability [32], due to the lack of available data [160]. Faes *et al.* further introduced a non-uniform approach to detect the nonlinear Granger causality in a multivariate time series by adopting a step-by-step composition of embedding vectors to reduce the Conditional Entropy [104]. In the method proposed by Wang *et al.* [140], a simple and intuitive strategy was applied to give the complexity estimation, in terms of irregularity, by constructing the phase space with different structures to generate unique probability distribution for each channel. Here, inspired by [140], based on Sample Entropy, a new multivariate entropy method is proposed, and is referred to as the variational embedding Multiscale Sample Entropy (veMSE). This new method offers the following advantages over the existing Multivariate Multiscale Sample Entropy (MMSE) algorithm:

1. Complexity estimates at a higher embedding dimension are better defined, even with a limited data size.
2. The requirement for the number of data points is lower than in current Sample Entropy based methods.
3. Strong noise-robustness is exhibited across all scales.
4. The overall performance of a multivariate estimate is independent on the quality of any single-channel within a multivariate data set.
5. Less computational time is required, owing to a straightforward and efficient implementation.

The aim of this chapter is to propose variational embedding Multiscale Sample Entropy (veMSE), a method which combines the different multi-source fusion methods, multivariate strategy, and variational embedding strategy, applied on Sample Entropy, as well as to demonstrate the merits of the new proposed veMSE. The remainder of the chapter is organized as follows. In Section 4.3, the proposed veMSE algorithm is outlined, and the key improvement of the variational embedding strategy is discussed in detail. Section 4.4 demonstrates the operation of veMSE on simulated signals, to give an initial insight, with regards to the choice of parameters. Then, based on the suggested parameter setting in Section 4.4, Section 4.5 considers and discusses the properties of veMSE, including noise robustness, directionality, and calculation efficiency. Next, veMSE is applied to real-world signals, such as wind and heart rate variability (in Section 4.6), and compared with the performance of the univariate MSE and MMSE. Finally, conclusions summarise the work in this article.

4.3 Variational Embedding Multiscale Sample Entropy

The steps of the proposed veMSE are given in Algorithm 22. The key advantages of veMSE are:

- It allows for variations in the setting of the embedding dimension for multi-channel signals, as shown in Table 4.1, while simultaneously maintaining the information within each channel. Compared to MMSE (see Algorithm 19 in Chapter 2), despite the fact that different embedding values can also be achieved for each channel by MMSE, veMSE focuses more on the information across data channels. Indeed, the composite delay vector in MMSE successfully combines the embedding vectors of the individual channels within multichannel signals. However, there is a discrepancy and bias between the similarity among the recombined composite delay vectors and those among embedding vectors of the original data sets. The proposed veMSE estimates the complexity information of signals in multiple channels, without involving the cross-information between data channels.
- Due to the varying embedding dimensions, veMSE employs a weighted contribution of multiple channels; that is, the probability distribution of each channel is diverse,

which serves to reinforce chaotic features in measured signals of a multi-channel system.

- Since the probability of occurrences of similar patterns with veMSE is processed multiple times, based on the number of channels, and the summation process is performed before the logarithm operation in the last step, the veMSE is theoretically able to unveil the complexity properties under higher embedding dimension, but with the same data length, as compared with existing algorithms based on Sample Entropy.

To provide insight into the performance of the proposed veMSE, both synthetic signals and real world data sets are considered in the following sections, while the popular MMSE multichannel entropy algorithm is employed as a benchmark for veMSE, under the same conditions.

Table 4.1: Relation between the embedding dimension (m) and index of channel (c)

Index of channel (c)	1	2	...	c
Embedding dimension ($m(c)$)	m	$m+1$	...	$m+c+1$

4.4 Results of veMSE on Simulated Signals

In this section, synthetic signals, generated based on five benchmark models, are utilized to illustrate the performance of veMSE. These include white Gaussian noise (WGN), flicker noise (coloured noise), and autoregressive (AR) models AR(1), AR(2), and AR(3). Standard deviations for all the generated signals were set to $\sigma = 1$, while the coefficients of AR models are given in Table 4.2. The role of the parameters discussed in the following subsections include the embedding dimension, data length, tolerance, number of channel, and scale factor. Essentially, the temporal span is jointly controlled by the embedding dimension and time delay. Here, to avoid unknown influence of control variables, time

Algorithm 22. Variational Embedding Multiscale Sample Entropy

Assume that there are P channels measured from a system, where a signal, recorded from the c^{th} channel, is denoted by $x_c(i)$ and is of length N, where $1 \leq c \leq P, 1 \leq i \leq N$. Parameters involved in the veMSE algorithm are the tolerance quotient (r), embedding dimension (m), scale factor (τ), and time lag (L). The detailed steps of veMSE are shown below.

1. Coarse graining is firstly applied to the original data sets for all the channels, whereby the scaled time series are calculated as $y^{(\tau)}(j) = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x(i),\ 1 \leq j \leq N_t,\ N_t = \frac{N}{\tau}$.
2. For each channel, the embedding dimension, m is set as a variable. The dimension for the c^{th} channel is calculated as $m(c) = m + c - 1$, as listed in Table 4.1. Therefore, combined with the time delay, L, the embedding delay vector of data $y^{(\tau)}(i)$ $(1 \leq i \leq N_t)$ for channel c is designated as a template, $Y_c^{(\tau)}(i)$, and calculated as $Y_c^{(\tau)}(i) = [y^{(\tau)}(i), y^{(\tau)}(i+L)\ , \ldots ,\ y^{(\tau)}(i{+}n_c)],\ n_c = (m(c) - 1)L$.
3. Compute the Chebyshev distance between templates, $Y_c^{(\tau)}$, where the distance is denoted according to the amplitude of the embedding vector, as $d = max\{Y_c^{(\tau)}(i{+}k)) - Y_c^{(\tau)}(j + k)\}$, where $0 \leq k \leq m(c) - 1,\ 1 \leq i, j \leq N_t - n_c,\ i \neq j$.
4. For each channel, the number of segments, $Y_c^{(\tau)}(j)$, within the tolerance level, r, of $Y_c^{(\tau)}(i)$, is recorded as $B_c(i)$. In other words, $B_c(i)$ is the number of templates matched or the number of similar patterns in the data set, based on the boundary set, where the boundary is defined by the tolerance, r. Therefore, the local probability of the occurrence of template match for channel c is $R_c(i) = \frac{B_c(i)}{(N_t - n_c - 1)}$.
5. The global probability of the occurrence of template match for a channel c is calculated as $\Phi(c) = \frac{1}{(N_t - n_c)}\sum_{i=1}^{N_t - n_c} R_c(i)$.
6. For all data channels, compute the sum of the global probability as $\Phi = \sum_{c=1}^{P} \Phi(c)$. Recall that, unlike the MMSE algorithm (see Algorithm 19), the embedding dimension, $m(c)$, is not a fixed parameter and is varied, along with the index of channel, c.
7. Modify the embedding dimension to $(m(c){+}1)$. Hence, n_c is adjusted to $m(c) \times L$, and Steps 2-6 are repeated to obtain the global probability for the so-increased embedding dimension $\Phi(m + 1)$.
8. Variational embedding Multiscale Sample Entropy is finally obtained, as $veMSE = -ln\frac{\Phi(m+1)}{\Phi(m)}$.

lags were set to $L = 1$ for all operations, to make the temporal span fully defined by the modification of the embedding dimension.

Figures in each subsection are presented in pairs, to illustrate the structural complexity results for the five benchmark models, and compared with the standard MMSE algorithm. Upper panels give the curves for white and flicker noise, while the panels in the bottom depict the results of the AR models, in contrast to white noise. Complexity curves of entropy values are plotted as error bars, based on standard deviation, averaged over outcomes of 20 independent realizations for each model.

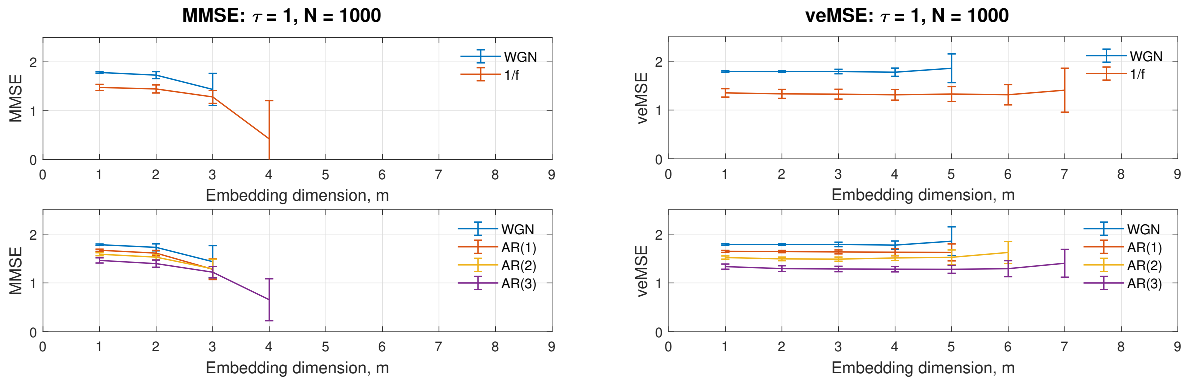
Table 4.2: Coefficients of AR models

Coefficients	a_1	a_2	a_3
AR (1)	0.5	–	–
AR (2)	0.5	0.25	–
AR (3)	0.5	0.25	0.125

4.4.1 Varied Embedding Dimension (m)

Usually, for the implementation of SampEn-based algorithms, the embedding dimension and data length are the two parameters that are interdependent and mutually coupled, with the data size restricted to between $(10^m - 30^m)$, as a rule of thumb [162]. In the real world, the recorded signals do not have infinite length and are generally limited by operation time and memory space. Therefore, the embedding dimension is commonly set to quite low values of $m = 2$ or $m = 3$ for a signal with 1000 samples [125]. Higher values of the embedding dimension for a small data size will cause unstable estimation, as in standard MMSE, shown in Figure 4.1(a).

The corresponding results for the veMSE, as a function of embedding dimension, are shown in Figure 4.1(b). Each entropy estimate was calculated based on signals from two channels. Except the independent variable, m, other parameters, such as the scale factor and data length, were set as constant values (1 and 1000, respectively). The tolerance, r,



(**a**) (**b**)

Figure 4.1: Behaviour of the single-scale (**a**) MMSE and (**b**) veMSE, on the estimation of WGN, 1/f noise, and AR processes, as a function of the embedding dimension, m.

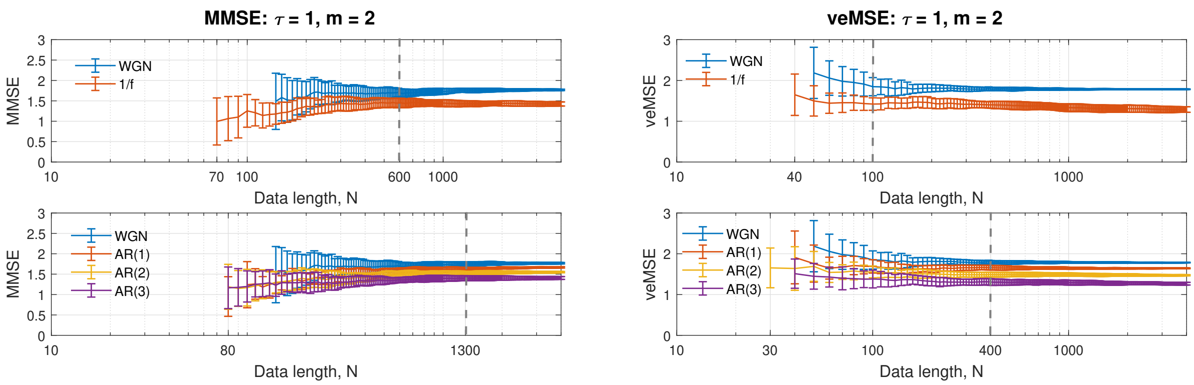
was varying, according to the total variance of the covariance matrix of processed data sets, as $r \times tr(S)$; here, the tolerance quotient was fixed to $r = 0.15$.

Figure 4.1(b), where the embedding dimension ranges from 1 to 9, shows that, unlike MMSE, the veMSE analysis was able to give a defined entropy value, even at high embedding dimensions and for a complex correlated structure, as, e.g., the AR(3) process at the scale of 7. Additionally, with standard MMSE, signals with higher randomness are more likely to yield unstable estimates as the embedding dimension increases. However, even in the case of white Gaussian noise with the highest randomness, the veMSE with the embedding dimension $m = 5$ was able to successfully and stably process the data. On the other hand, traditional multiscale sample entropy methods fail to give a defined value with the embedding dimension higher than 3, under the same condition [15]. Therefore, from the viewpoint of estimation stability, veMSE exhibits a marked improvement, when it comes to assessing complex information in high dimensions.

4.4.2 Varied Data Length (N)

The data length, N, of the signal is another limitation, in addition to the embedding dimension, when implementing entropy-based complexity calculations, particularly in real world processes. Indeed, amplitude distance-based entropy algorithms require at least 1000 data points to guarantee a consistent estimation, such as with the Multiscale Sample Entropy (MSE) and Multiscale Fuzzy Entropy (MFE) [112]. However, in real world data, as in the analysis of heart rate variability, for example, to obtain the required data size for RR intervals, a minimum of 5 min of the raw electrocardiograph (ECG) signals are needed. In practice, the implementation of such a long-time recording in a controlled state is hard to be satisfied. Compared to amplitude distance-based entropy methods, space distance-based entropy algorithms, such as cosine similarity entropy, show less restriction to data length, with a minimum of 700 samples required [15].

Figure 3.5(b) illustrates the performance of a single scale of veMSE, as a function of data length, N, in a logarithmic scale. The embedding dimension was set to $m = 2$, and the choice of tolerance was the same as before. The values of veMSE for white and $1/f$ noise were not defined before $N = 40$, while for AR(2), they were not defined when N was smaller than 30. The smallest sample sizes for the veMSE to reliably compute entropy estimates for AR(1) and AR(3) were also $N = 40$.



(**a**) (**B**)

Figure 4.2: Behaviour of the single-scale (**a**) MMSE and (**b**) veMSE, on the estimation of WGN, 1/f noise, and AR processes, as a function of the data length, N.

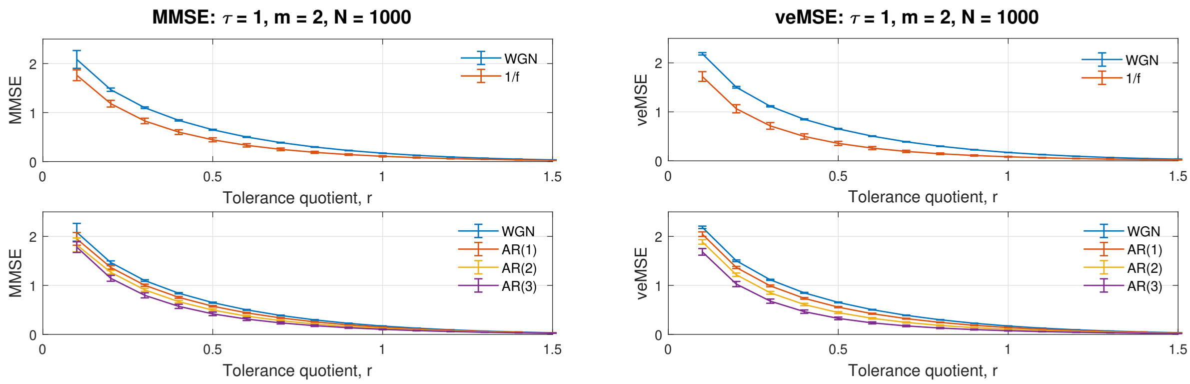
Observe that the standard deviation of the entropy results is gradually decreasing with an increase in data length, while the range for each error bar is increasing from WGN to AR(3). As a result, a system with more structure (AR(3)) reveals a larger standard deviation. In addition, the consistency of the estimation can be guaranteed as evidenced by the relative position of curves in each graph being unchanged as data length, N, increases. More importantly, when analysing the white and flicker noise, in the top panel in Figure 4.2(b), the estimation at the $N = 100$ sample length could successfully separate the complexity degrees of the two signals, while with veMSE, in the bottom panel in Figure 4.2(b), when the data length reaches $N = 400$ samples, there is no intersection region among entropy values from different models. This illustrates that the requirement of data length when applying veMSE is much lower than in other entropy methods, e.g., MMSE in Figure 4.2(a) requires $N = 1300$ for the separation of the four models. Subsequently, this property enables us to reveal the complexity information under high scales, even at a limited data length, but with a stable estimation, which better serves the balance between the dimension and data size.

4.4.3 Varied Tolerance (r)

The tolerance, r, can be explained as the boundary of the similarity degrees among comparing templates. The SampEn-based algorithms limit the tolerance to a hard threshold as a Heaviside function related to the standard deviation of the original data. However, for a multivariate case with multichannel data sets, only a single tolerance value is allowed in standard algorithms. As in [136], the choice of tolerance of veMSE is dependent on the total variance of the covariance matrix, S, of the analysed data sets. Therefore, the tolerance was set as $r \times tr(S)$.

Figure 4.3(b) illustrates single-scale entropy estimation, as a function of the tolerance parameter, r, varying from 0.1 to 1.5, at 0.1 increments. The data length and embedding dimension were fixed at $N = 1000$ and $m = 2$, to show the influence of the varied tolerance setting. Observe from the figures that, for all curves, the increase of tolerance quotient results in a monotonic decrease in complexity estimation, which is the same behaviour as with MMSE, in Figure 4.3(a). All curves can be obviously distinguished from each other, before $r = 1$ in veMSE, with the values after $r = 1$ too small for differentiation. The gap among different complexity estimations in the two figures obviously narrows down after

$r = 0.5$, therefore, supporting the choice for the value of tolerance quotient to be chosen below $r = 0.5$.



(**a**) (**b**)

Figure 4.3: Behaviour of the single-scale (**a**) MMSE and (**b**) veMSE, on the estimation of WGN, 1/f noise, and AR processes, as a function of the tolerance, r.

4.4.4 Varied Number of Channels (P)

The number of channels, P, is closely related to the setting of the embedding dimension, m, and the performance of veMSE. We have showed that the veMSE is able to give complexity estimation under higher embedding dimension in Figure 4.1(b), the higher embedding dimension will be assigned to consecutive data sub-channels. Figure 4.4(a)&(b) illustrate the influence of the number of data channels on MMSE and veMSE. The default parameters were set to $m = 2$, $N = 1000$, and $\tau = 1$.

Figure 4.4 illustrates single-scale estimation, as a function of the channel number, P, varying from 1 to 8. Observe that all the curves exhibit a decreasing trend for both algorithms, yet for different reasons. The drop in MMSE is resulted from the low level of correlation within the signals. Hence, the increase of channel number in MMSE has the same effect as the increase of scale factor, exhibited in the next sub-section. The decrease in veMSE is caused by the absence of information, hidden in high embedding dimensions, and simplicity of the system structure; that is, the extra channels fail to give more information, which could contribute to the complexity estimation. Despite this

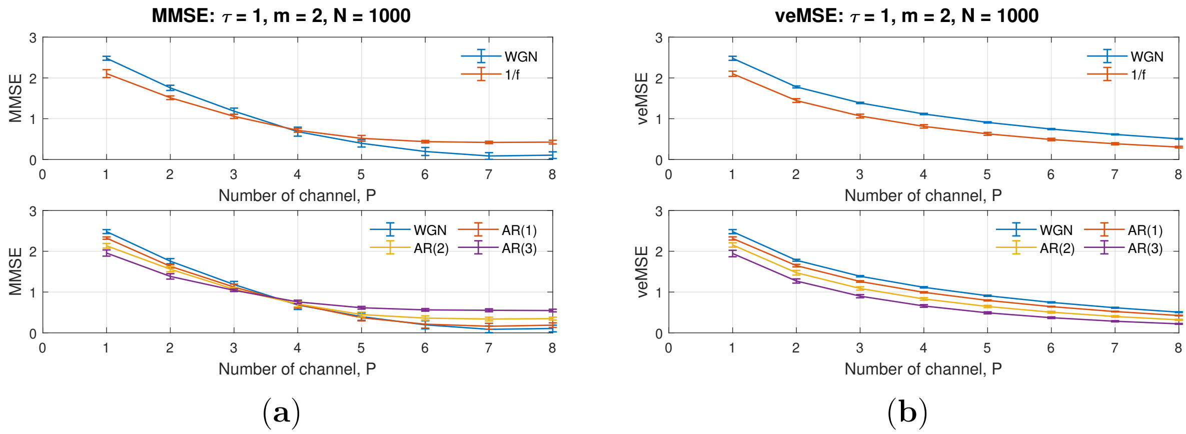


Figure 4.4: Behaviour of the single-scale (**a**) MMSE and (**b**) veMSE, on the estimation of WGN, 1/f noise, and AR processes, as a function of the channel number, P.

decrease, the estimation of veMSE is consistent for all the situations. The gap among different signals in Figure 4.4 keeps at a certain level for a high number of channels. Although the increasing channel number has limited impact on the ability of veMSE to separate systems, with various degrees of regularity, a too high embedding dimension, given by the large number of channel, might possibly fail to provide new information, due to no similar patterns in the following sub-channels. Therefore, according to the discussion in Section 4.4.1, the ideal uncorrelated WGN holds a well-defined value for embedding dimension, $m = 5$. We suggest the number of channels, ranging from 1 to 5, for the best performance of veMSE. In short, veMSE is more suitable for analysis of systems with limited number of channels and strongly correlated inter-channel structure.

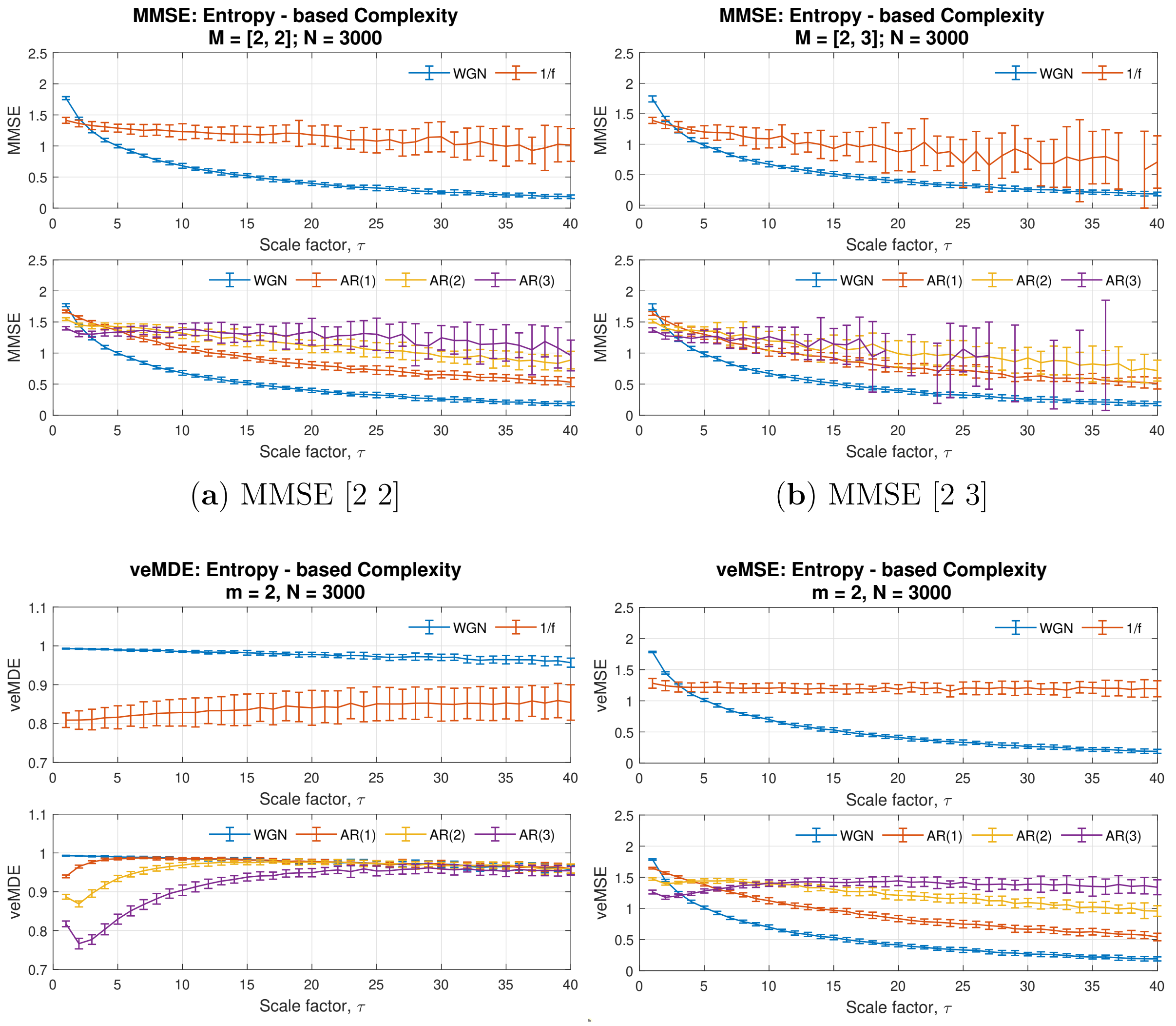
4.4.5 Varied Scale Factor

As noted by Costa et al. in [120], the multiscale analysis by integrating consecutive coarse-graining is of importance in signal processing associated with hidden correlation structure in data. Based on the aforementioned analysis of parameters involved in veMSE, the embedding dimension was set to $m = 2$, and the tolerance was chosen as $r = 0.15$, multiplied by the total variance of the covariance matrix. With regard to performance of multiscale analysis, graphs of multichannel entropy results are presented, in response to the scale factor, varying from $\tau = 1$ to $\tau = 40$. Dual channel (bivariate) data, with $N = 3000$ data points for each model, were considered.

To further elucidate the extent of improvements of the proposed veMSE, over the MMSE, its performances were compared against the proposed veMSE method. With the same data size, due to the varying embedding dimension feature of veMSE, two different settings of the parameter, related to embedding dimension for MMSE, were applied and are shown in Figure 4.5(a)&(b). In addition, the performance of the veMDE is also given in Figure 4.5(c), with the embedding dimension set to $m = 2$, while the results of veMSE are presented in Figure 4.5(d).

These figures demonstrate that complex structure, hidden in higher dimensions, is hard to unveiled via MMSE, as the standard deviation of those independent 20 realizations grows steadily as the scale factor increases. Overall, the embedding dimension pair $[2, 2]$ gives a better performance in MMSE for the considered restricted data length. However, even under the optimal dimension settings, as in Figure 4.5(a), the complexity of AR(3) in purple and AR(2) in yellow fails to be distinguished in multi-scale cases. As for the veMDE in Figure 4.5(c), Diversity Entropy is developed based on angular distance and Shannon Entropy, which measures a complex system from a different perspective (by the amount of information [31]). As can be seen in the graph, veMDE reveals a consistent estimation for each system considered, which exhibit short-term correlation. Our analysis focuses on the improvements of variational embedding fusion methodology over multivariate multiscale strategy, based on Sample Entropy. Details of the veMDE improvements by the variational embedding strategy, compared to other existing entropy methods, can be found in [140].

On the other hand, the merits of veMSE can be clearly seen from Figure 4.5(d). To better specify the improvement, the optimal dimension setting $[2, 2]$ of MMSE in Figure 4.5(a) was utilized for a comparison with veMSE. Observe from the top panel of the two types of noise in each figure. Although both of the two algorithms were able to distinguish



(**a**) MMSE [2 2]

(**b**) MMSE [2 3]

(**c**) veMDE

(**d**) veMSE

Figure 4.5: Behaviour of (**a**) & (**b**) MMSE, (**c**) veMDE, and (**d**) the proposed veMSE, for WGN, 1/f noise, and AR processes, with the embedding dimension set as (**a**) [2, 2] and (**b**) [2, 3] for MMSE, $m = 2$ for veMDE and veMSE.

between the two models, the complexity of white noise went down, while the flicker noise maintained a certain complexity level, in spite of the increasing scale. The range of entropy in error bars for flicker noise based on veMSE was much narrower than that in basis of MMSE, especially in large scales. Secondly, in the bottom panel in each figure, values of AR(2) and AR(3) (in yellow and purple) fail to be fully separated by MMSE in the cases of high scale, as stated above, while with the same data length, the separability of AR models of different orders was successfully accomplished in the proposed veMSE. It is critical to apply the entropy calculation under the multiscale situation, since the long-range correlation of the system is largely ignored in the analysis under low scale. Next, it can be observed that minor differences exist between the complexity estimation of the two models, namely white noise and AR(1) (blue and red line in bottom graph). Instead, the enhancement properties of veMSE is particularly revealed in the analysis of highly correlated and structured signals, as well as systems with higher structural complexity.

Overall, the comparison of veMSE and MMSE, based on the above five models, shows that the veMSE provides a more stable estimation that can better demonstrate complex temporal fluctuations. In addition, veMSE is especially suitable for multiscale analysis of highly correlated signals, which exhibit variation of spatial–temporal patterns over a range of scales.

4.5 Properties of veMSE

We now elaborate on the three desired properties of the proposed entropy based veMSE algorithm: noise robustness, directionality, and calculation efficiency. The parameter setting was as follows: data size $N = 3000$; embedding dimension $m = 2$; tolerance $r = 0.15$; and scale factor $\tau = 1$. A bivariate system was considered in the analysis. The results which depict the noise analysis and directionality analysis based on the proposed veMSE are presented in Figures 4.6(b) and 4.7(b), respectively. The corresponding performance of MMSE is shown in Figures 4.6(a) and 4.7(a). The time requirement for the calculations of veMSE and MMSE is shown in Figure 4.8.

4.5.1 Noise Robustness

Robustness against noise and artefacts is of critical importance in any estimation. Given that it is infeasible to avoid the noise associated with recording equipment and the ubiquity of artefacts in biosignals, for instance, muscle and electro-magnetic artefacts exist in EEG-based monitoring [163], the noise-robustness property was tested, by comparing the complexity estimation for AR models with and without corrupting noise. In Figure 4.6(b), the top panel presents the curves for uncorrelated white Gaussian noise (WGN), correlated flicker noise ($1/f$ noise), and coloured noise, containing both WGN and $1/f$ noise. Observe that the three systems with different degrees of correlation were successfully separated by the veMSE. Adding white noise will enhance short-term correlation, as shown at the left of the top panel in Figure 4.6(b), where the yellow line (`1/f + WGN`) is as high as the blue line (`WGN`), while the long-term correlation is lower as the scale factor increases (`1/f` $>$ `1/f+WGN` $>$ `WGN`). The top panel in Figure 4.6(b) reveals that veMSE could correctly yield complexity estimation, in line with the theoretical analysis, on the basis of uncorrelated and correlated noise.

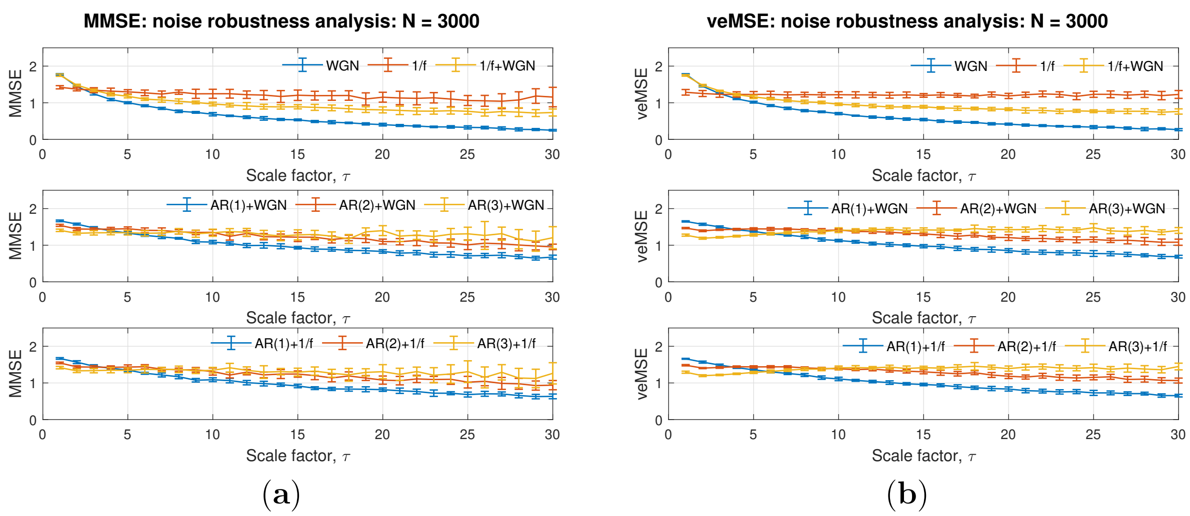
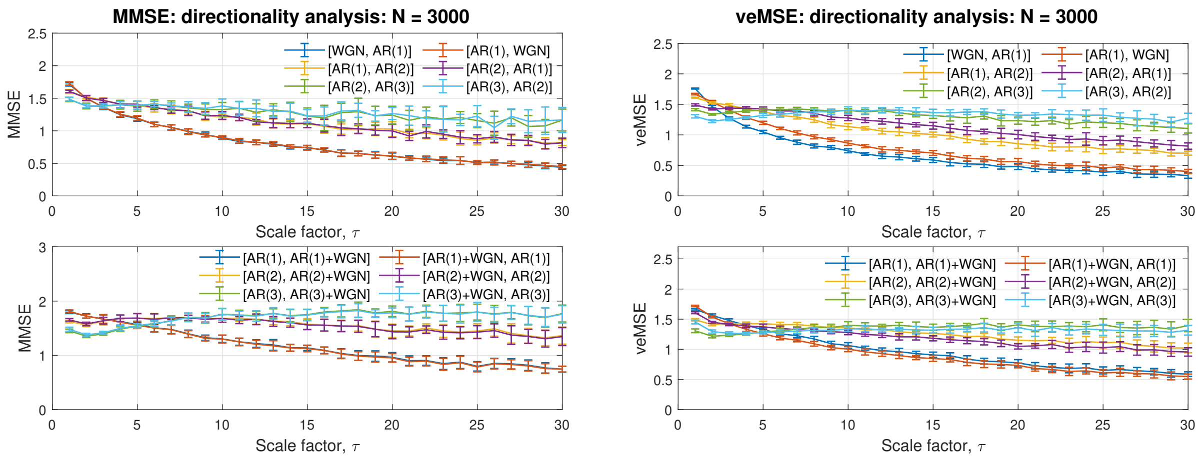


Figure 4.6: Illustration of robustness, with respect to noise on (**a**) standard MMSE and (**b**) the proposed veMSE.

In the middle and bottom panel in Figure 4.6(b), the results of veMSE for AR models with uncorrelated white noise and correlated flicker noise are presented, to contrast to the outcomes of pure AR signals in Figure 4.5(d). The amplitude of the added noise signal was set to 20% of that for the AR signals. Compared to Figure 4.5(d), the gaps between the complexity curves for the AR models of varying order decrease with noise. However, even that the gap among distinct models is narrowed down, separation can still be achieved at high scales in 20% of the noisy scenarios, while in case of MMSE, noisy AR signals with different complexity cannot be well separated and the impact of noise is clearly shown in Figure 4.6(a). Given these points, the performance of complexity estimation, based on veMSE, is consistent with cases without noise, a unique feature of veMSE that is not present in the other MSE algorithms, thus demonstrating the potential in practical recording data sets.

4.5.2 Directionality

For multivariate analysis, the directionality refers to the issue that the optimal ordering of the input channels is unknown; it is, therefore, desirable that an algorithm is independent of channel ordering. Yet, without prior knowledge related to the optimal channel order, the performance of the estimation will be impacted in standard entropy-based algorithms. To this end, the directionality of the veMSE is next analysed for bivariate systems. Figure 4.7(b) shows two graphs, each containing three pairs of curves. The top panel depicts the results for white noise with AR(1), AR(1) with AR(2), and AR(2) with AR(3), with the order of input shown by legend (first present, first processed). As can be seen from this figure, the estimates at a lower scale are mainly influenced by the first input signal. For example, the blue line `[WGN, AR(1)]` can be clearly recognized as lower than the red one `[AR(1), WGN]` at the beginning, especially in the single-scale case. As the scale increases, the two lines approach the same level; a similar trend can be seen for the other two pairs.



(**a**) (**b**)

Figure 4.7: Influence of directionality on (**a**) standard MMSE and (**b**) the proposed veMSE. Each figure comprises three pairs of bivariate systems with an inverted order of channels. The smaller the difference in each pair, the less biased the method would be by the order of channels.

In the bottom panel in Figure 4.7(b), the analysed signals are AR(1), AR(2), and AR(3), with one of the signals in each system associated with white noise. The legend `[AR, AR+WGN]` refers to cases where noise-free signals are the first variate, followed by noisy signals, and vice versa for `[AR+WGN, AR]`. This setting of the inputs was used to simulate real world scenarios when dealing with multi-channel signals, whereby one of the constituent channel represents a poor recording with noise. Considering the noise-robustness property of veMSE, the amplitude of noise signals in this subsection was enlarged to the same level as for the AR signals, to demonstrate a clear difference when the input order is altered. As shown in the figure, the inverted input orders can be reflected by different start levels, while the complexity curves then approach each other, as well as ending with similar estimates. Therefore, regardless of the input order, the separation of complexity levels of AR models was successfully achieved with the proposed veMSE algorithm, as observed in Figure 4.7(a). In the case of MMSE, the inverted input exhibited no influence on the resulted curves in small scales, while similar performance as veMSE in the larger scale.

As demonstrated in Figure 4.7(b), for the proposed veMSE, the reversed order had little influence on the estimation at high scales as all the paired curves approach to the same three regions, so that, in spite of the modified order, the three models were separated. However, a similar phenomenon is shown in the top panel in Figure 4.7(b), the varying order of the input signals will generate entropy values with different degrees at small scale analysis when the input signals contain distinct structure. Therefore, the direction of the input order needs to be carefully considered when applying small scale analysis, and such considerations can be ignored at high scale analysis with identical system measurements.

4.5.3 Computational Complexity

Entropy analysis based on multichannel signals is more time-consuming than univariate estimation, so that calculation efficiency becomes one of the critical factors that needs to be carefully considered. Therefore, in this subsection, time consumption of veMSE is discussed and compared with the commonly used MMSE.

Figure 4.8 shows the processing time, as a function of various modified parameters, when implementing the veMSE (blue line) and MMSE (red line). All the curves are produced as an average over 10 independent realizations in MATLAB conducted on a laptop with a CPU base frequency of 1.60GHz. Each graph is designed to reflect the behaviours for only one modified parameter, with the independent variables in the following figures, from the left- to the right-hand side, as the scale factor, length of data (in log-log scale), number of channels, and embedding dimension. All the entropy calculations were set as single-scale and bivariate processing by default. The data length and embedding dimension were irrelevant variables and were fixed to $N = 5000$ and $m = 2$.

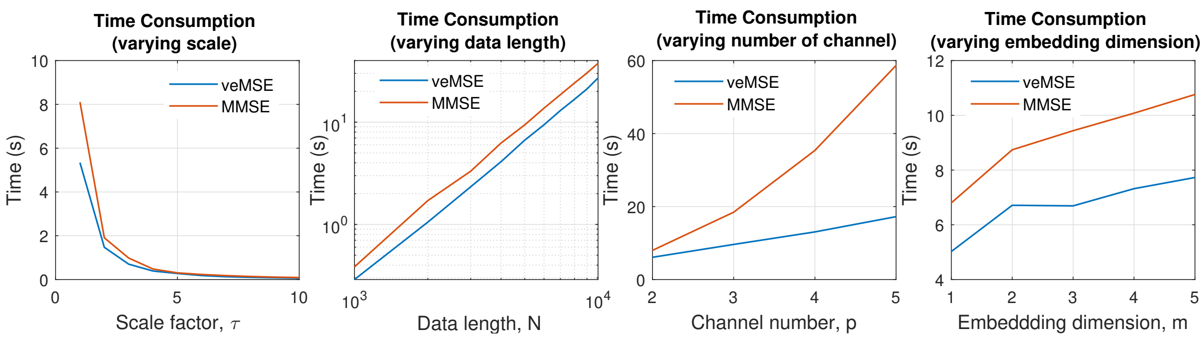


Figure 4.8: Computation time for MMSE and the proposed veMSE, with the modified parameters for a white Gaussian input.

Overall, the red line, representing the computational time of the standard MMSE, was above the blue, that of veMSE, for all the scenarios. The increase of scale factor reflects that when the data size, after 'coarse-graining' procedure, is lower than $N = 1000$, the times needed for the two calculations are similar, as shown in the left most graph where the scale factor is higher than 5. The relation between the computational load and data length in the second graph from the left indicates that veMSE has the dependency of $O(N^2)$ on the data size as standard MMSE, given in the log-log scale plot. As for the influence of the number of channels, demonstrated in the third graph from the left, it is reasonable that MMSE needs more time as the channel number increases, because the key step for Sample Entropy is the ratio of conditional probability for similar patterns between the embedding dimension, m, and its increment, $m+1$, whereby the number of possible ways to apply the $(m+1)$-dimension is equal to the number of channels involved when forming the composite delay vector in MMSE. Therefore, the calculation, with an increased embedding dimension, will be repeated c times in MMSE, where c denotes the number of data channels. Finally, in the relationship between an increased embedding dimension and computation time in the right-most panel, the time difference roughly maintains a fixed value, in spite of the embedding dimension changing.

In summary, compared to the widely used MMSE, the time needed for the same amount of data with the proposed veMSE is shorter. Therefore, the calculations efficiency of veMSE is higher than that of MMSE, which gives it high potential in real-time monitoring of human states.

4.6 Performance of veMSE on Real Signals

4.6.1 Wind Dynamics

Having illustrated the advantages of veMSE on synthetic signals, we shall now examine the performance of veMSE in real-world systems. We first considered trivariate wind data of different dynamics. The long-term correlations in wind dynamics has been revealed in previous studies by using Fetrended Fluctuation Analysis (DFA) [164] and standard MMSE [136].

Here, the proposed veMSE showed an improved performance in characterizing different dynamical complexity of wind regimes.

Wind data was recorded by 3D ultrasonic anemometers, at a sampling frequency of 50 Hz. The recording process was implemented in the Institute of Industrial Science of the University of Tokyo [136]. Three channels, containing wind speed into the east-west, north-south, and vertical direction, were recorded. The wind regimes, containing different dynamics, were defined as the low, medium, and high by the magnitude of wind speed, as shown in Figure 4.9. The parameters involved in veMSE analysis were set to $m = 2$, $L = 1$, $r = 0.2 \times tr(S)$, and $N = 3000$. Illustrations of entropy results were averaged over 10 trials for each channel and displayed as error bars, based on the standard error of the mean. In addition, time-shuffled wind data sets were also tested and compared with the obtained wind dynamics. The shuffled wind samples were generated from the recorded wind signals but with a randomised order. In this case, the possible correlations within signals were broken down but the statistical properties were preserved.

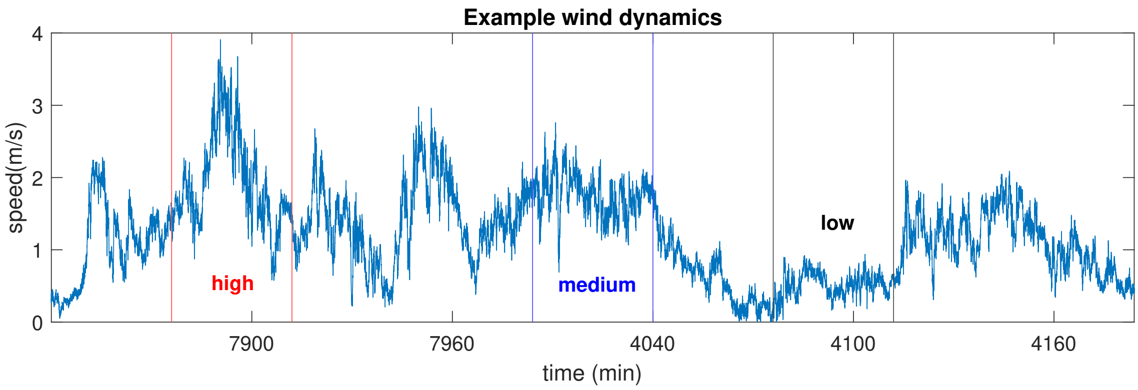


Figure 4.9: Magnitude of the wind signal considered. The wind segments are defined as the low, medium, and high regimes.

Before performing the comparison between veMSE and MMSE, univariate MSE was firstly applied to give an insight into the complexity of the dynamical system to be analysed. As shown in Figure 4.10(a), the randomized data sets exhibit a white noise-like behaviour, in terms of structural complexity, which is expected, as the correlation was destroyed while shuffling data samples. The MSE results of wind dynamics assigned the highest complexity to the medium dynamics regime, which is expected, since the medium wind speed held the fewest constrains. The complexity of low wind regime was quantified as higher than that of the high regime, which is against the intuition, as the high wind regime also contains components with medium speed. Besides, each wind dynamics exhibited lower complexity than the their randomized series, wrongly suggesting the absence of long-term correlations within the wind data sets.

The same trails were analysed by MMSE, as shown in Figure 4.10(b), which revealed the relation of the three wind regimes in an expected way; that is, with mild wind exhibiting the highest complexity, followed by the high and low dynamics regimes. Further, all the three wind regimes were able to present higher long-range correlation than their surrogate data sets, after $\tau = 5$; however, the overlapped areas can be found across all the scales. Observe from Figure 4.10(c) that, as desired, although the scale factor need to be set up to 11, the long-range correlation of different wind speeds was detectable by the proposed veMSE, resulting in a more obvious separation among the different wind regimes, as compared to MSE and MMSE.

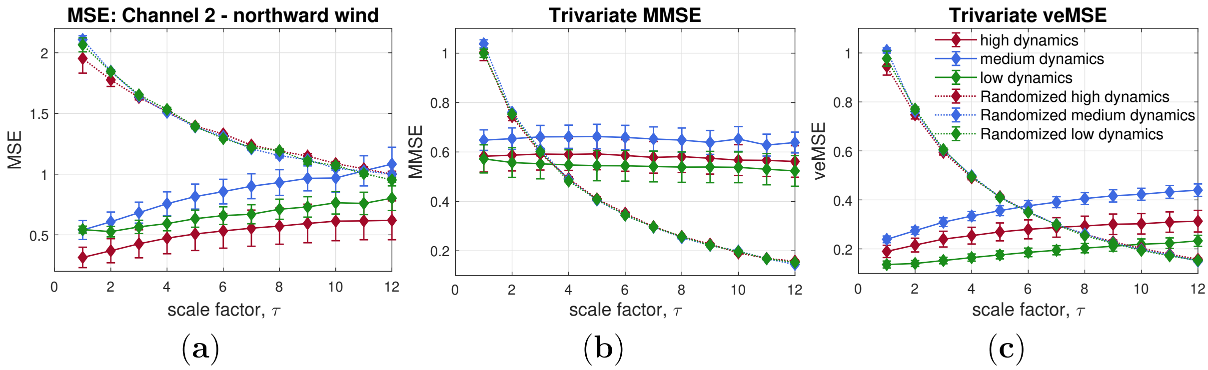


Figure 4.10: Complexity analyses of real-world wind data; (**a**) Standard univariate MSE; (**b**) Standard multi-channel MMSE; (**c**) Proposed multivariate veMSE.

4.6.2 Physiological Data

The physiological signals utilized were from the Fantasia database [165], which includes the RR intervals (RRI) and interbreath intervals (IBI), extracted from the electrocardiogram (ECG) and respiration signals, respectively. The structure of long-term correlation in heart rate variability (HRV) and respiratory dynamics has been examined by traditional methods, namely detrended fluctuation analysis (DFA) [166] and standard MMSE [136].

The Fantasia database contains recordings from 20 young people (the age of 21–34) and 20 elderly participants (ages from 68–85) who were rigorously screened for 120 min; we chose 10 subsets from each group. The ECG and respiration signals were recorded at the sample frequency of 250 Hz. Then, RRI and IBI signals were extracted and aligned, and the surrogate series with randomised order were analysed, along with the RRI and IBI signals for each methodology, to reveal the effectiveness of the complexity analysis. The parameters in entropy analysis were set to $m = 2$, $L = 1$, $r = 0.15 \times tr(S)$, and $N = 4000$.

In the first place, to give an overall view of the signals in each channel, univariate MSE was applied to single channels of RRI and IBI, separately, as shown in Figures 4.11 and 4.12. As complexity loss theory [45] states, the adaptive capacity of bio-systems is damaged by disease and ageing. The complexity of the RRI and IBI signals recorded from young people is, therefore, supposed to be higher than that from elderly subjects, when considering long-term correlation. Regarding short-term correlation in RRI, the MSE exhibited correct estimation under certain low scales. However, the MSE for RRI and IBI was unable to reveal the correct relation between the dynamics of elderly individuals and young people in long-range terms.

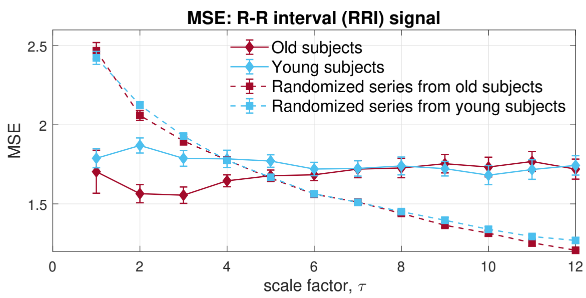


Figure 4.11: Behaviour of univariate MSE, applied to RR intervals within the Fantasia database.

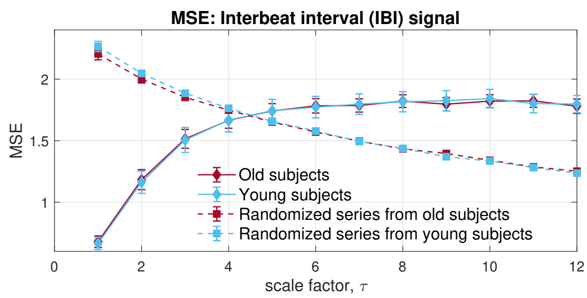


Figure 4.12: Behaviour of univariate MSE for the interbreath intervals (IBI) within the Fantasia database.

Next, the MMSE was applied to the bivariate channels composed of the RRI and IBI signals. The resulted error bars indicate higher dynamical complexity in young participants than of that in elderly individuals, which conforms with the complexity loss theory due to ageing [45]. Moreover, both physiological data sets showed higher long-range correlation than the randomized surrogate series in an expected way; yet, as scale increases, an overlapping area can be observed, due to the lack of sufficient sample length at larger scales, owing to the coarse graining process. In contrast with MMSE, the veMSE in Figure 4.14

gave the correct and clear estimation, with enhanced stability, particularly in large scales. In practical scenarios, the size of recorded signals is mostly limited; hence, the veMSE better facilitates the analysis of physical and physiological signals when identifying dynamical differences based on nonlinear features.

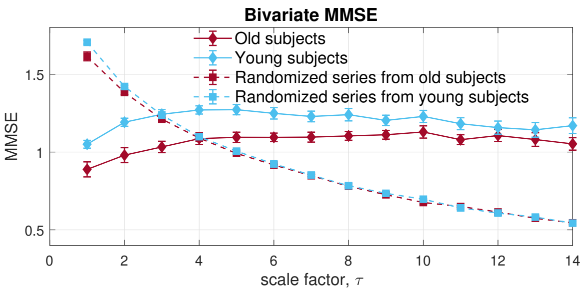


Figure 4.13: Behaviour of MMSE for the bivariate [RRI, IBI] data from the Fantasia database.

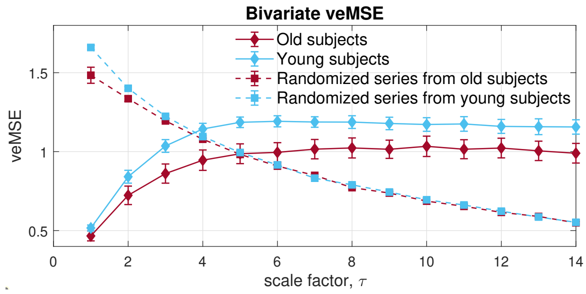


Figure 4.14: Behaviour of the proposed veMSE for the bivariate [RRI, IBI] data from the Fantasia database.

4.7 Conclusions

The variational embedding Multiscale Sample Entropy method has been introduced for robust structural complexity analysis of real-world data. It has been shown that veMSE is capable of assessing the complex features of the system at large scales and with higher embedding dimensions, compared to the standard MMSE. In addition, the utilization of multivariate analysis via veMSE guarantees an improvement over univariate analysis, regardless of the quality of the recorded signals in sub-channels. The veMSE has also been shown to exhibit a strong noise robustness and lower computational complexity than MMSE, under the same conditions. As desired, this improvement is apparent as the number of available channel increases within a certain range. The higher calculation efficiency within veMSE is of high interest when applying entropy analysis in scenarios, which requires near real-time processing or synchronized monitoring.

However, a very large number of channels will lead to inefficient measures, while the common problem of the Sample Entropy-based method still remains; that is, the irregularity of the signal is insufficient to quantify the complexity of the system. Therefore, future work can be considered from several aspects. First, the scaling process, applied in this method, will be further examined, and the multi-scale procedure can be further improved by balancing on computational efficiency. Secondly, the choice of time lag needs to be further explored, to contribute to optimal temporal span. Thirdly, as already stated, to further explore the property of complexity, the estimation need to be considered from various perspectives, other than the current irregularity-based analysis, such as causality, determinism, nonlinearity, and other features. Lastly, it should be noted that the veMSE method is restricted to amplitude-based distance. Future research will focus on angular distance-based algorithms, which employ the variational embedding dimension methodology.

Chapter 5

Multivariate Multiscale Cosine Similarity Entropy

5.1 Chapter Summary

The extension of Sample Entropy methodologies to multivariate signals has received considerable attention, with traditional univariate entropy methods, such as Sample Entropy (SampEn) and Fuzzy Entropy (FuzzyEn), introduced to measure the complexity of chaotic systems in terms of irregularity and randomness. The corresponding multivariate methods, Multivariate Multiscale Sample Entropy (MMSE) and Multivariate Multiscale Fuzzy Entropy (MMFE), were developed to explore the structural richness within signals at high scales. However, the requirement of high scale limits the selection of embedding dimension and thus, the performance is unavoidably restricted by the trade-off between the data size and the required high scale. More importantly, the scale of interest in different situations is varying, yet little is known about the optimal setting of the scale range in MMSE and MMFE. To this end, we extend the univariate Cosine Similarity Entropy (CSE) method to the multivariate case, and show that the resulting Multivariate Multiscale Cosine Similarity Entropy (MMCSE) is capable of quantifying structural complexity through the degree of self-correlation within signals. The proposed approach relaxes the prohibitive constraints between the embedding dimension and data length, and aims to quantify the structural complexity based on the degree of self-correlation at low scales. The proposed MMCSE is applied to the examination of the complex and quaternion circularity properties of signals with varying correlation behaviours, and simulations show the MMCSE outperforming the standard methods, MMSE and MMFE.

5.2 Introduction

The investigation of entropy has been a considerably important topic in nonlinear analysis, especially in relation to the quantification of the degree of irregularity and multiple correlations within time series. Entropy-based complexity studies have been implemented in many physical and physiological studies, mostly based on the widely accepted complexity loss theory (CLT) [59]. The CLT states that the physiological responses of the human body in a pathological state (resulting from, for example, stress, illness, or ageing) exhibit a loss in structural complexity, while the highest degree of irregularity is associated with healthy and young bio-systems [37, 59].

The concept of entropy derives from thermodynamics to measure the probability of microstates in physics [167]. When it comes to Shannon's information theory, entropy is employed to describe the amount of information and the generation rate of new information [25, 70]. Among numerous methods to characterize signals in terms of entropy, two commonly used techniques are Approximate Entropy (ApEn) [105] and Sample Entropy (SampEn) [26]. Both employ the Chebyshev distance and utilize estimates of the conditional probability of similar patterns within the time series. SampEn is usually considered an enhancement of ApEn, as it gives less biased estimation and higher robustness [26]. Subsequently, SampEn has found applications in both physical and physiological systems [49, 80, 145, 168–170]. For example, research based on SampEn of the electrocardiogram (ECG) has been utilized to estimate different degrees of atrial fibrillation [110], or the analysis of stress in public speaking [111]. A further improvement of SampEn, called Fuzzy Entropy (FuzzyEn), employs fuzzy functions as a substitute for the hard thresholding within SampEn in the definition of similar patterns [112]. Given a continuous relation between the similarity and the obtained distance, FuzzyEn is capable of giving a more reliable and well-defined estimation, while requiring shorter data length compared to SampEn [112].

Despite the broad use of SampEn and FuzzyEn [1, 112, 137, 171, 172], their shortcomings still remain when it comes to real-world data. The problems include the following:

- The distance measure utilized in SampEn and FuzzyEn is based on the amplitude of the time series, where the unlimited range would lead to undefined values.
- Due to the signal related parameter setting, when the tolerance is set according to the value of standard deviation, the data normalization is required in advance, which would re-scale the original signal and cause information loss.
- Considering the requirement of re-scaling, this affects the noise robustness of SampEn and FuzzyEn, in particular, when processing short time series.
- Long data length is required for traditional SampEn and FuzzyEn analyses when exploring complexity with a high embedding dimension.

Recently, a new theory associated with self-correlation states that pathology is simultaneously followed with an increase in structural complexity (e.g., correlation within time series), which can be considered a complementary theory to the traditional complexity loss theory (CLT) [15]. This promises to help resolve the problems with SampEn- and FuzzyEn-based methodologies, which fail to give a holistic quantification of the self-correlation of the target signals. Unlike other existing entropy methods, Cosine Similarity Entropy (CSE) was introduced to quantify the structural complexity of time series in terms of self-correlations based on angular distance [15].

The CSE provides a robust and more meaningful estimation of the structural complexity of chaotic systems, in contrast to SampEn and FuzzyEn [15]. The extension of CSE to multivariate case is introduced in this chapter, whereby the proposed Multivariate Multiscale Cosine Similarity Entropy (MMCSE) is based on the idea of the composite delay vector introduced in Multivariate Multiscale Sample Entropy (MMSE) [136] and the rule of angular distance applied in CSE [15]. The performance of MMCSE is first evaluated over all parameter settings, and then applied to the detection of complex and quaternion non-circularity in multi-channel systems, whereby different correlation behaviours were examined and discussed by multivariate entropy methods. The virtues of MMCSE are further illustrated by correctly quantifying the meaningful structural complexity for varying conditions of correlation degrees and power in data channels.

The reminder of the chapter is organized as follows. Related work and development of SampEn and FuzzyEn-based entropy are reviewed in Section 5.3. To contrast the structure of the CSE method, traditional SampEn and FuzzyEn along with their multivariate extended forms are presented and reviewed in Section 5.4. The details of the extension of CSE to the multivariate case are given in Section 5.5. Section 5.6 discusses the selection of parameters and provides the suggested values for the manually set parameters when applying MMCSE. Section 5.7 illustrates the application of MMCSE in the detection of circularity in terms of varying correlation behaviours. Finally, the conclusions are given in the last section.

5.3 Related Work and Development of Existing Sample Entropy and Fuzzy Entropy

Although FuzzyEn exhibits a higher consistency at the expense of computational load, single-scale processing has shown limitations in real-world entropy analysis. Hence, the 'coarse-graining process' (CGP) was subsequently introduced in traditional entropy methodologies to yield multiscale entropy [120], which allows for the examination of information at high scales. This is achieved by taking the average over neighbouring data points. However, it was realized that CGP is not the optimal method to obtain the scaled signal because the traditional CGP in the frequency domain operates as a low-pass filter with large stop band ripples, resulting in biased scaled signals containing artefacts [157]. In addition, the averaging procedure of CGP requires a large data set and comes with a loss of the high-frequency component in data [132]. To overcome the first drawback of CGP, refined multiscale entropy was introduced by employing a low-pass Butterworth filter to generate unbiased scaled signals [157]. As a refinement of CGP in terms of the second flaw, several improved methods have been proposed. For instance, composite Multiscale Sample Entropy operates by implementing CGP at various starting points and taking the average over the resulting entropy values [158], while generalized multiscale entropy employs higher moments (e.g., variance) to give the dynamics over multiple scales in place of the mean [122]. Another multiscale strategy that has been applied in entropy analysis is named intrinsic Mode Entropy (IMEn), and employs Intrinsic Mode Functions (IMFs) obtained by Empirical Mode Decomposition (EMD), whereby the scaled signals in IMEn are obtained by cumulative sums of IMFs to introduce robustness toward low-frequency changes [173]. However, the additional processes used to improve CGP inevitably increase the computational time. Despite these drawbacks of CGP, it is widely used in real-world applications since it balances computation efficiency with reasonable entropy estimates.

Multivariate Multiscale Sample Entropy (MMSE) was proposed as an extension to SampEn in order to simultaneously process information from multiple channels and relax the requirement of long data length for univariate entropy [136]. For multi-channel data, MMSE provides a more powerful and accurate estimation, even for limited data length [136]. Multivariate Multiscale Fuzzy Entropy (MMFE) was subsequently introduced, which maintains the merits of FuzzyEn in contrast to SampEn [137], in the multivariate setting. Both FuzzyEn- and SampEn-based methods are built based on the concept of amplitude distance, whereby outliers manifest themselves as spurious peaks and have a severe impact on the accuracy of estimation of structural complexity [15]. Further, the complexity quantified by SampEn or FuzzyEn is based on randomness and irregularity, which is one aspect of complex systems, while other properties, such as long-range correlations, fail to be demonstrated by single-scale SampEn and FuzzyEn [13]. Note that the scale of interest in multiscale entropy, based on amplitude distance, has no unified measure indicators as criteria when the scale is guaranteed to be set as a certain value to exhibit the long-range correlation of the signal.

5.4 Multivariate Multiscale Sample Entropy and Multivariate Multiscale Fuzzy Entropy

In this section, two traditional entropies, SampEn and FuzzyEn, are introduced and compared. Starting from a univariate single-scale process to the multivariate multi-scale approach, both SampEn and FuzzEn are developed based on the amplitude distance and the conditional probability of similar patterns.

5.4.1 Sample Entropy and Fuzzy Entropy

The details of SampEn are given in Algorithm 16 in Chapter 2. In standard SampEn, the original signal is first reconstructed in the phase space following Takens' embedding theory [66]. Each embedding vector is jointly described by the embedding dimension, m, and time delay, l. Then, the Chebyshev distance is applied to pairs of the embedding vectors to give the degree of similarity, where a similar pair is defined by the Heaviside function with the tolerance coefficient, r. In the final stage, the probability of similar patterns is produced, and the estimated result is generated by the ratio of the probabilities given in increased phase spaces.

Similarly to SampEn, FuzzyEn employs the Chebyshev distance to give the similarity between paired embedding vectors. Here, the embedding matrix needs to be centered prior to the reconstruction. Next, the embedding matrix, $\mathbf{X}_m(i)$, is formed and the distance matrix $\mathbf{D}_m(i,j)$ is computed. The details of FuzzyEn are presented in Algorithm 17 in Chapter 2. The main modification of FuzzyEn from SampEn is the involvement of a continuous transformation from the distance, $d_m(i,j)$, to the similarity, $s_m^{r_f}(i,j)$, through the fuzzy membership functions as given in the Step 3 in Algorithm 17. Benefiting from the employment of the fuzzy membership functions, undefined estimation due to the absence of matching pairs is largely avoided. However, a higher computational load is required compared to SampEn, and the performance of FuzzyEn generally shows improved but also similar trends as with SampEn.

5.4.2 Multiscale Entropy

The information provided by single-scale entropy is limited, as it does not account for long-range correlations and temporal fluctuations over multiple scales [120]. To this end, by employing the consecutive coarse-graining process (CGP) with a new parameter denoted as the scale factor, τ, Costa et al. proposed multiscale entropy [120], given in Algorithm 23.

Algorithm 23. Multiscale Sample Entropy and Multiscale Fuzzy Entropy.

Assume that a univariate data set, $\{x(i)\}_{i=1}^{N}$, is of length, N, and the coarse graining scale factor is donated as τ.

1. Obtain the scaled time series, $\{y(j)\}_{j=1}^{N/\tau}$, by the coarse graining process as
$$y^{(\tau)}(j) = \frac{1}{\tau}\sum_{i=j-\tau/2-1}^{j+\tau/2-1} x(i), \quad 1 \leq j \leq \frac{N}{\tau}.$$
2. Apply the scaled data set, $y^{(\tau)}(j)$, as the input into SampEn given in Algorithm 16 or FuzzyEn in Algorithm 17 to obtain the complexity estimation based on multiscale sample entropy and multiscale fuzzy entropy, respectively.

5.4.3 Multivariate Multiscale Entropy

The implementation of CGP reveals the hidden information buried in multiple scales. However, the averaging process requires multiple times larger data size. To this end, multivariate multiscale entropy based on SampEn (MMSE) was proposed [136], followed by Multivariate Multiscale Fuzzy Entropy [137]. There are two main modifications in multivariate entropy algorithms based on amplitude distance. One is that the embedding vector in the univariate SampEn is replaced by the composite delay vector (CDV) composed by the embedding vectors from each channel in multivariate entropy. Figure 5.1 shows an example of composite delay vector construction for $m = 3$ and $l = 1$. In addition, due to the different range of amplitudes across channels, the obtained distance between pairwise composite delay vectors can be biased. Therefore, the second modification of multivariate entropy is the requirement of normalization of input data sets. As stated in [137], the performance of MMFE is proved to be more consistent than MMSE, particularly in high-dimensional phase spaces. The processes of MMSE and MMFE are summarized in Algorithm 24.

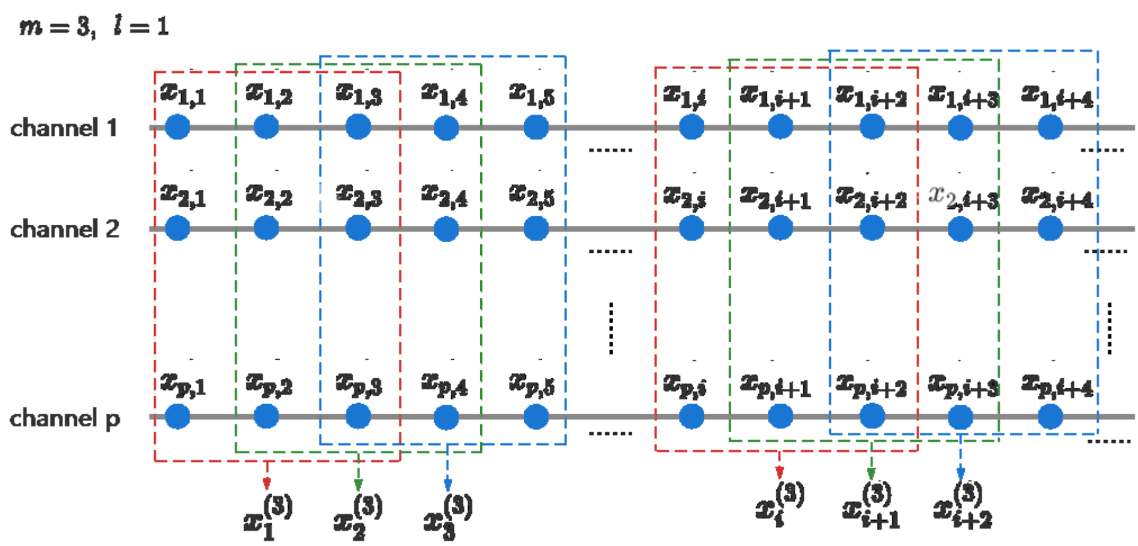


Figure 5.1: Example of composite delay vectors construction in multiscale entropy

5.5 Cosine Similarity Entropy and Multivariate Approach

Multiscale cosine similarity entropy (MCSE) was proposed [15], whereby instead of amplitude based distance, CSE employs the angular distance in phase space to define the difference among embedding vectors. The angular distance offers advantages, especially

Algorithm 24. Multivariate Multiscale Sample Entropy and Multivariate Multiscale Fuzzy Entropy

Given a multivariate data set with P channels $\{x_{k,i}\}_{i=1}^{N}$, $1 \leq k \leq p$, of length N. Parameters involved are the embedding dimension, $M = [m_1, m_2, \dots, m_p]$, tolerance, r, time delay, $L = [l_1, l_2, \dots, l_p]$, and the scale factor, τ.

1. Standardize the original multivariate data sets by subtracting the mean and dividing by standard deviation for each channel.
2. Coarse graining is next applied to the normalized data sets, $\{y_{k,i}^{(\tau)}\}_{j=1}^{N/\tau}$, for each channel following Algorithm 23.
3. Form the composite delay matrix, $\mathbf{Y}_M(i)$, according to the embedding dimension, M, and the time delay, L, in the form

$$\begin{aligned}\mathbf{Y}_M(i) = [&y_{1,i}, y_{1,i+l_1}, \dots, y_{1,i+(m_1-1)l_1}, \\ &y_{2,i}, y_{2,i+l_2}, \dots, y_{2,i+(m_2-1)l_2}, \\ &\vdots \\ &y_{p,i}, y_{p,i+l_p}, \dots, y_{p,i+(m_p-1)l_p},]\end{aligned}$$

4. Apply the scaled composite delay matrix, $\mathbf{Y}_M(i)$, as the input into the SampEn presented in Algorithm 16 or FuzzyEn in Algorithm 17 to obtain the measure of MMSE and MMFE, respectively.

regarding the sensitivity to outliers or sharp changes in time series that amplitude-distance-based calculation suffers from. Additionally, the angular distance, restricted by the maximal length of 2π, is more robust, and is less prone to generating undefined estimates in the presence of noise. In amplitude-based entropy calculation, because of the large distance that is obtained by amplitude, the tolerance, r, is set as a ratio associated with the standard deviation of the input data set, where parameters driven by data can be unstable for signals with high variance. Given that CSE is restricted to values within $[0, 1]$, the tolerance for CSE is selected, irrelevant to the variance of original signal. Hence, the process of CSE exhibits enhanced stability, especially when dealing with highly dynamical signals. The virtue of CSE is therefore that entropies can be assessed more stably in multivariate analysis when the data sets across channels have different variance.

The extension of MCSE to multivariate scenarios employs the construction of the composite delay vector in MMSE [136] and follows the main procedure of CSE [15]. Note that MMCSE is sensitive to the DC offset, as it requires the origin coordination to project the angular distance within the tolerance angle. Therefore, the DC offset or a long-term trend needs to be removed in MMCSE. Here, we apply the zero-median approach to remove the global trend because of its robustness against outliers. It is worth highlighting that the definition of CSE is developed based on Shannon Entropy, where the calculation only relies on the probability of phase space, m, without the involvement of phase space $m+1$ [15], which is less computationally expensive. The proposed MMCSE is presented in Algorithm 25.

Algorithm 25. Multivariate Multiscale Cosine Similarity Entropy.

Given a multivariate data set with P channels $\{x_{k,i}\}_{i=1}^{N}$, $1 \leq k \leq p$, of length N. Parameters involved are designated as the embedding dimension, $M = [m_1, m_2, \ldots, m_p]$, tolerance, r_c, time delay, $L = [l_1, l_2, \ldots, l_p]$, and the scale factor, τ.

1. Remove the DC offset of the original data sets $\{x_{k,i}\}_{i=1}^{N}$ by subtracting the median for each channel.
2. Obtain the scaled time series $\{y_{k,i}^{(\tau)}\}_{j=1}^{N/\tau}$ following the instruction of Algorithm 23.
3. Reconstruct composite delay vectors $\mathbf{Y}_M(i)$ as demonstrated in Algorithm 24.
4. Calculate the angular distance based on cosine similarity for all pairwise composite delay vectors, $Y_M(i)$ & $Y_M(j)$, as

$$d_M(i,j) = \frac{1}{\pi} \cos^{-1}\left(\frac{Y_M(i) \cdot Y_M(j)}{|Y_M(i)||Y_M(j)|}\right), \quad i \neq j$$

5. Compute the number of similar patterns defined as similar pair, $B_M^r(i)$, that satisfy the criterion $d_M(i,j) \leq r_c$.
6. Compute the estimated local probability of $B_M^{r_c}(i)$ by $C_M^{r_c}(i) = \frac{B_M^{r_c}(i)}{N-n-1}$, where $n = max(M) * max(L)$.
7. Compute the estimated global probability of $B_M^{r_c}(i)$ as $\Phi_M^{r_c} = \frac{\sum_{i=1}^{N-n} C_M^{r_c}(i)}{N-n}$.
8. MMCSE is defined as

$$MMCSE(M, L, r_c, N) = -[\Phi_M^{r_c} \log_2 \Phi_M^{r_c} + (1 - \Phi_M^{r_c}) \log_2 (1 - \Phi_M^{r_c})].$$

5.6 Value of Parameters

As pointed out in the MMCSE algorithm, several parameters need to be manually selected. This section provides a comprehensive discussion regarding the choice of the parameters, including the tolerance, r, embedding dimension, m, and data length, N.

The corresponding performances of trivariate MMSE and MMFE are examined against MMCSE. Throughout the analysis, the time delay was set to $L = 1$ for all simulations to avoid unknown influence and also make the temporal span fully controlled by the modification of the embedding dimension. Unless otherwise stated, the default values of parameters for MMSE and MMFE are set to $r = 0.15 * tri(S)$, $M = 2$ and $N = 10{,}000$, where $tri(S)$ refers to the total variance of the covariance matrix, S. The mean complexity curves in each subsection are plotted with error bars representing the standard deviation. Both the mean and standard deviation were calculated over outcomes of 10 independent realizations for each system.

5.6.1 Tolerance, r

Unlike MMSE and MMFE, the selection of tolerance, r, in MMCSE is independent of the variance of the input data sets. However, due to multiple channels, the influence of the number of variate needs to be considered when selecting the value of tolerance in MMCSE. To explore the relationship of the number of channels, p, and the tolerance, r, multi-channel white Gaussian noise (WGN) was utilized as the reference signal due to its complete randomness and simplest structure. The performances of multivariate entropy were first evaluated by varying tolerance, and then the mathematical relation was given by the best fit of the curve between r and p.

The left panel of Figure 5.2a shows the outcomes of multivariate single-scale CSE based on WGN as a function of tolerance for a varying number of input channels, from a univariate to 16-variate case. The tolerance was initially set to vary from 0.01 to 1 at intervals of 0.05 and was then linearly interpolated using values ranging from 0 to 1, at 0.01 intervals. Observe, in all curves, a rise in entropy with tolerance from 0 to 0.5 and a decrease in entropy with tolerance from 0.5 to 1. Due to this symmetry, only the range of 0 to 0.5 was considered to give the common relationship that the larger the tolerance, the more similar patterns are found shown as higher entropy. As discussed in [15], the tolerance for univariate MCSE was empirically set to $r = 0.07$, whereby the performance of MCSE applied to WGN resulted in the entropy value of 0.365. Therefore, the relationship between the number of variates and tolerance, r, was estimated by setting the entropy as 0.365 as a standard (the black dashed line shown in the left panel in Figure 5.2a), as revealed in the right panel of Figure 5.2b using piece-wise cubic interpolation fitting curve. It is found that the best fit was for

$$r_{cse} = -0.4(p^{-0.71}) + 0.47, \tag{5.1}$$

where p donates the number of variates. Therefore, the tolerance in MMCSE is decided

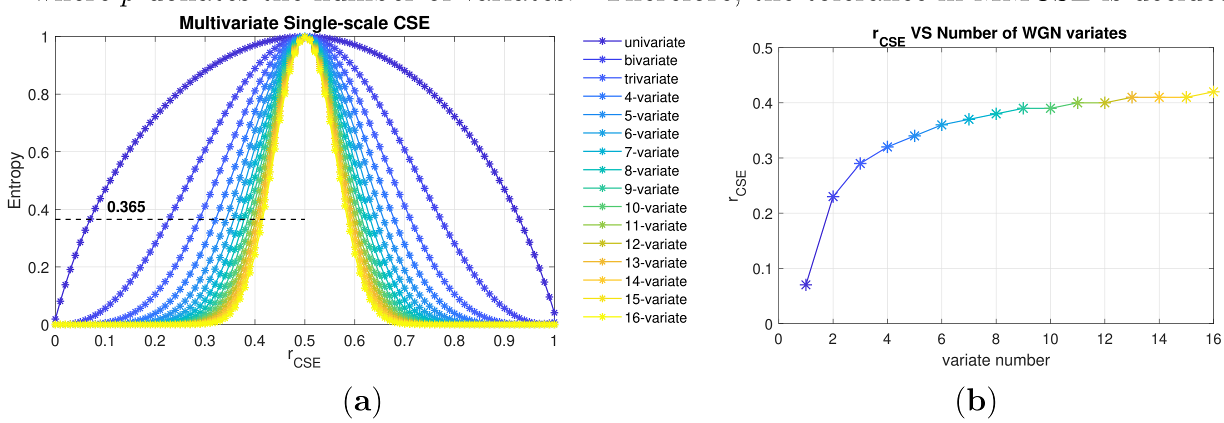


Figure 5.2: Choice of parameters in MCSE. (**a**) Behaviour of multivariate single-scale CSE on the estimation of white Gaussian noise (WGN), as a function of the tolerance, r. The default parameters were set as N = 10,000, M = 2, and L = 1 for all channels. The error bars designate the standard deviation over 10 realizations. (**b**) The optimal tolerance, r, as a function of the number of variates.

as a function of the number of variates. This relationship is given in Equation (5.1) to maintain the performance of multivariate CSE for a varying channel number. For the commonly applied situation when the channel number is 2 (bivariate data), the typical setting value of tolerance is $r = 0.225$ and for the trivariate case, it is $r = 0.287$, according to the equation.

5.6.2 Embedding Dimension, m, and Data Length, N

The embedding dimension, m, in amplitude-distance-based methods, is usually restricted by the data length, N. As a rule of thumb, N should scale as 10^m [23,174]. For FuzzyEn, the limitation is relaxed such that N should exceed 50 samples to give a defined estimate. That is due to the continuous boundary of the transfer function between similarity and tolerance at the expense of computational load [112]. As a result, limited by the sample size and computational time, the structure of complexity in high embedding dimensions is usually poorly estimated by current entropy methods. The trade-off between the embedding dimension and sample length means that the entropy methods are difficult to implement in real time. The influences of the embedding dimension, m, and the data length, N, of these multivariate entropy methods are discussed here.

The performance of the three multivariate entropy algorithms, MMSE, MMFE, and MMCSE, was tested on white Gaussian noise (WGN) and autoregressive (AR) models AR(1), AR(2), and AR(3) in three-channel multivariate systems. A total of 10 independent realizations were generated for each model with a constant sample size of $N = 2000$ points for all operations, and the results are shown in Figure 5.3. The mean entropies are shown, and error bars designated the standard deviation. Coefficients of the considered AR models are given in Table 5.1.

Table 5.1: Coefficients of the AR models used.

Coefficients	$\boldsymbol{a_1}$	$\boldsymbol{a_2}$	$\boldsymbol{a_3}$
AR (1)	0.9	–	–
AR (2)	0.7	0.25	–
AR (3)	0.6	0.25	0.125

The mean values of entropy are shown as a function of the embedding dimension, m, in Figure 5.3. The given embedding dimension ranges from 1 to 10 with increments of 1 for MMSE and MMFE, and the range of used embedding dimension for MMCSE was set to 2 to 10, due to the minimum requirement of 2 when calculating angular distance. Observe that in the left panel of Figure 5.3, the selection of m for MMSE is limited to lower than 7 to give a defined estimation for the most complex signal AR(3). On the other hand, for uncorrelated WGN, the maximum embedding dimension with a defined positive MMSE value is $m = 2$. In the second panel, MMFE produces a more stable estimation at higher scales, in contrast to MMSE. Yet, the differences between the four synthetic signals decrease with increasing the embedding dimension, and when m comes to 9, the complexity estimates overlap. In addition, resulting from the single-scale analysis, both MMSE and MMFE fail to order the complexities of the four simulated signals. The highest complexity among the given models ought to be assigned to AR(3). In comparison, the MMCSE in the last panel gives a positive, defined estimate for all the scenarios with the correct rank based on the structural complexity of the signals, where complexity(AR(3)) > complexity(AR(2)) > complexity(AR(1)) > complexity(WGN). Moreover, MMCSE illustrates a consistent estimation with a good separation among the four models. Therefore, MMCSE exhibits a more stable and consistent performance when applying complexity evaluation under high embedding dimensions in these simulations.

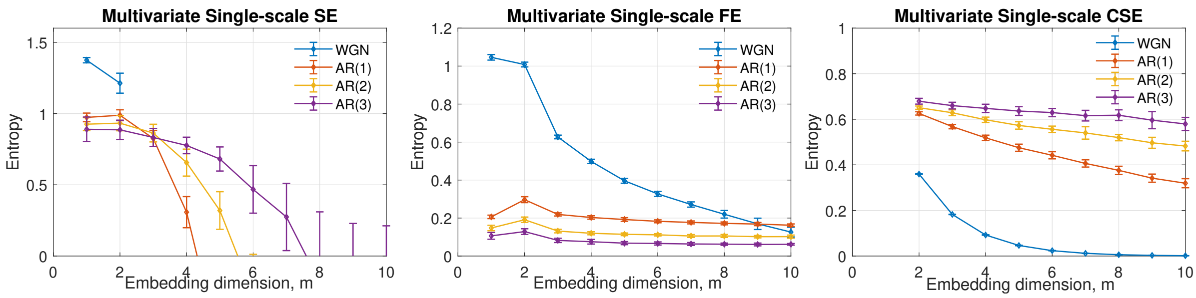


Figure 5.3: behaviour of multivariate single-scale SE, FE and CSE as a function of the embedding dimension, m. The default parameters were set as N = 2000, L = [1,1,1], r = 0.45 for multivariate SE/FE, and r = 0.287 for multivariate CSE.

The second significant parameter that controls the performance of the complexity estimation based on entropy is the measured data length, N, which is difficult to control for many real-world signals. Figure 5.4 exhibits the performance of the three single-scale multivariate entropy methods. The default parameters were set to M = [2,2,2], L = [1,1,1], and the tolerance values were selected following the Equation (5.1). In the left panel in Figure 5.4, multivariate SampEn performs poorly when $N < 60$ on the most complex signal here, AR(3). Generally it gives wide error bars with little separation between the four models. Multivariate FuzzyEn and multivariate CSE exhibit a more stable estimate and a less strict requirement for the minimal data length compared to that given by multivariate SampEn. However, in the middle panel and left panel, the most structurally complex model, AR(3), is wrongly assigned the lowest degree of complexity. The improvement of FuzzyEn from SampEn comes at the expense of calculation efficiency, with more time consumption required, as presented in the algorithms. The CSE-based algorithm manages to reduce the minimal data length required without increasing the computational load.

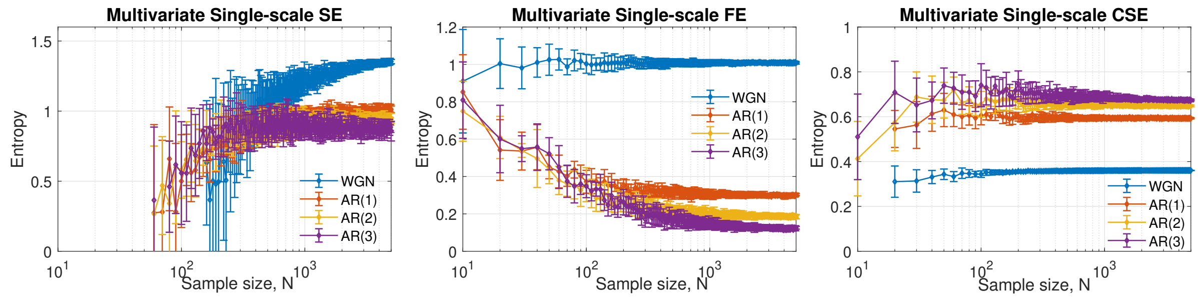


Figure 5.4: behaviour of multivariate single-scale SE, FE and CSE as a function of the data length, N. The default parameters were set as M = [2,2,2], L = [1,1,1], r = 0.45 for multivariate SE/multivariate FE, and r = 0.287 for multivariate CSE.

5.6.3 Complexity Profile of MMCSE, MMSE and MMFE

The complexity profiles of multivariate multiscale entropy algorithms were first calculated for simulated linear and nonlinear signals. The same auto-regressive (AR) models were applied together with a commonly encountered nonlinear signal, the 1/f noise. With the default parameters set to M = [2,2,2] and N = 10,000, the mean complexity of 10 independent realizations with three channels are illustrated in Figure 5.5. Error bars represent the standard deviation over the 10 realizations.

From the left to the right panel in Figure 5.5, the complexity profiles of the MMSE, MMFE and MMCSE are presented. With the same setting, MMSE is observed to be capable of distinguishing signals with self-correlations, including AR models and 1/f noise, from the completely uncorrelated WGN (in blue line). However, the random WGN is wrongly assigned to exhibit a structural complexity at low scales, and the separation among correlated signal is also difficult to be achieved as overlapped error bars shown across all the scales by MMSE. In contrast with MMSE, MMFE correctly demonstrates the complexity difference of the five models (AR models from WGN and 1/f noise from

WGN respectively) with a narrower variance.

Observe from the middle panel in Figure 5.5, that the 1/f noise is expected to be a truly complex signal across all scales, which is correctly given by MMFE. However, the problem of the wrong estimation of the complexity of uncorrelated WGN still remains by applying MMFE at small scales. When estimated by MMCSE in the right panel in Figure 5.5, the complexity measures based on WGN remain the lowest across all scales, reflecting the uncorrelated structure of an ideal random signal. In regard to the AR processes with different orders, a system with the higher order means more coefficients involved when generating signals associated with a higher degree of freedom. Hence, the correct relation of structural complexity of simulated AR processes is supposed to be AR(3) > AR(2) > AR(1), which can be only observed in the graph given by MMCSE. Meanwhile, the estimation of the complexity given by MMCSE is consistent in the comparison between AR processes with various orders, and between 1/f noise and WGN based on self-correlation. To this end, the proposed MMCSE yields stable estimates at a high coarse-graining scale, thus making it possible to examine structural complexity of real world processes with long-range correlations.

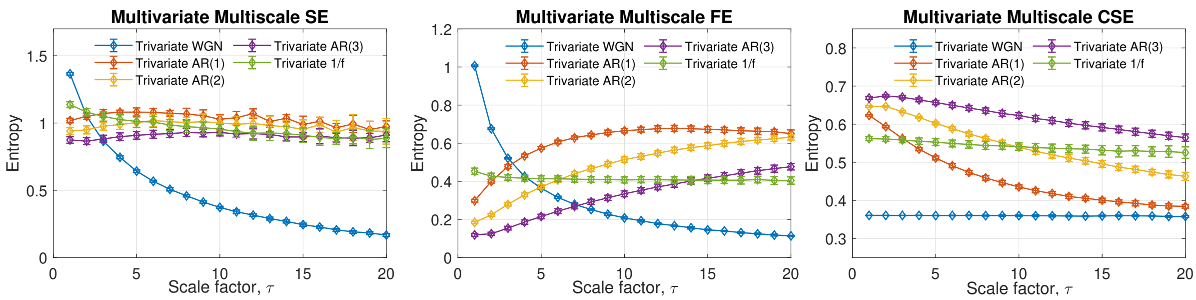


Figure 5.5: Complexity profile of MMSE, MMFE, and MMCSE. The default parameters were set as M = [2,2,2], N = 10,000, L = [1,1,1], r = 0.45 for MMSE/MMFE, and r = 0.287 for MMCSE.

5.7 Detection of Circularity

We next examine the application of multivariate entropy in the detection of complex and quaternion non-circularity, that is, the rotation dependence of probability density function (pdf). The non-circularity manifests itself through a degree of correlation and power imbalance of the system channels, which can be reflected in a scatter plot as the inclined angle and dispersed degree of the distribution. The long-range correlation of a time series measured by multiscale entropy reflects the influence of past states on the generation of future signals [46], while the correlation of a multi-channel system is a measure of association among channels, a summary of dependence strength for multivariate systems, and a feature that could reflect the structural richness of the system [19, 175]. A strongly correlated system is expected to show a high structural complexity, while a less correlated system would approach the performance of WGN to give a low structural complexity. As for the other critical property, the power of the signal has been widely used as an indicator of dynamics in complex systems, such as the analysis based on heart rate variability in physical systems [51], where the ratio of the low-band frequency power and high-band frequency power was used to reveal the sympatho-vagal balance of the autonomous nervous system [3].

5.7.1 Examination of Non-Circularity of Bivariate Systems

To examine the ability of MMCSE simulations on bivariate (complex-valued) data, were classified into three groups: (a) correlated WGN with equal power, (b) uncorrelated WGN with unequal power and (c) correlated WGN with unequal power. The dual-channel inputs, $\mathbf{Y}$, were generated according to Equation (5.2), where $\mathbf{X}$ donates the bivariate uncorrelated WGN system, $\mathbf{X} = [x_1, x_2] = [\varepsilon_1, \varepsilon_2]$, $\varepsilon(i) \sim \mathcal{N}(0, 1)$, and $Chol(\mathbf{C})$ refers to the Cholesky decomposition based on the correlation matrix, $\mathbf{C}$. Thus, the dual channel of the generated system, $\mathbf{Y} = [y_1, y_2]$, becomes $y_1 = x_1$ and $y_2 = q(px_1 + (\sqrt{1-p^2})x_2)$. We modified the value of coefficients, p and q, to control the behaviours of the system correlation during the examination, that is

$$\mathbf{Y} = \mathbf{X} \cdot chol(\mathbf{C}), \quad \mathbf{C} = \begin{bmatrix} 1 & pq \\ pq & q^2 \end{bmatrix} \tag{5.2}$$

For each setting, 30 random realizations were generated with sample size $N = 1000$, to give the performance of estimation shown as error bar plots based on the standard deviation. The default parameters of multivariate multiscale entropy were set to $M = [2, 2]$, $L = [1, 1]$, $r_{se} = r_{fe} = 0.15 * tri(S)$, and $r_{cse} = 0.225$ according to Equation (5.1). The synthetic data generations and simulations were implemented in MATLAB, where the graphs based on circularity statistics were plotted by using online MATLAB toolbox referred in [176].

i) Correlated WGN with Equal Power

We considered five systems with equal power, $q = 1$, but different correlation levels, where $p = [0, 0.6, 0.8, 0.9, 0.99]$ to give different degrees of complex non-circularity. To illustrate the behaviours of the synthetic systems, angular densities plotted on the unit circle are shown in the first row of Figure 5.6 as well as angular histograms are in the second row. The number in blue in each angular histogram designates the frequency density of the largest bin in the graph. The system in the first column represents the situation of a random (circular) system, where two channels are independent of each other as the points on the unit circle is loosely distributed with flat histogram in Figure 5.6. With fixed power, the correlation increases from the left to the right system as the density gradually concentrates on specific non-zero angles, $\pi/4$ and $-3\pi/4$. Therefore, the structural complexity is expected to increase as p becomes larger, and the bivariate distribution more non-circular.

Scatter plots are plotted in the first row of Figure 5.7 to give the space distribution (non-circularity of pdf) of each bivariate system. The second row of Figure 5.7 exhibits the complexity estimation given by MMSE, MMFE and MMCSE accordingly. Complexity profiles are shown as entropy in a function of scale factor. Observe that MMSE and MMCSE (in blue and black) present ascending levels with increasing p, while a limited difference can be seen based on MMFE with varying correlations (in red). Although both MMSE and MMCSE are capable of demonstrating the different structural

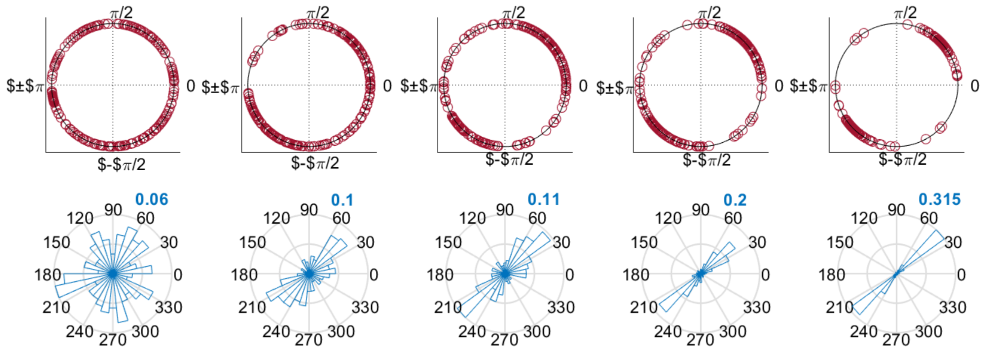


Figure 5.6: Angular density shown as points on the unit circle and angular histogram of bivariate (complex valued) input correlated WGN with equal powers in data channels.

richness, MMSE drops down with increasing scale, while MMCSE exhibits a consistent estimation across all scales.

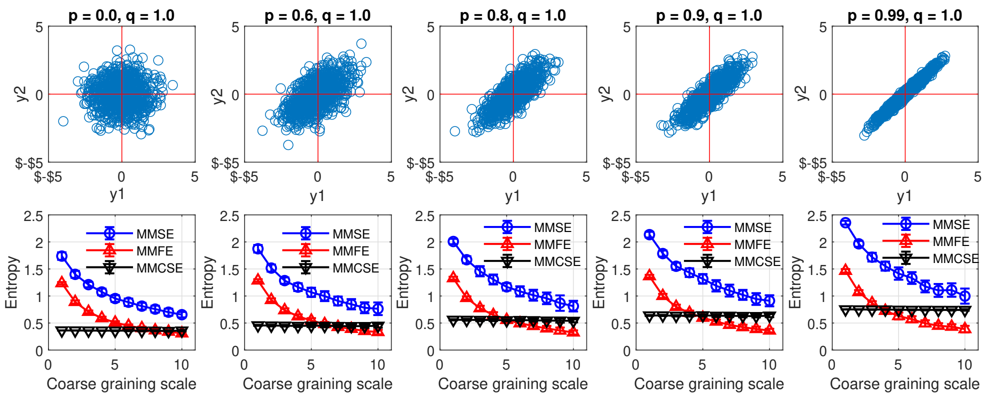


Figure 5.7: Scatter diagrams and entropy-based complexity (non-circularity) estimations of bivariate input (complex valued) correlated WGN with equal powers in data channels.

ii) Uncorrelated WGN with Unequal Power

We considered five uncorrelated dual-channel (complex valued) WGN systems, $p = 0$, with different power with $q = [1, 0.4, 0.3, 0.2, 0.1]$. The density and histogram are shown in Figure 5.8, while the corresponding scatter plots are given in Figure 5.9. The data points are gathered around horizontal line and the angular density approaches zero as power decreases. Therefore, according to the space distribution and angular density plots, the structure richness of the system (degree of non-circularity) is ascending as q decreases.

The estimates of multivariate multiscale entropy are accordingly shown in Figure 5.9

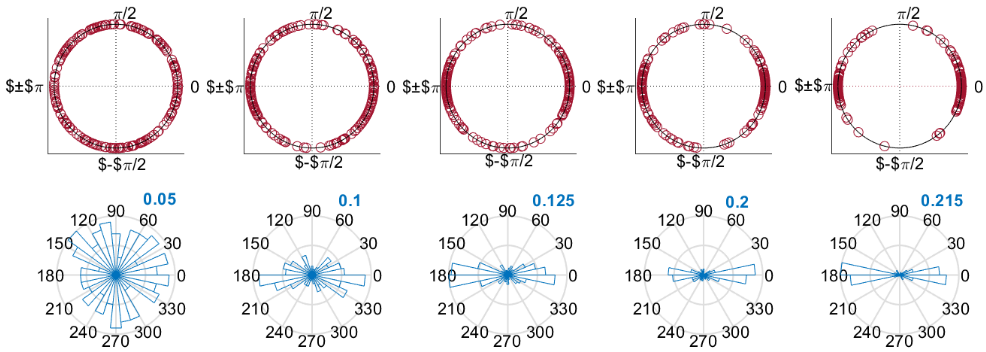


Figure 5.8: Angular density shown as points on the unit circle and angular histogram of bivariate (complex valued) input uncorrelated WGN with unequal powers in data channels.

below the scatter plot of each system. Both MMSE and MMFE (in blue and red) failed to reflect the change of power with identical complexity and tendency. In contrast, MMCSE (in black) exhibited an increasing structural richness from the left to the right, thus correctly reflecting the change in the degree of non-circularity.

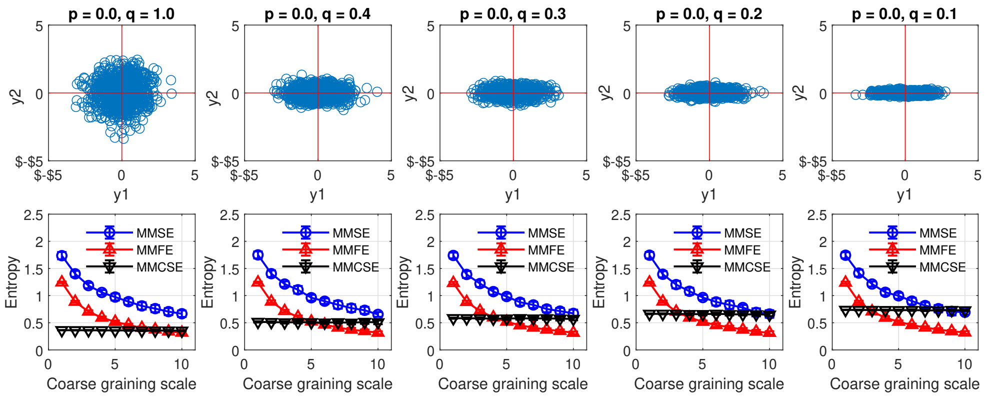


Figure 5.9: Scatter diagrams and entropy-based complexity estimations of bivariate (complex valued) input uncorrelated WGN with unequal powers in data channels.

iii) Correlated WGN with Unequal Power

Next, we applied the entropy-based estimation on four correlated WGN systems, $p = 0.6$, with unequal power, $q = [1, 0.5, 0.3, 0.1]$ and uncorrelated random system was included as a reference, $p = 0$ and $q = 1$. Figure 5.10 shows that, despite the fixed correlation coefficient, the decreasing power pulls the angular density back to the horizontal level. Therefore, the structural richness enhances either with an increasing correlation, p, or decreasing power, q, which is also reflected in the scatter plots in Figure 5.11.

Figure 5.11 plots the results of entropy in the second row. Observe that MMSE and MMFE were able to separate completely random systems in the leftmost panel from the rest of the correlated systems. However, due to the need for normalization, the power difference could not be detected by the complexity estimation given by either MMSE or MMFE. In contrast, the proposed MMCSE managed to evaluate the structural change of the system either caused by the modification of power or the behaviour of correlation.

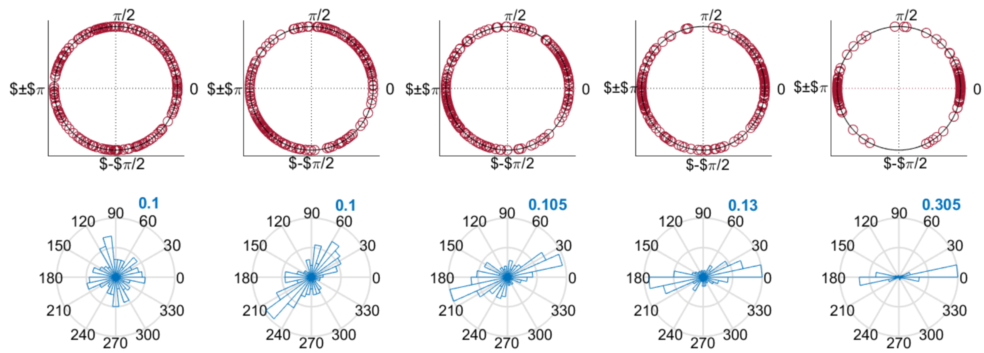


Figure 5.10: Angular density shown as points on the unit circle and angular histogram of bivariate (complex valued) input correlated WGN with unequal powers in data channels.

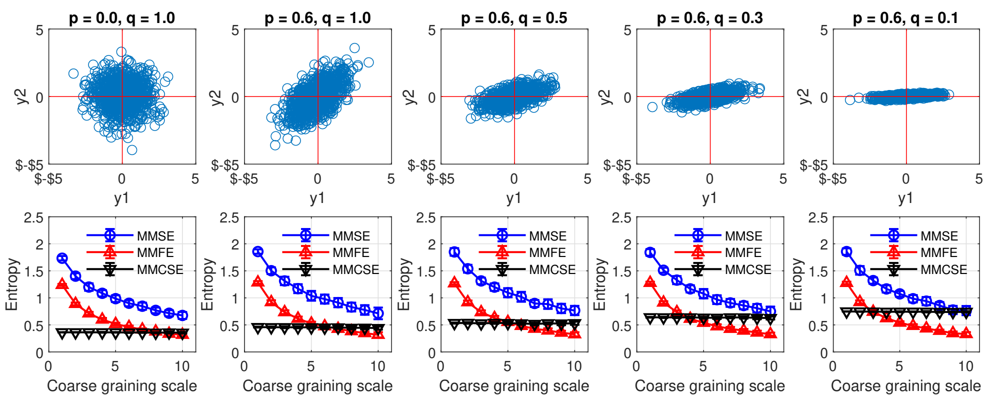


Figure 5.11: Scatter diagrams and entropy-based complexity estimations of bivariate (complex valued) input correlated WGN with unequal powers in data channels.

5.7.2 Examination of Non-Circularity of Trivariate Systems

The ability of standard MMSE, MMFE and the proposed MMCSE in the detection of circularity of multivariate signals is next simulated and discussed in trivariate systems. Systems constructed by trivariate correlated WGN, $\mathbf{Y}$, are discussed here and generated by the correlation matrix, $\mathbf{C}$, shown in Equations (5.3) and (5.4). The correlation behaviour is jointly controlled by the coefficients, p and q, while the correlated power is determined by the factor, q, that is

$$\mathbf{Y} = \mathbf{X} \cdot chol(\mathbf{C}) \tag{5.3}$$

$$\mathbf{C} = \begin{bmatrix} 1 & pq & pq \\ pq & q^2 & pq^2 \\ pq & pq^2 & q^2 \end{bmatrix}, \quad chol(\mathbf{C}) = \begin{bmatrix} 1 & pq & pq \\ 0 & q\sqrt{1-p^2} & \frac{pq\sqrt{1-p}}{\sqrt{1+p}} \\ 0 & 0 & q\sqrt{1-\frac{2p^2}{1+p}} \end{bmatrix} \tag{5.4}$$

Here, $\mathbf{X}$ denotes a trivariate uncorrelated WGN system, $\mathbf{X} = [\varepsilon_1, \varepsilon_2, \varepsilon_3]$, $\varepsilon(i) \sim \mathcal{N}(0, 1)$, and $Chol(\mathbf{C})$ refers to the Cholesky decomposition based on the correlation matrix, $\mathbf{C}$.

The trivariate systems with varying correlation degree, p, and power, q, were analysed by MMSE, MMFE, and the proposed MMCSE. The default parameters were set to $M = [2,2,2]$, $L = [1,1,1]$, $N = 1000$, and the maximum scale factor was chosen as $\tau = 10$. The tolerance, r, for MMSE and MMFE was set as $0.15 * tr(S)$, where $tr(S)$ refers to the total variance of the input signals. The choice of tolerance for MMCSE was calculated according to Equation (5.1). The error bars designate the average and standard deviation over 30 realizations for each model.

i) Correlated WGN with Equal Power

Trivariate input signals associated with five different correlation degrees and equal power were first generated as scatter plots shown in Figure 5.12. The first model in the left graph with $p = 0$, $q = 1$, gave the uncorrelated WGN system. The four considered cases were set as the fixed power, $q = 1$, and the correlated relations among input channels, $\mathbf{Y}$, were hence determined by one varying coefficient, $p \in \{0.4, 0.6, 0.75, 0.99\}$. Observe the varying degree of non-circularity, from the rotation-invariant (circular) case on the left panel through to a high degree of non-circularity (narrow scatter plot) on the right-most panel.

Figure 5.13 illustrates the performance of the multivariate multiscale analysis based on the SampEn, FuzzyEn and CSE on the task of the detection of the degree of non-circularity in trivariate (quaternion) data. Observe that MMSE and MMCSE, in the left and right panels respectively, were capable of distinguishing between the varying settings of coefficients, p, shown as different trends of lines, while MMFE in the middle panel was only prone to separating the case with maximum correlation (in amber) from the rest of the scenarios. As p increases, the structural richness of the system (degree of non-circularity) rises, as shown in Figure 5.12. Hence, the structural complexity for signals with large p is supposed to be greater than that of signals generated from small p, which is correctly evaluated by both MMSE and MMCSE. However, as scale, τ, increases, MMSE exhibits a descending trend and the gaps among curves narrow down accordingly, whereby the ability of MMSE to classify varying correlations is negatively influenced by the increasing scale. On the other hand, for MMCSE in the right panel, the increasing scale has no impact on the separation of different models as shown by the flat lines.

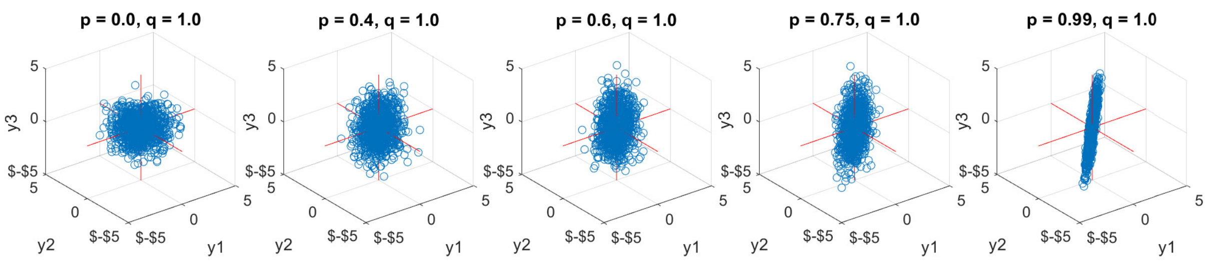


Figure 5.12: Scatter diagrams of trivariate (quaternion) input correlated WGN (pure quaternion) with equal powers in the three data channels, but different degrees of correlation among them. Observe an increase in the degree of non-circularity from left to right.

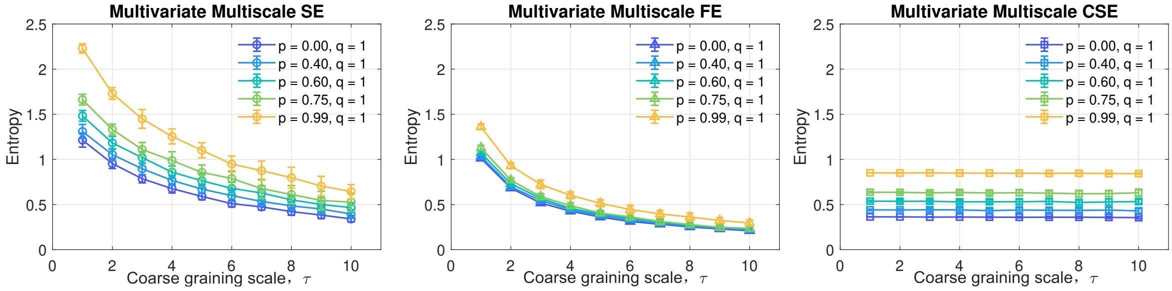


Figure 5.13: behaviour of standard MMSE, MMFE, and the proposed MMCSE applied to correlated trivariate (quaternion) WGN with equal power.

ii) Uncorrelated WGN with Unequal Power

Next, five models of trivariate (pure quaternion) uncorrelated WGN with unequal power were simulated and analysed by multivariate entropy. Figure 5.14 demonstrates the scatter diagrams of five uncorrelated input systems, where the dynamics in one channel is independent of the change of others. The coefficient, p, was fixed as zero (to make the $chol(\mathbf{C})$ a diagonal matrix) and q was varying in this case among $\{1, 0.6, 0.45, 0.35, 0.1\}$ to control the power of the system. As shown, the above two channels exhibit no correlation, with flat scatter diagrams and the power imbalance increase from the left to the right when the power reaches its maximum and is fully balanced, which is the case of white noise as given in the first graph. Figure 5.15 shows the performance of the three multivariate entropy methods.

The left and middle panels of Figure 5.15 give the estimation by MMSE and MMFE, which fail to reflect the effects of different powers of the signals due to data normalization, as given in the first step of amplitude distance-based algorithms shown in Algorithm 24. In other words, the information about power imbalance is lost when normalizing the data for SampEn and FuzzyEn-based analyses, thus resulting in a decrease in entropy values approaching the behaviours of random white noise. However, data normalization is one of the key pre-processing steps for amplitude-based entropy analyses, which cannot be avoided. In contrast, the MMCSE in the right panel is observed to detect the discrepancy of powers among the data channels of the uncorrelated systems, resulting from non-unified variance towards the input multivariate data sets based on the angular distance.

In addition, the non-descending trends in the right panel remain constant, irrespective of the varying power, where the expected performance of structural complexity estimation can be observed at single scale by multivariate CSE.

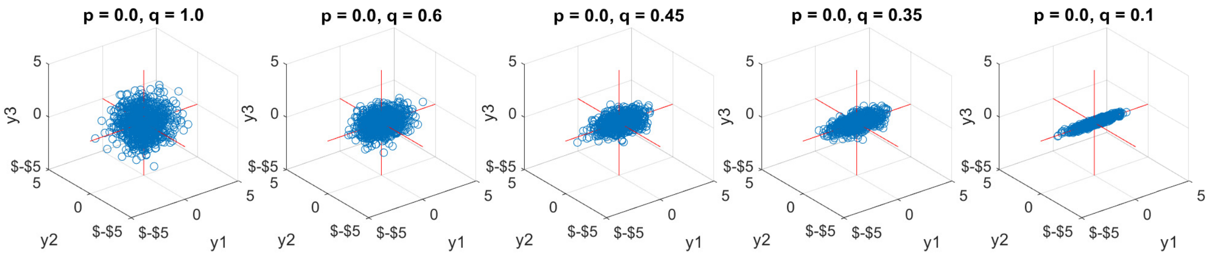


Figure 5.14: Scatter diagrams of trivariate (pure quaternion) input uncorrelated WGN with unequal powers in data channels. Observe the increase in the degree of non-circularity, from the left to the right panel, with the increase in power imbalance between the data channels.

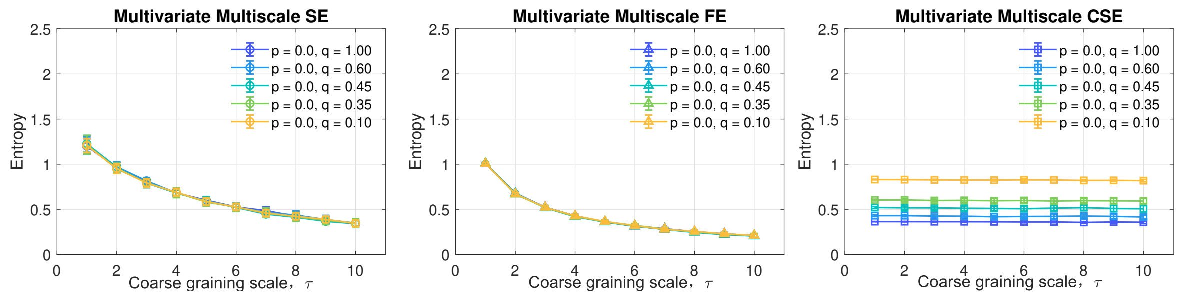


Figure 5.15: behaviour of standard MMSE, MMFE, and the proposed MMCSE applied to trivariate (quaternion) uncorrelated WGN with unequal powers in data channels.

iii) Correlated WGN with unequal power

Next, we jointly considered the two above-discussed scenarios for systems of trivariate (pure quaternion) correlated WGN associated with unequal powers in data channels. In terms of the correlated input, the scatter diagrams are illustrated in Figure 5.16, where q is varying within $\{1, 0.5, 0.35, 0.1\}$, given p a non-zero value as $p = 0.6$. For comparison, the uncorrelated WGN, $p = 0$ & $q = 1$, was also produced as the model with least structural complexity (circular) as shown in the first diagram in Figure 5.16. Figure 5.17 demonstrates the complexity estimation based on the models generated.

With a fixed, non-zero coefficient p, the increasing power imbalance (non-circularity) yields a higher structural complexity with stronger correlation as shown in the scatter

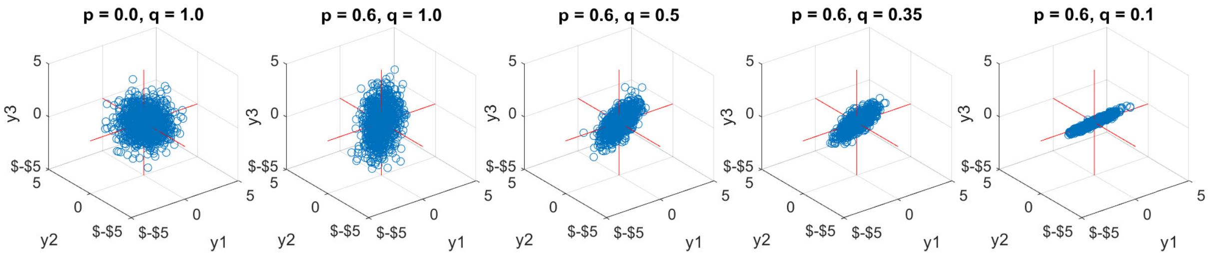


Figure 5.16: Scatterdiagrams of trivariate (quaternion) input correlated WGN with unequal power.

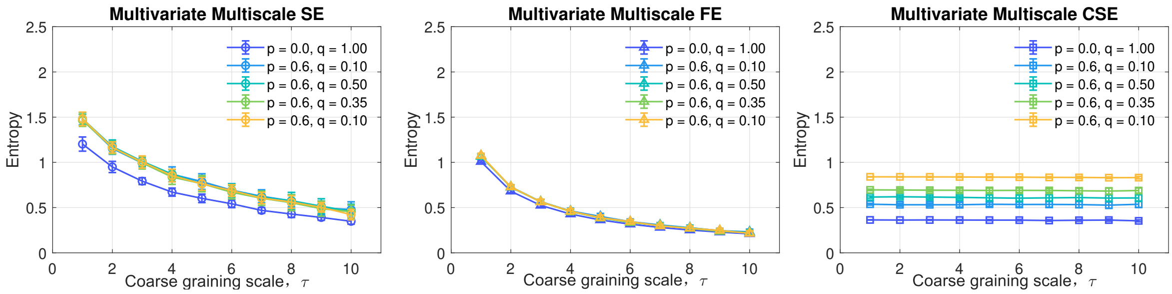


Figure 5.17: behaviour of standard MMSE, MMFE, and the proposed MMCSE applied to correlated trivariate (quaternion) WGN with unequal power.

plots from the left to the right in Figure 5.16. Regarding the complexity measures, observe that MMSE in the left panel of Figure 5.17 could only separate the completely uncorrelated signals (in blue) from the correlated signals, while MMFE fails to detect the significant discrepancy of structural complexity within all the simulated models. In terms of the proposed MMCSE, signals with various powers and correlation structures (different degrees of non-circularity) are separated in a consistent way in the right panel in Figure 5.17. To further explore the ability of complexity estimation for signals with jointly varying coefficients p and q, the heatmaps of entropy estimated via multivariate single-scale SampEn, multivariate single-scale FuzzyEn and multivariate single-scale CSE are displayed in Figure 5.18, where p in the y-axis was set to range from 0.05 to 0.99 at 0.05 intervals, and power, q, in the x-axis ranged from 0.05 to 1 at 0.05 intervals. Observe that SampEn- and FuzzyEn-based multivariate analyses are sensitive to the correlation coefficient p, but are insensitive to signal power, q. Comparing the two amplitude entropy methods, multivariate SE exhibits higher sensitivity to changes of p than multivariate FE. In contrast, the proposed CSE-based multivariate method manages to detect changes in both p and q as shown in the right heatmap in Figure 5.18.

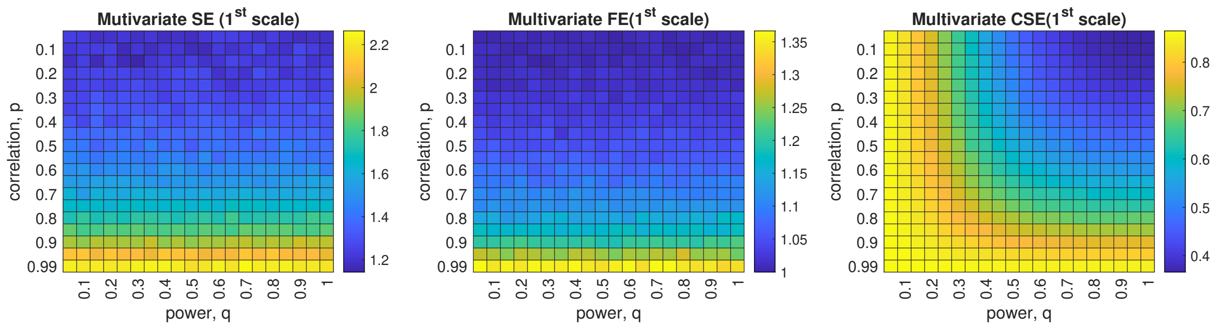


Figure 5.18: Heatmap of complexity estimation based on standard multivariate SE, multivariate FE, and the proposed multivariate CSE in a function of varying coefficients p and q.

5.8 Conclusions

This work extended the univariate cosine similarity entropy to the multivariate case. It was shown that the proposed Multivariate Multiscale Cosine Similarity Entropy method is capable of estimating structural complexity based on self-correlation with higher stability at large scales and lower requirement on data length than the existing methods. In contrast with standard MMSE and MMFE, the MMCSE has also exhibited higher consistency in multiscale analysis, where the long-range correlation of signals can be correctly measured at low scales. The performance of MMCSE was first examined on five benchmark signals to reveal the improved estimation against MMSE and MMFE, and then tested on multivariate correlated systems associated with unequal power among data channels (complex and quaternion non-circularity), to detect the properties of circularity including the degree of correlation among channels and the strength of power in the system. As desired, due to the angle-based distance and the relaxation of the requirement of data normalization, the proposed MMCSE exhibited noticeable improvement in detecting the correlation degree of the multivariate system and also showed the unique feature to detect the change of power in the system, unlike MMSE and MMFE. The results indicated the wide range of potential applications of MMCSE in complexity science for practical scenarios. However, regarding the limitation of the coarse-graining average, the multiscale procedure can be further improved by employing the enhanced scaling process. It should be noted that this study is restricted to the analysis on synthetic signals. Therefore, future work will focus on the analysis of real-world data sets and involve statistical methodologies.

Chapter 6

Complexity-based Financial Stress Evaluation

6.1 Chapter Summary

Financial markets typically exhibit dynamically complex properties as they undergo continuous interactions with economic and environmental factors. For example, the Efficient Market Hypothesis (EMH) suggests a considerable difference in the structural complexity of security prices between normal (stable markets) and abnormal (financial crises) situations. Considering the analogy between market undulation of price time series and physiological stress of bio-signals, we investigate whether physical stress indices in bio-systems can be adopted and modified so as to measure 'standard stress' in financial markets. This is achieved by employing structural complexity analysis, based on variants of univariate and Multivariate Sample Entropy, to estimate the stress level of both financial markets on the whole and the performance of the individual financial indices. Further, we propose a novel graphical framework to establish the sensitivity of individual assets and stock markets to financial crises. This is performed through Catastrophe Theory and entropy-based stress evaluations, indicating the unique performance of each index/individual stock in response to different types of crisis. Four major indices and four individual equities with gold prices are considered over 31 years, from 1991-2021. Our findings both demonstrate the feasibility of measuring financial stress and reveal the structural complexity relations among economic indices and within each price time series.

6.2 Introduction

Over the last several decades, structural Complexity characteristics of stock markets have been investigated owing to their potential efficient indicators associated with financial health and economic stabilization [177, 178]. The cornerstone of modern financial theory, the Efficient Market Hypothesis (EMH), states that the underlying value of an asset incorporates all the available information to make the stock always trade at a fair value [179]. This implies that in 'normal' situations, markets are barely influenced by any specific event and evenly respond to the continuous stimuli of economic change. In the context of complexity science, the security prices in 'normal' situations exhibit high randomness and uncertainty, together with low predictability from their historical values. On the other hand, when an economic crisis occurs, corresponding to an 'abnormal' situation, markets are easily driven by irrational fluctuations (e.g. panic buying/selling), showing a lower degree of randomness and higher determinism [53, 180].

Analogous to the well-known automatic 'fight-or-flight' response in physiological stress studies from human cognitive science [181], financial stress can also be considered as a deviation from the normal functioning of financial markets [54]. Considering the similarity between the Complexity Loss Theory (CLT) in human body functions [45] and the implications of EMH in financial markets, the concept of sympatho-vagal balance in bio-systems can also be used to describe the 'acceleration-stabilization' type of behaviour in financial systems [3, 51]. In human-centred sciences, the sympatho-vagal balance refers to the joint influence of the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) [51], which respectively accelerates and decelerates body functions. Regarding the analysis of financial systems, the acceleration-stabilization' behaviour is driven and sustained by supply and demand, which results in market expansions and recessions [180].

By virtue of the analogy between financial systems and bio-systems, the non-parametric analyses, that have been applied in human-centred science to reveal complex features in the nonlinear domain, can also be employed to estimate and determine the financial stress of markets [19, 56, 137, 182]. Nonlinear methods utilized in the following analyses include entropy-based univariate/multivariate algorithms. Examples include the Recurrence Quantification Analysis (RQA) [183] and Assessment of Latent Index of Stress using instantaneous amplitude (iA-ALIS) methods [3]. Traditional entropy methods aiming at signal irregularity quantification have been applied in financial investigations due to their model-independent and prejudice-free analysis, including Approximate Entropy [19, 182], Sample Entropy [184], Permutation Entropy [84, 185] and other recently developed algorithms of a kind [186–188]. In addition to univariate entropy analysis, enhanced multivariate methodologies have been initially involved such as Multivariate Multiscale Sample Entropy [2]. Recurrence Quantification Analysis (RQA) is another popular nonlinear measure method in quantifying determinism via predictability, which has been widely considered in physiological studies [189, 190]. When it comes to financial systems, stock market volatility is a well-established concept referring to asset price fluctuations and reflects the degrees of uncertainty of future prices [19, 53, 182]. Closely relevant to the volatility measure, the high degree of determinism (DET) given by RQA implies small volatility and high predictability [191, 192]. Assessment of latent index of stress (ALIS) was introduced to detect financial crises primarily in stock markets by examining the power in Low-Frequency and High-Frequency bands of dynamical prices [53, 180]. Expanding on the ALIS, an assessment based on the instantaneous amplitude (iA-ALIS) was proposed in place of the power of the signal as a more reliable indicator of financial stress [3].

As stated in EMH, the asset price contains all available information, and in order to obtain the dynamics of price change, most financial studies have been implemented based on the return time series, the difference between the two consecutive prices, $x(i + 1)$ and

$x(i)$. In this work, to maintain maximal information of price time series in the signal, we first apply a Moving Average (MA) filter to detrend data. Given the potential of nonlinear methods to reveal financial stress, we employ Complexity Science (CS), that is structural complexity analysis based on historical data to predict the occurrence of an 'abnormal' situation, that is, a financial crash. Such a sudden change in the behaviour of the financial system results from the smooth changes which arise jointly from both economic and non-economic factors. This phenomenon has been conceptualized by René Thom, who termed it Catastrophe Theory [193]. While applicable to any dynamical system, the use of Catastrophe Theory in economics has been limited to a small number of studies [55, 194–196]. We therefore set out to demonstrate that the approaches of financial stress measurement offer a very insightful and practical way for quantitative analysis. To this end, we considered 31 years of historical data, from January 1991 to December 2021, with one data point per weekday. The main investigation is conducted over 4 stock indices (Dow Jones Industrial Average, NASDAQ Composite, Russell 2000 and Standard & Poor's 500) and 4 individual equities from different industries (Apple Inc., Microsoft, McDonald's and American International Group). The Catastrophe plots exhibit the unique performance of each index/individual stock in response to different crises in the last several decades.

The remainder of the chapter is organized as follows. Details of nonlinear algorithms and methodologies are presented in Section 6.3. To give an overview of the asset price time series and analytical flow, a summary of data and methods is given in Section 6.4. Section 6.5 illustrates the results of measures and analyses the nonlinear properties of financial markets. Section 6.6 proposes the framework of Catastrophe Theory in financial investigation and provides initial analyses based on Catastrophe Theory. Finally, Conclusions are given in the last section.

6.3 Algorithm and Methods

Algorithms used in this study are first introduced and described here.

6.3.1 Modified Univariate Multiscale Sample Entropy & Modified Multivariate Multiscale Sample Entropy

Moving Average Filter To obtain the scaled and detrended signal, $y^{(\tau)}(j)$, a Moving Average (MA) filter is first utilized to remove the local trend, $s^{(\tau)}(j)$, from the original time series, $\{x(i)\}_{i=1}^{N}$, as

$$s^{(\tau)}(j) = \frac{1}{\tau} \sum_{i=j-\tau/2-1}^{j+\tau/2-1} x(i), \quad 1 \leq j \leq N - \tau + 1. \tag{6.1}$$

$$y^{(\tau)}(j) = x(i) - s^{(\tau)}(j) \tag{6.2}$$

where τ is a pre-defined scale factor. Observe that, different from the traditional Coarse Graining Process [120], the scaling given by the MA filter is capable of maintaining the original signal length and is a better fit to the intrinsic properties of the applied data.

Multivariate Sample Entropy Sample Entropy is a standard approach to evaluate the irregularity and randomness degree of time series based on their temporal dynamics, which has been widely applied in real-world complex systems [26, 125]. Sample Entropy is built based on the probability of similarity between embedding (delay) vectors, where their higher similarity demonstrates the higher predictability at multiple scales. To this end, we applied the MA filter to implement the scaling process for each time series, which was termed the Modified Multiscale Entropy (Mod-MSE) [127]. Due to the intrinsic properties of financial data, the pre-defined scale factor is set as one week, $\tau = 5$. The Mod-MSE was applied on every index and stock individually to give the stress evaluation of each index and equity, thus indicating their response towards the external environment. However, any individual stock or single index is not sufficient to represent the performance of the whole financial market. To this end, the enhanced Multivariate Entropy is employed to estimate the overall stress of the financial market. The multivariate entropy method accounts for the cross-channel dependencies in multivariate data by constructing Composite Delay Vectors (CDV), $\mathbf{X}_M(i)$, derived from the original p-channel signal, as demonstrated in Algorithm 19 in Chapter 2. Based on a combination of the MA filter (as a scaling process) with Multivariate Sample Entropy, the Mod-MMSE was employed across multiple channels as a standard stress estimation method over the whole market, with the details of Mod-MSE and Mod-MMSE given in Algorithm 26.

Algorithm 26. Modified Univariate Multiscale Sample Entropy (Mod-MSE) & Modified Multivariate Multiscale Sample Entropy (Mod-MMSE)

Given a multivariate data set with P channels $\{x_{k,i}\}_{i=1}^{N}$, $1 \leq k \leq p$, of length N, or a univariate data set with $P = 1$.

1. Standardize the original data sets by subtracting the mean and dividing by the standard deviation for each channel.
2. Scale the normalized data sets, $\{y_{k,i}^{(\tau)}\}_{j=1}^{N-\tau+1}$, for each channel following on (6.1) and (6.2).
3. Apply the standard MMSE as shown in Algorithm 19 in Chapter 2.

6.3.2 Recurrence Quantification Analysis

The Recurrence Plot (RP) is a traditional methodology to identify hidden correlations in multi-dimensional spaces, without the limitation of data stationarity and size restriction [22,197]. By using Takens' embedding theorem [66], the univariate time series, $\{x(i)\}_{i=1}^{N}$, is reconstructed into a phase space according to the optimal embedding dimension, m, and time delay, l, as

$$x_m(i) = [x(i), x(i+l), \ldots, x(i+(m-1)l)] \tag{6.3}$$

The optimal combination of embedding dimension and time delay can be selected via different methods, such as False Nearest Neighbours for embedding dimension [198] and Minimum Mutual information for delay factor [199]. In this work, we choose the joint selection of optimal embedding dimension and time delay by the differential entropy-based method introduced in [174].

The outcome of RP is a matrix summarizing the distance among the Delay Vectors (DVs).

Given a threshold, ε, the RP matrix can be plotted as a grey image, where every element in the matrix is converted into a pixel colour based on the relation between ε and the distances between DVs. Several probabilistic measures can be implemented according to the RP matrix, such as Degree of Determinism (DET) and Laminarity (LAM). Both could be used as measures of the inverse of the volatility, where DET is the percentage of recurrent points forming diagonal line structures and LAM is that of forming vertical lines [191]. In financial markets, volatility is an important property that shows the implicit risk and is generally referred to as the degree of uncertainty about the future price [182], hence, it also reflects the degree of predictability. Here, we employed the DET index in RQA, where the length of a diagonal line in RP reflects the number of consecutive recurrent states. Considering the inverse relation of DET and volatility, we expect that high determinism refers to the high predictability of the future price, representing the 'abnormal' situation in the financial market [199]. To this end, the change of DET is positively related to the stress level of the estimated index. Algorithm 27 outlines the process of DET calculation.

Algorithm 27. Recurrence Quantification Analysis (RQA)

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N.

1. Construct the delay vectors (DVs), x_m, according to Takens' embedding theorem, as in Equation (6.3).
2. Generate the Recurrence Plot (RP) matrix, composed of pairwise Euclidean distances between DVs, as

$$RP(i,j) = \Theta(\varepsilon - ||x_m(i) - x_m(j)||, \; i,j = 1,\ldots,N-n-1, \; i \neq j,$$

where $||\cdot||$ designates the Euclidean distance, $\Theta(\cdot)$ refers to the Heaviside function, and ε denotes the threshold when defining the similarity between DVs, which is set as 60% of the mean Euclidean distance of the DVs, and $n = (m-1) * l$.

3. The degree of Determinism (DET) can be calculated as the percentage of recurrence points that form diagonal lines in the RP matrix, that is

$$DET = \frac{\sum_{j=j_{min}}^{N-n-1} j \cdot P(j)}{\sum_{j=1}^{N-n-1} j \cdot P(j)}$$

where $P(j)$ is the number of diagonal lines of the length j, and j_{min} is the minimal number of points to be considered as a diagonal line, which is set as $j_{min} = 2$.

6.3.3 Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS)

Assessment of Latent Index of Stress (ALIS) was especially proposed to quantify the 'stress level of a financial organism' in [53]. The original ALIS utilised the low-frequency band power and high-frequency band power of detrended price after the MA filter. By taking into account the intrinsic properties of financial data, the Low-Frequency band is considered to occupy the frequency band below 0.0042Hz ($=\frac{1}{240}$) with a time window of one year, while the High-Frequency band is set to between 0.0167Hz ($=\frac{1}{60}$) and 0.2Hz ($=\frac{1}{5}$), corresponding to 2 months and 5 days, respectively. Recently, by employing instantaneous amplitude via the Hilbert transform, ALIS has been enhanced with correct signal power estimation, as discussed in [3]. We therefore employed the Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS) based on the detrended financial time series. The higher the iA-ALIS index the more stressful the stock, with the threshold between 'stressed' and 'normal' derived based on the median value. The details of iA-ALIS are summarised in Algorithm 28.

Algorithm 28. Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS)

Given a univariate data set $\{x(i)\}_{i=1}^{N}$ of length N.

1. Remove the trend of the data by a Moving Average filter with a window of 1 year, and obtain the detrended data, $\{z(i)\}_{i=1}^{N}$.
2. Bandpass-filter the detrended data, $z(i)$, into the Low-Frequency (LF) Band and High-Frequency (HF) bands. Apply the Hilbert transform to LF and HF, and obtain the instantaneous amplitude based on the analytic signals given by the Hilbert transform at every time point, denoted as iA_{LF} and iA_{HF}.
3. Take 4 years as the window length and 1 month as an increment. In every time window, exclude the 20% largest and smallest values, to remove the outliers and then calculate the mean iA for every window, denoted as $LF(d)$ and $HF(d)$, where d refers to a month.
4. Normalise the $LF(d)$ and $HF(d)$ series by subtracting the mean and dividing by standard deviation in order to alleviate the problem of scaling. Remove the offset of $LF(d)$ and $HF(d)$.
5. The ALIS is given by $ALIS(d) = LF(d) + HF(d)$. The median value in the $ALIS(d)$ is a threshold for the stress in the market.

6.4 Data Overview and Methods Summary

We applied several methodologies to four groups of indices/stocks over the last 31 years, spanning the period between 1991.01.01 and 2021.12.31. These are:

- Stock market index
  - Dow Jones Industrial Average (DJIA/DOW): 30 large companies;
  - NASDAQ Composite (NAS): Mid- and large-caps;
  - Russell 2000 (RUS): Smaller companies;
  - Standard & Poor's 500 (SNP): 500 large companies.
- Equity
  - Apple Inc.: Large technology company;
  - Microsoft: Large technology company;
  - McDonald's: Fast food company;
  - American International Group (AIG): Insurance company.
- Price of Metal
  - Gold (Au);
  - Silver (Ag);
  - Copper (Cu);
  - Platinum (Pt).
- Currency
  - EUR-GBP;
  - GBP-JPY;
  - GBP-USD;
  - USD-JPY.

Figures 6.1 provides an overall view, and exhibits the original price in the upper panels and the detrended price in the bottom panels of six indices/stocks. The detrended price in blue in each figure was produced by a MA filter at scale = 5, where the detrended dynamical time series were the main signals involved in the following analyses. When a crisis arises, the price becomes more dynamic, which can be clearly observed in detrended signals with the removal of the local mean. In Figure 6.1, the S&P 500 and DJIA are collections of large companies in the US market. The NASDAQ and Apple Inc. illustrate the performance of the technology industry, where Apple Inc. is the largest stock in the NASDAQ index. And the bottom panels in Figure 6.1 represent the food industry (McDonald's) and real estate (AIG Insurance), respectively.

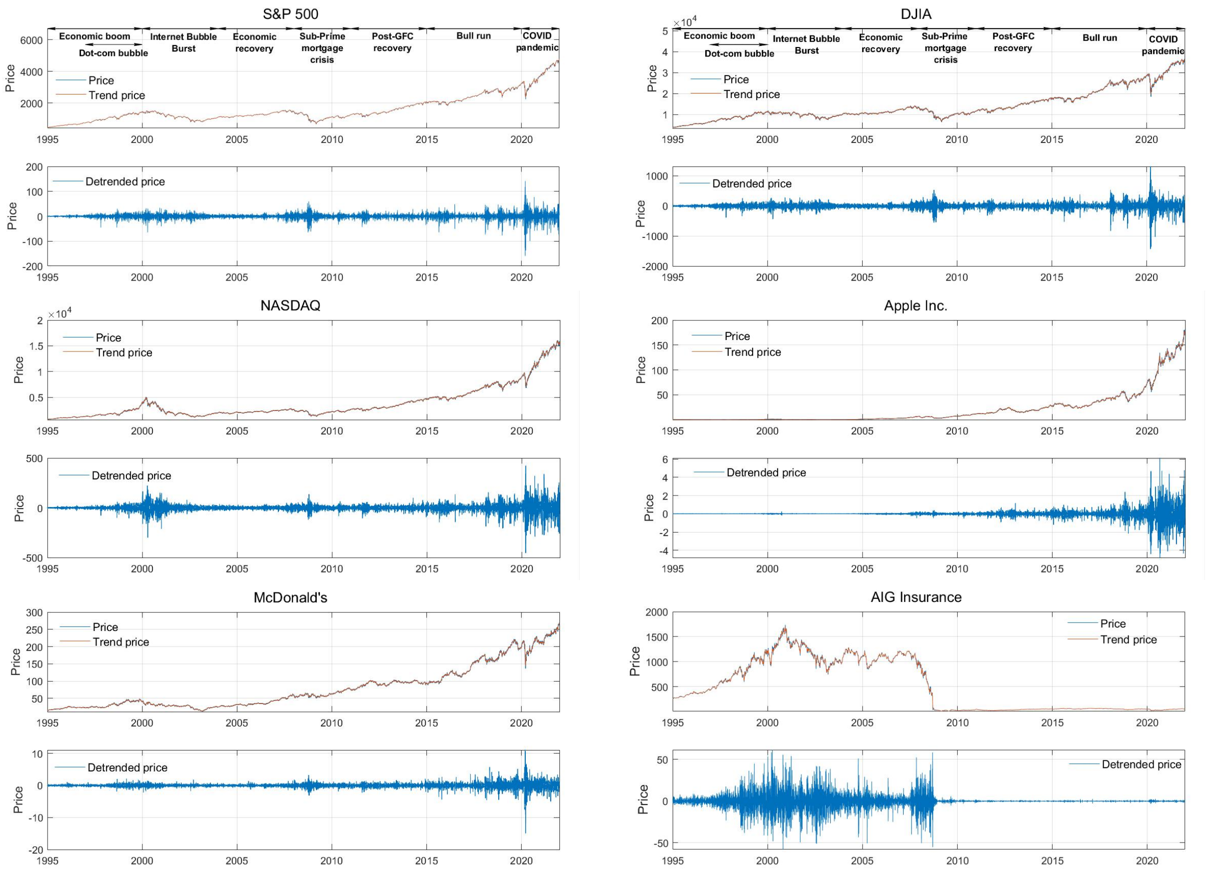


Figure 6.1: Exemplary price time series and their detrended dynamical signals over 1995-2021. Detrended signals in blue highlight the dynamic of the original signals with removal of trend.

Over the last 30 years, seven consecutive periods of different natures were identified based on the world economies [200–202], and are marked at the top of each figure. These are:

1. Economic boom/Dot-com bubble: 1997.01.01 to 1999.12.31.
2. Internet bubble burst (Crisis-1): 2000.01.01 to 2003.12.31.
3. Economic recovery: 2004.01.01 to 2007.12.31.
4. Sub-Prime mortgage crisis (Crisis-2): 2008.01.01 to 2011.12.31.
5. Post-Global Financial Crisis (GFC) recovery: 2012.01.01 to 2014.12.31.
6. Bull run: 2015.01.01 to 2019.12.31.
7. COVID Pandemic (Crisis-3): 2020.01.01 to 2021.12.31.

Observe that different crises show varying influences on different industries, which is reflected in the following complexity analysis. Apart from the impact of the crisis, the trend of price generally increases as time goes by, except the AIG insurance, which was severely attacked by the Sub-Prime mortgage crisis around 2008.

6.5 Results and Analysis

Due to the properties of the price time series (one data point per weekday), the length of one year contains 261 points. The analyses for univariate time series include Modified Univariate Multiscale Sample Entropy (Mod-MSE), Recurrence Plot Analysis (RQA) and Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS). The analysis for multivariate time series was generated using Modified Multivariate Multiscale Sample Entropy (Mod-MMSE). Furthermore, from the Mod-MSE and Mod-MMSE analyses, the catastrophe analyses the performance of individual price series and the overall US financial market. Figure 6.2 shows the analysis framework using the algorithms presented in Section 6.3.

6.5.1 Modified Univariate Multiscale Sample Entropy (Mod-MSE) and Modified Multivariate Multiscale Sample Entropy (Mod-MMSE)

Different settings of the pre-defined scale factors, τ, have been discussed in [53], where the various scale factors exhibited the same trend. Here, the short-term MA filter with

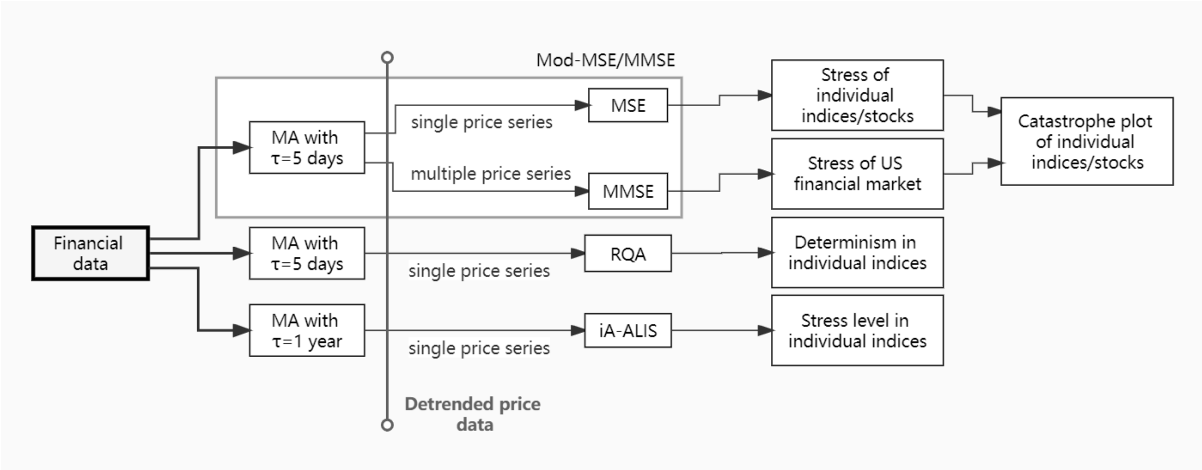


Figure 6.2: The employed analysis framework

$\tau = 5$ is selected, which was able to give the most distinct tracking of the financial stress evolution. The analysis window was set as $N = 1044$ (261 points×4 years) with a 1-day increment. Therefore, given the data from 1991, the complexity plot starts from 1995, whereby each entropy value was calculated based on the history price in the past 4 years. The default parameters of Mod-MSE/MMSE were set as the embedding dimension $m = 2$ and delay factor $l = 1$.

Figure 6.3 and 6.4 show respectively the complexity of the 4 indices and 4 equities considered as well as currency and metal price via Mod-MSE. The figures show the reciprocal of Mod-MSE, representing the level of stress according to the complexity-loss hypothesis. The increase of Mod-MSE indicates higher randomness referring to a 'normal' period and a decrease of stress when the dynamics of price are balanced and influenced by multiple factors. In terms of indices in Figure 6.3, the complexity plots validate the hypothesis that when financial crises arise (i.e., the Internet bubble burst, the Sub-Prime mortgage crisis and the COVID pandemic), all the indices showed different degrees of stress increase. As the collections of technology markets, NASDAQ (in green) showed the largest impact from the Internet bubble, while the Sub-Prime mortgage crisis had less influence. As for the COVID pandemic, all the indices illustrated a moderate increase in stress levels at the beginning of 2020 due to the general struggle of all industries. Among them, Russell 2000 took a relatively long time to recover back to normal, compared to the other three larger indices.

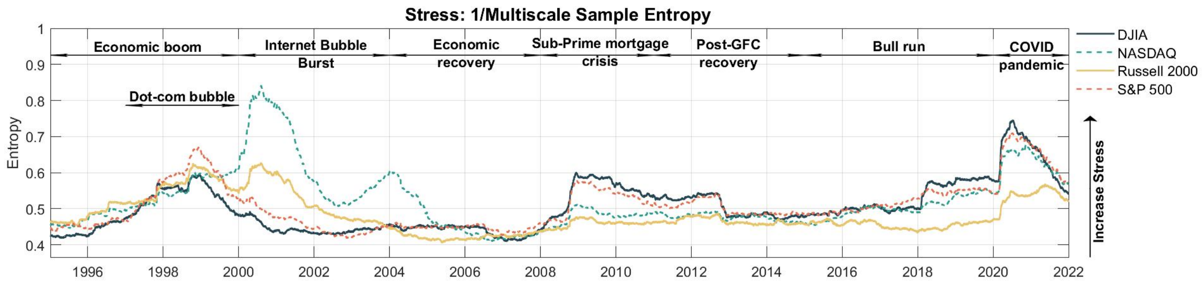


Figure 6.3: Financial stress of four indices (DJIA, NASDAQ, Russell 2000 and S&P 500) estimated by Mod-MSE over 1995-2021.

As for the individual equities in Figure 6.4, observe that the most stable index was the price of gold (in red) which kept a low level of stress over time, including the COVID pandemic. The same trend can be found in McDonald's, the world's largest fast-food chain, where the stress shown by the price dynamics of McDonald's stock was as low as the gold price. Apple Inc. is the most significant individual equity in the NASDAQ index. Hence, the trend of Apple Inc. (in black) in Figure 6.4 is partly in line with the change of NASDAQ (in green) in Figure 6.3. For example, the peaks after 2000 and around 2004 are presented in both plots. Observe also the step change of Apple Inc between 1996-2006, which could be the influence of new product releases. In contrast, this effect has been reduced in the recent 10 years.

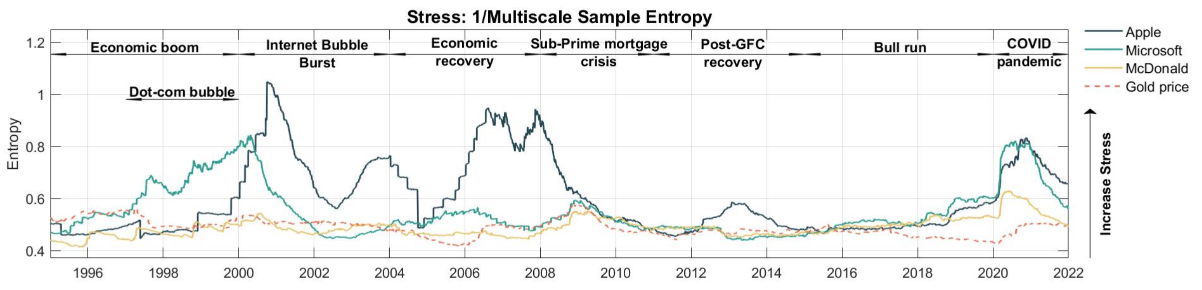


Figure 6.4: Financial stress of four leading equities in their sector (Apple Inc., Microsoft, McDonald's and Gold price), estimated by Mod-MSE over 1995-2021.

Next, we have applied univariate analysis on currency indices and metal price time series. The value of a currency can be influenced by a number of factors including the government's economic policies and their national central bank [180], and national central banks are closely associated with their metallic reserves [203].

The four currency indices are EUR-GBP, GBP-JPY, GBP-USD and USD-JPY. To demonstrate the stress level, the reciprocal of Mod-MSE is plotted in Figure 6.5 in line with the complexity-loss theory. The stress level given by Mod-MSE showed that USD-JPY (in red) was at a high-stress level before the end of the Sub-Prime Mortgage crisis. GBP-JPY (in green) and EUR-GBP (in black) exhibited higher stress levels in Sub-Prime mortgages than that in the Internet bubble burst. Generally, the Forex market was less impacted by the COVID pandemic compared to the previous two global crises.

Figure 6.6 presents the stress levels of four mental prices: Gold (Au), Silver (Ag), Copper (Cu) and Platinum (Pt). In general, metal prices were less influenced by the Internet bubble burst than other stocks given in Section 6.5. The gold price (in black) and platinum price (in red) are at a low-stress level over time as expected. The copper price has shown the most sensitivity during Sub-Prime mortgage crisis, while silver received more impact from the attack of the COVID pandemic.

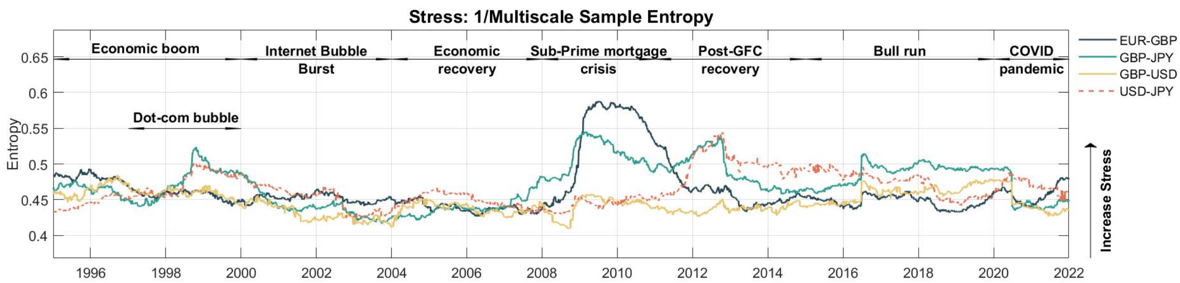


Figure 6.5: Financial stress of four currencies (EUR-GBP, GBP-JPY, GBP-USD and USD-JPY) estimated by Mod-MSE over 1995-2021.

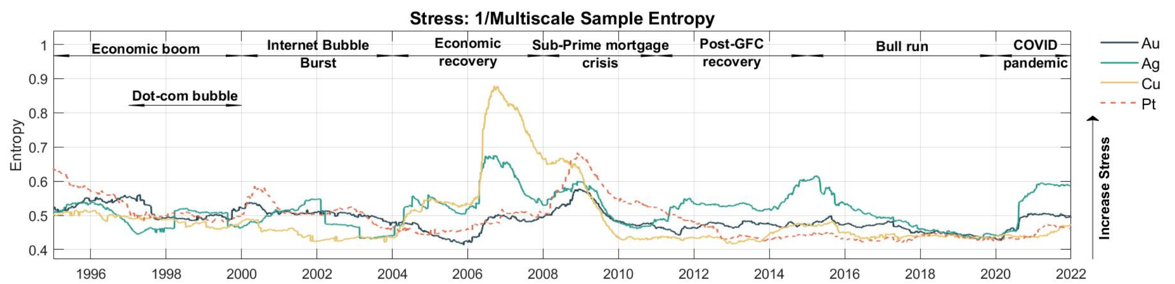


Figure 6.6: Financial stress of four metal prices (Gold-Au, Silver-Ag, Copper-Cu and Platinum-Pt) estimated by Mod-MSE over 1995-2021.

As shown in Figure 6.3 and Figure 6.4, entropy analyses implemented based on individual indices/stocks are able to assess the internal stress from the dynamics of a signal.

Although S&P 500 and DJIA are generally considered leading indices, single-channel analysis is sub-optimal to evaluate the overall performance of the US financial market. To this end, we applied the multivariate Mod-MSE with the 4 representative indices in Figure 6.3 as multi-channel data. To visualize the advantages of multivariate analysis, the Mod-MSE of the leading index (S&P 500) is jointly plotted with Mod-MMSE in Figure 6.7. Observe that during the Dot-com bubble, Sub-Prime mortgage and COVID pandemic, the stress in multi-channel increased accordingly, while for the 'normal' situations, it showed a flat stress change. With time, the stress levels of 'normal' periods increased. For example, the period of Post-GFC recovery after the Sub-Prime mortgage crisis is higher than the economic recovery after the internet bubble burst around 2006. Overall, the performance of Mod-MMSE was in line with Mod-MSE based on S&P 500 but was more sensitive and exhibited a stronger response towards the change of environment. Therefore, Mod-MMSE was able to reflect the overall stress of the US financial market better than any individual indices/stocks.

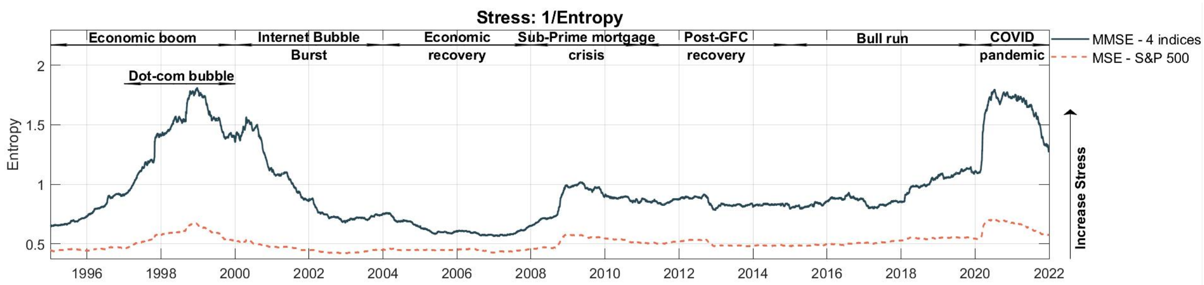


Figure 6.7: Financial stress of US market estimated by Mod-MSE and Mod-MMSE over 1995-2021. The red dashed line represents the SNP index estimated by MSE, while the black solid line depicts the stress estimation MMSE, which includes four variables: DJIA, NAS, RUS, and SNP.

6.5.2 Recurrence Quantification Analysis

Next, we applied the Recurrence Quantification Analysis (RQA) to yield the determinism degree (DET) of each index, which is another quantitative way to evaluate the stress level of the financial market. Recall that the EMH implies that during 'normal' financial regimes, stock prices behave in a random (uncertain) way [191]. The lower the determinism, the more stochastic components the system contains, referring to a 'normal' situation in line with the high randomness in entropy analysis. Therefore, the degree of determinism (DET) metric is consistent with the stress level of the index/stock, that is, we expect low determinism in normal situations and high determinism during the financial crisis.

In the analysis, the pre-defined scale factor, τ, was set to 5 days, and the window for RQA to 4 years. The optimal combination of the embedding dimension, m, and delay parameter, l, were jointly selected by the differential entropy-based method, proposed in [174].

Figure 6.8 shows the DET of 4 indices, DJIA, NASDAQ, Russell 2000 and S&P 500. As expected, DJIA and S&P 500 showed similar curves with high determinism during the Dot-com bubble, Sub-Prime mortgage crisis and COVID pandemic, as general responses towards environmental stress changes. On the other hand, NASDAQ in green exhibited high stress during the Internet bubble burst than the Sub-Prime mortgage crisis due to the intrinsic features of the technology market which have no relevance to housing investment. As for small companies (Russell 2000 in yellow), they showed delayed responses to crises compared to big companies, whereby the yellow line rose up following the other three indices at the beginning of the three crises, similarly, it took a long time to recover from crises. For example, around 2002, when DJIA and S&P 500 had largely dropped, it was still at a high level; also at the end of 2021, Russell 2000 (in yellow) was still in a highly deterministic state, while the stress of other indices had mostly decreased.

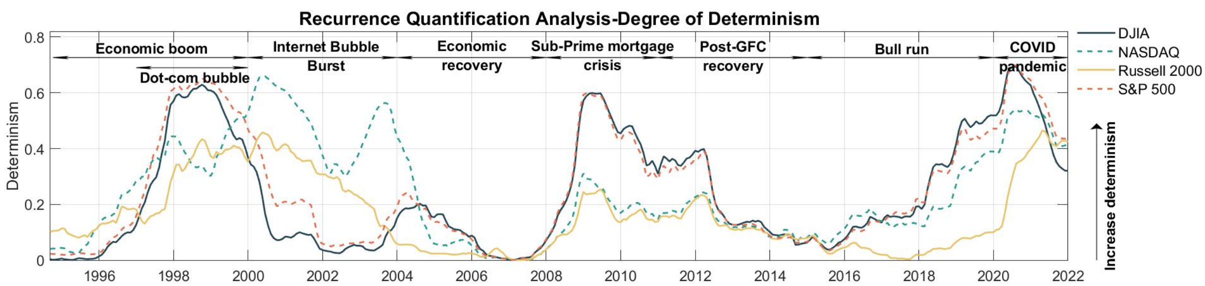


Figure 6.8: Degree of determinism of four indices (DJIA, NASDAQ, Russell 2000 and S&P 500), estimated by RQA over 1995-2021.

6.5.3 Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS)

The third methodology to estimate the stress level of financial indices/stocks is the Assessment of Latent Index of Stress with Instantaneous Amplitude (iA-ALIS) [3]. The ALIS was originally proposed to examine the 'economic organism' through the complexity-loss hypothesis, whereby a high-stress level is indicated by the high value of ALIS [53]. Here, we applied the enhanced iA-ALIS, whereby When estimating the 'power' of the low-frequency band and high-frequency band, the instantaneous amplitude is used in place of absolute power. Following the work in [53], a four-year sliding window was employed with a 1-day increment and the detrended data was given by MA filter on a scale of one year.

Figure 6.9 shows the stress levels of every index considered (DJIA, NASDAQ, Russell 2000 and S&P 500) given by iA-ALIS. The black dashed line in the bottom figure represents the threshold, which is the median value among all the indices over time. Observe that the iA-ALIS exhibited substantially high levels during the internet bubble burst and Sub-Prime mortgage crisis. However, when considering the impact of the COVID Pandemic, the dramatically high stress given by the global pandemic among all industries made measures of the previous two crises less significant when visualising. During the internet bubble burst, NASDAQ demonstrated a higher stress level than the other three indices as expected, which is supported by the Mod-MSE and RQA analyses.

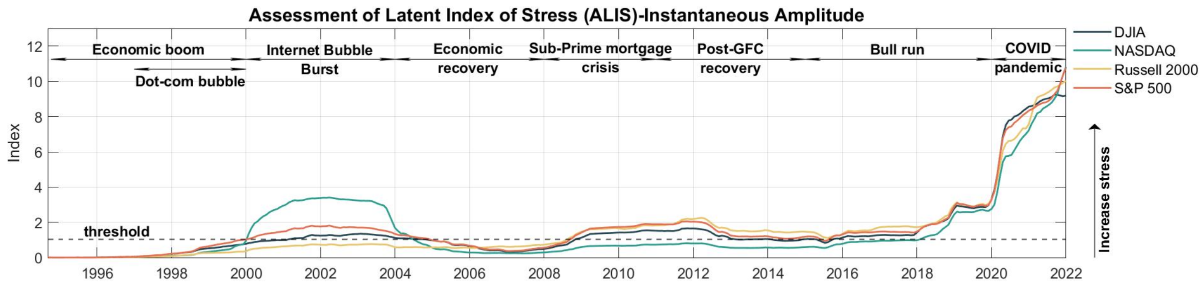


Figure 6.9: Financial stress of four indices (DJIA, NASDAQ, Russell 2000 and S&P 500) estimated by iA-ALIS over 1995-2021.

Observe also the problem with iA-ALIS whereby the highly dynamic changes in daily price were hard to distinguish. With the same resolution, Mod-MSE in Figure 6.3 and RQA in Figure 6.8 were able to give more details of the stress evolution, while the iA-ALIS measure, as shown in Figure 6.9, evaluated the stress levels in a smooth way. Although the periods of crisis could be marked by iA-ALIS in 2000, 2008 and 2020, limited information can be observed from iA-ALIS for further analysis.

6.6 Catastrophe Plots based on entropy

Complexity theory suggests that all aspects of complex problems have the characteristics of catastrophes [195]. Catastrophe theory is a branch of applied mathematics, developed by Rene Thom in the late 1960s [193], whereby the basic idea of Catastrophe Theory aims to explain the breakdown of relationships in variables of a dynamic system [204]. More specifically, it offers a potential method to describe the ability of a smooth change in system parameters to generate catastrophic behaviours (i.e., abrupt, discontinuous, sudden change) in a dependent variable, termed critical points [55, 194, 196, 204]. Therefore, Catastrophe theory has the potential for describing all aspects of natural phenomena due to their complex properties; it embodies a theory of great generality which is perceived as a state of mind [194, 205].

In the realm of physiology, Hardy & Fazey (1987) state that physiological arousal is related to performance in an Inverted-U hypothesis when the athlete is not worried or has low cognitive anxiety. If cognitive anxiety is high, the increases in arousal pass a point of optimal arousal and a rapid decline in performance occurs [206]. Catastrophe Theory has also been involved in brain modelling [207], however, the catastrophe model still remains in its conceptual framework state without computational analysis [206]. When it comes to financial applications, Catastrophe Theory represents a unique hypothesis made up of different mathematical structures, in contrast to the Efficient Market Hypothesis (EMH) [194]. Indeed, Catastrophe Theory has been tentatively employed to explain discontinuous jumps in bank investment [55] and studies have shown that it could explain the crash of the stock exchanges better than other models [196]. Yet, the number of studies using catastrophe theory in economics is fairly small and mostly built based on qualitative descriptions rather than quantitative applications [55, 196]. Considering the high complexity and unpredictability of the stock market and its chaotic and uncertain behaviours, Catastrophe Theory offers the potential to explain the occurrence of financial events [194]. To this end, we propose a practical framework for the application of Catastrophe Theory on basis of the complexity estimation through entropy.

6.6.1 Arousal-Performance plot of index

Considering the similarity between the financial and physiological systems, the analogy between catastrophe theory applied to athlete performance and index/stock performance can be drawn. The catastrophe plot in physiological systems reflects the relation between physiological arousal (anxiety) and performance. Accordingly, in the financial market, we observe the performance of individual indices/stocks evaluated by Mod-MSE; that is, the higher the stress level given by lower Mod-MSE the lower the performance. While the arousal in the physiological system is an internal trigger that determines performance, due to the intrinsic property of a financial system, we modelled the dependent factor (arousal) as the external stress imposed by the external environment on the whole market and quantified by the Mod-MMSE. Therefore, the catastrophe framework in the financial market describes the relationship between the level of overall market anxiety (external stress) and the performance of individual indices/stocks (internal stress).

Two indices (S& P 500 and NASDAQ), three individual stocks in different industries (Apple Inc., AIG and McDonald's) and the price of gold (AU) were selected to demonstrate the Arousal-Performance Plot as shown in Figure 6.10. Good performance is reflected in low internal stress indicated as high Sample Entropy, while strong arousal/stimulus is designated as high external stress given by the reciprocal of Multiscale Multivariate Sample Entropy (MMSE), as discussed in Figure 6.7. The line plots in Figure 6.10 depict a relationship between the performance of each index/individual stock and external stimuli during the whole 27 years (1995-2021), where light colours refer to early years and dark colours represent late years; all the plots were adjusted to the same axis scale.

The Catastrophe plots of two indices (S&P 500 and NASDAQ) are shown in the first two panels in Figure 6.10. As a collection of big companies, S&P 500 and NASDAQ exhibit a similar tendency of performance increase as the external stress decreases. Observe that the curve of S&P 500 tends towards more of a linear relationship compared to the curve of NASDAQ, showing the higher predictability/regularity of S&P 500. The larger slope of the NASDAQ curve in light green shows that the technology market was under higher stress (with sub-optimal performance) in the early years than in recent years.

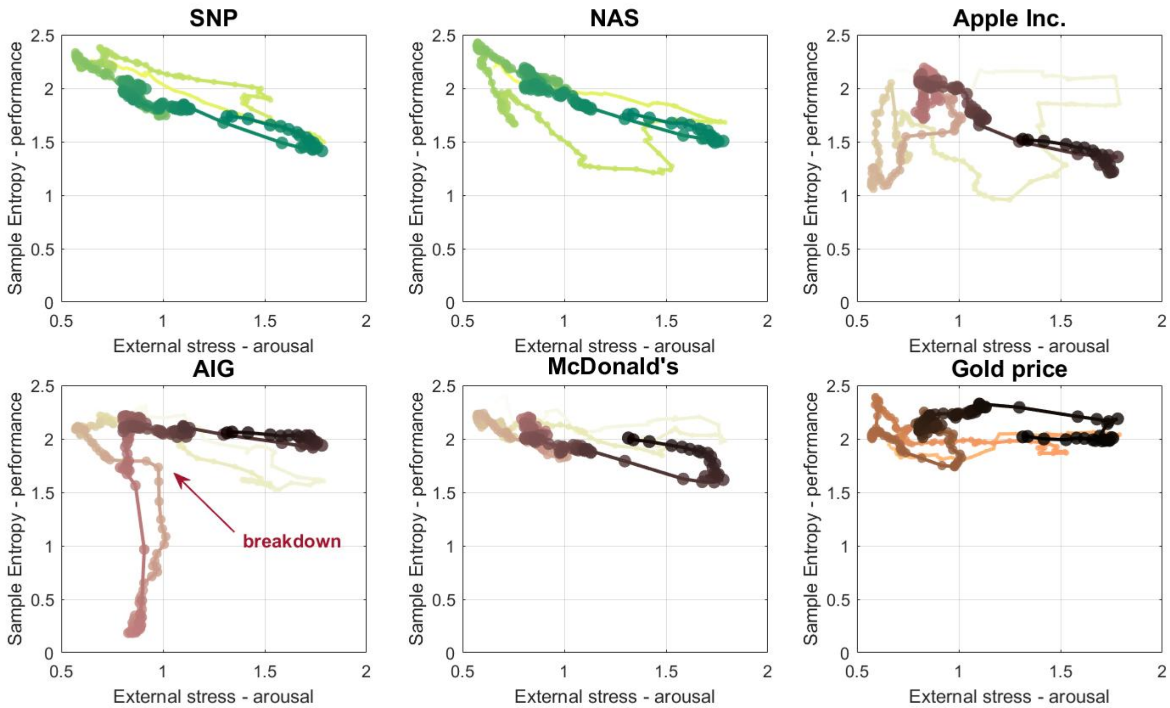


Figure 6.10: Evaluation of Catastrophe Plots of 6 indices (SNP, NAS, Apple Inc., AIG, McDonald's and gold price) over the period 1995-2021. The plots depict the stress of an individual asset in response to the stress of the whole market (sensitivity), where the performance is measured by the entropy of each index, while the external stress is estimated using MMSE shown in Figure 6.7. The plots are colour-coded, starting from the lightest shade and ending in the darkest shade. A good performance is characterized by a minimal decrease as stress levels increase, and a linear relationship between performance and stress signifies high robustness of the index.

Next, the curve of the most significant stock in NASDAQ, Apple Inc. is in line with the change of NASDAQ in recent years as shown in dark colours; this is due to the leading role of Apple Inc. in the technology market in the US. As for American International Group (AIG), the only individual stock that has experienced catastrophe change among the given indices is plotted in the first graph of the second row in Figure 6.10. Observe that, in the early and late years, AIG exhibits a relatively stable response to the overall external stress. What is striking is the sharp drop in the middle years (during the Sub-Prime mortgage crisis), which reflects the sharp increase of internal stress of the individual stock in response to a small change in external stress. Therefore, AIG equity exhibited two critical points in Catastrophe Plot: i) the lowest performance/highest stress level the stock could sustain; and ii) the highest performance point that could manage to bring the equity back to a normal state. Notice also that the recovering point is higher than the breakdown point. In terms of highly robust stocks such as McDonald's and gold price, both exhibit a relatively flat relationship between arousal and performance. The flat curves show the highly stable performance of the equity/index under large changes in external stress. McDonald's stock indicates a decrease in stability in recent years compared to the early years as evidenced by the more inclined tendency, while the price of gold keeps a high performance without apparent influence by the stress caused by environmental change.

6.6.2 Arousal-Performance plot of the index in crises

Next, we extracted the three specific 2-year periods of crisis from the previous Catastrophe Plot to discuss the response of each index/individual stock to different crises. These are the Internet Bubble Burst (IBB) between 2002-2002 in green, the Sub-Prime mortgage crisis between 2008-2010 in brown, and the COVID pandemic between 2020-2022 in red. To indicate the direction of each segment, we used the gradient colour, where the lightest colour refers to the start of the selected period of time. The selected Catastrophe Plots are illustrated in Figure 6.11.

According to the Catastrophe Theory, with a fixed increase of external stress (arousal), a good response should be shown as a small decrease or non-decrease of Sample Entropy (performance). Therefore, the slope of lines could quantitatively reflect the overall performance of the index in response to different crises. Three indices and three individual stocks are examined in Figure 6.11, with the same axis scale. The index in the first panel indicated that S&P 500 showed stable responses to all three crises with different degrees of robustness. Observe that, among the selected segments, the Sub-Prime mortgage crisis had the largest impact on the S&P 500 index (in red) with a sharp decrease in performance, while for NASDAQ and Russell 2000, the Internet bubble burst exhibited a larger influence than the other two crises. NASDAQ, in particular, was critically affected due to the technological domination of the index.

In terms of the equities in the second row in Figure 6.11, in line with the plot in Figure 6.10, the Catastrophe plot of the insurance company, AIG, exhibits a breaking point during the Sub-Prime mortgage crisis. It reverted back to a stable state during the COVID pandemic. As for Apple Inc. stock, the non-decreasing tendency during the Sub-Prime mortgage crisis (in red) demonstrates the limited influence of the specific stimuli (with non-decreasing performance). The most significant crisis for Apple Inc. was the Internet Bubble Burst (in green) as expected. The gold price is given in the last panel. As the most stable index, the performance of gold price has been at a relatively high level throughout three crises, with a slight drop in the Sub-Prime mortgage crisis and the COVID pandemic.

6.6.3 Arousal-Performance plot of crises

Finally, to compare the performance of indices/individual stocks in each of the crises, we selected the arousal-performance plots of five indices/individual stocks in three crises separately, as shown in Figure 6.12. The left panel shows the performance during the Internet Bubble Burst between 2002-2002. Observe that the most stable two indices are the S&P and gold prices (in blue and red, respectively) with flattening tendencies. Those which were the most significantly impacted are the NASDAQ and Apple Inc. in brown and grey. The Sub-Prime mortgage crisis in the middle graph shows that the stock that suffered most is AIG (in green) which exhibits an elbow. The remaining stocks showed similar slopes apart from Apple Inc., which exhibited an increasing performance as the

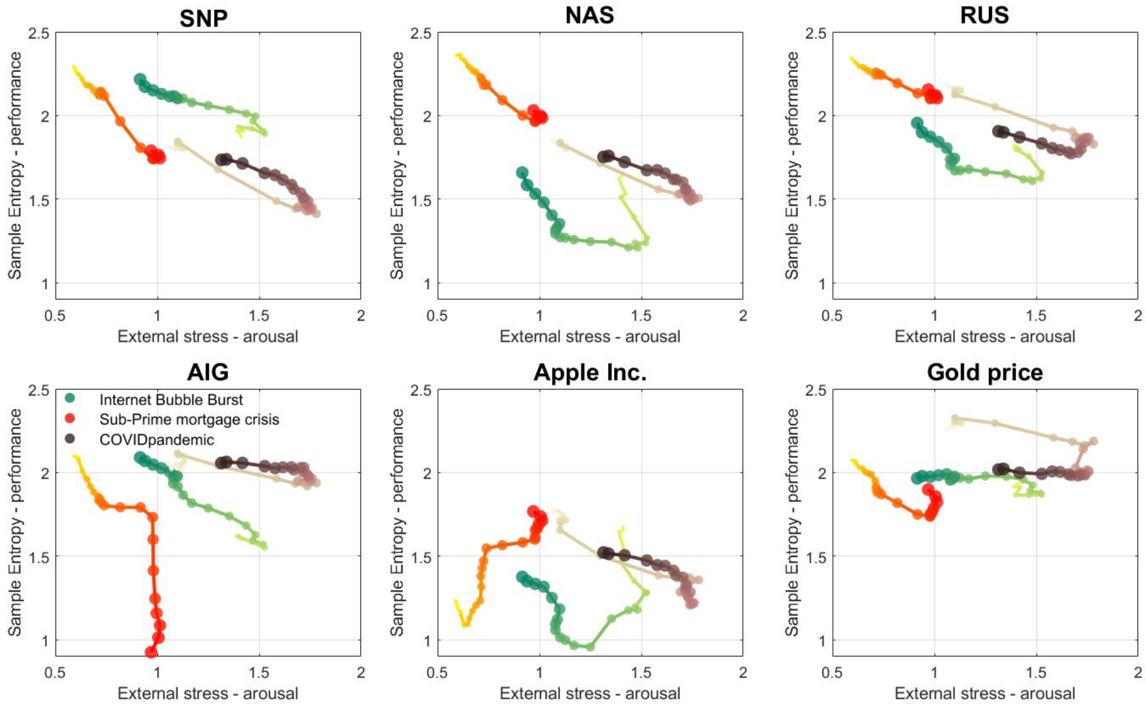


Figure 6.11: Catastrophe Plots of 6 indices (S&P 500, NASDAQ, Apple Inc., AIG, McDonald's and gold price) during three selected periods of crises (Internet Bubble Burst, Sub-Prime mortgage crisis and COVID pandemic). Each segment is colour-coded, starting from the lightest shade and ending in the darkest shade lasting a period of 2 years.

external stress rises.

As for the recent COVID pandemic in the right panel, all the indices/stocks display similar responses to the crisis; that is, performance decreases at the beginning of the crisis and then gradually raises back as the stress relaxes, and the elbow point in each curve means the stock started the recovery process from the impact of the crisis. The highly aligned tendencies of all the indices emphasize the general influence of the global pandemic on all the industries in the US financial market.

6.7 Conclusion

We have estimated financial stress from the viewpoint of nonlinear dynamics and have examined the significance of structural complexity features in financial analysis. This has been achieved for four stock market indices and four individual equities from 1991-

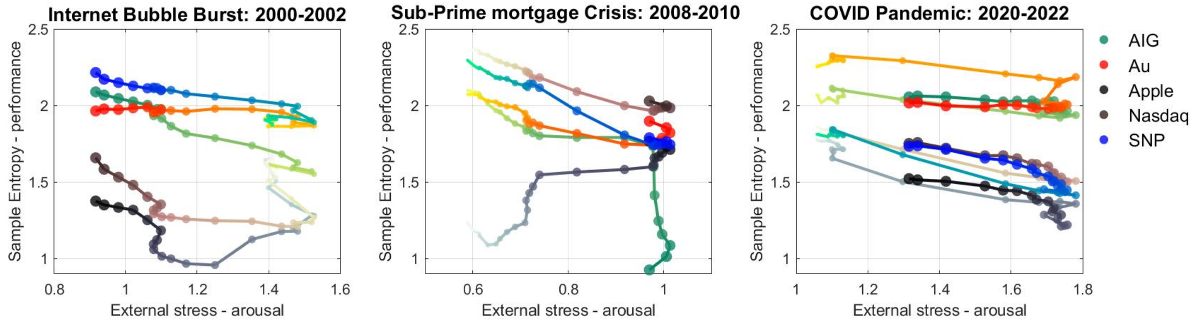


Figure 6.12: Catastrophe Plots of three 2-year periods of crises (Internet Bubble Burst, Sub-Prime mortgage crisis and COVID pandemic) for selected 5 indices (S&P 500, NASDAQ, Apple Inc., AIG and gold price). Each segment is colour-coded, starting from the lightest shade and ending in the darkest shade

2021. The financial stress has been estimated based on Mod-MSE in both univariate and multivariate cases. In addition, the univariate RQA and iA-ALIS methods have been shown to give the degree of determinism and stress change, in line with the entropy-based analysis. All three nonlinear approaches have demonstrated their ability to quantify financial stress, with multivariate entropy being the most information-rich and physically meaningful.

Further, a novel framework based on Catastrophe Theory has been proposed, where Arousal-Performance plots have been employed to visualise the response of each financial index/stock. We have adopted Mod-MMSE of four major indices as a metric of external arousal and Mod-MSE of each index/individual stock as a metric of performance. The analysis has demonstrated that the same crisis triggers different performance changes in various industries and that the same index/equity exhibits various robustness to different types of crises (the Internet Bubble Burst, Sub-Prime mortgage crisis and COVID pandemic crisis). Finally, through Catastrophe Plots, the 'performance' of index/individual stock has been qualitatively and quantitatively explored.

The arousal and performance in this study have employed entropy-based estimation methods. Future studies will focus on incorporating different stress quantification approaches based on the proposed framework.

Chapter 7

Robustness Evaluation: Young's Modulus in Finance

7.1 Chapter Summary

Financial signals exhibit highly dynamic behaviours, with the most extreme cases in the form of market crashes. Engineering disciplines often employ the stress-strain characteristics and the associated Young's modulus to ascertain whether a material/phenomenon is about to undergo a catastrophic collapse. This motivates us to investigate the feasibility of a "Financial Young's Modulus" which would indicate the level of "brittleness", "plasticity" and "elasticity" of financial markets. We implement and examine the modified Young's Modulus for Finance (YMF) on the time series of prices, including four leading indices and six individual stocks over the last 31 years, from 1991-2021. The results verify that the proposed YMF can be used to indicate and forecast the resilience of individual stocks and equities, especially during a crisis.

7.2 Introduction

Young's Modulus (YM) quantifies a material's deformation under stress, with higher values indicating greater stiffness and lower values suggesting increased flexibility [208, 209]. The fundamental expression for YM in material evaluation is the ratio of stress change to strain change. Applying pressure to a material reveals its stiffness, with stiffer materials exhibiting minimal strain (i.e., length change). Importantly, the linear correlation between stress and strain is confined to the elastic deformation regime. This condition prevails when the material's structure withstands pressure without permanent damage, ensuring the deformations are reversible.

In the financial domain, the reaction of an index or individual stock to market stress exhibits analogous behaviour. This chapter explores the correlation between the stiffness of a material and the resilience of the financial system using price time series. The proposed Young's Modulus for Finance (YMF) incorporates modifications to the formula to align with the intrinsic characteristics of financial signals. The pressure applied to the material finds its counterpart in the 'overall stress' experienced by the entire financial market, while the change in strain on the material aligns with the resulting price change of the index [4,53,210]. We presume that this relationship adheres to the elastic deformation regime. Therefore, under a specific stress scenario (such as a crisis), a minor change in price variance would yield a high Young's Modulus (YM), providing a measure of the index's resilience or the stock's robustness in the financial market.

To assess the effectiveness of the proposed Young's Modulus for Finance (YMF), we employed the YMF formula on four indices and six individual stocks spanning the 31-year period from January 1, 1991, to December 31, 2021. The entities analysed are:

- Stock market index
  - Dow Jones Industrial Average (DJIA/DOW): 30 large companies;
  - NASDAQ Composite (NAS): Mid- and large-caps;

  - Russell 2000 (RUS): Smaller companies;
  - Standard & Poor's 500 (SNP): 500 large companies.

- Individual stock

  - Apple Inc.: Large technology company;
  - Microsoft: Large technology company;
  - Amazon.com: Large technology company;
  - McDonald's: Fast food company;
  - American International Group (AIG): Insurance company.
  - Au: Gold price.

Over the past three decades, researchers have identified seven consecutive periods characterized by distinct economic conditions worldwide [200–202], which are presented in Chapter 6.4.

The subsequent sections of this chapter are structured as follows: Section 7.3 provides a detailed introduction to Young's Modulus for Finance. Section 7.4 implements the proposed formula and explores the robustness property within and across the provided financial time series. Concluding remarks are presented in Section 7.5.

7.3 Young's Modulus formulas in financial realm

Commencing with the concept of Young's Modulus (YM) in materials, Figure 7.1 illustrates a comparison between physical and financial systems. The left side introduces the original YM formula, delineating the modulus of toughness as the change in length ($\Delta strain$) under certain pressure ($\Delta stress$). This linear relationship exists only during elastic deformation without structural damage (before the breakpoints in the graph) [209]. On the right side, the 'strain' in financial analysis represents an index, resulting in a change in price as the response of the index/stock [179]. However, due to the intrinsic properties of financial markets, applying the original Young's Modulus formula in a complex system poses several challenges, as outlined below:

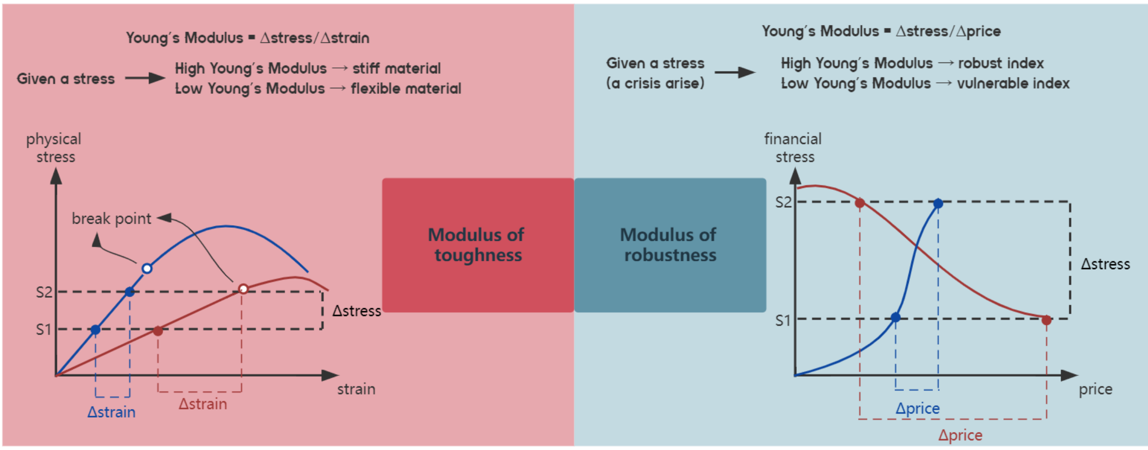


Figure 7.1: Analogy diagram between Young's Modulus in physical systems and financial systems. In the left graph, Young's Modulus is computed in the linear region before the breakpoint as the ratio of the change in stress, $\Delta stress$, and the change in strain, $\Delta strain$. Two materials are depicted (in blue and red), with the blue one representing a stiffer material possessing a higher YM. In the right graph, based on the assumption of an elastic relationship, the new Young's Modulus for Finance is determined by the change in stress and price, as illustrated on the right-hand side. Two indices are showcased (in blue and red, changing in different directions), with the blue index demonstrating a more robust property, exhibiting less fluctuation in price, $\Delta price$, under the same stress, $\Delta stress$.

- The first challenge lies in the fact that, unlike simplified physical systems, financial systems are open systems. Market stress is influenced by various interacting factors, encompassing both endogenous and exogenous elements, such as company news, unexpected events, and human behaviours demonstrating herding tendencies [4]. Unlike closed physical systems, the impact on financial systems cannot be distinctly separated and precisely modelled. Therefore, drawing inspiration from entropy-based financial stress evaluation [53,210], we employed complexity analysis to estimate the stress level of the financial market as an ideal measure of overall stress experienced by the market [19,56,182]. Notably, stress estimated by a single index or individual stock can only represent the stress exhibited by that specific stock in response to market shocks, termed idiosyncratic (i.e., stock-specific) risk. To address this, as suggested in [210], we applied Modified Multivariate Multiscale Sample Entropy (Mod-MMSE) to four leading indices (S&P 500, DJIA, Russell 2000, and NASDAQ) to ideally represent the global stress of the entire financial market. Mod-MMSE has demonstrated greater sensitivity to environmental changes compared to univariate entropy measurements on single indices or individual stocks. Details of Mod-MMSE can be found in Algorithm 26 in Chapter 6 [127, 136, 210], and it is established that the stress level inversely correlates with Sample Entropy-based evaluations, indicating the degree of randomness [53, 180].

- The second distinction arises from the decision to quantify the change in stress and price. In material analysis, an extension (i.e., a positive shift in stress) invariably results in an increase in length (i.e., a positive change in length). In essence, the signs of $\Delta stress$ and $\Delta strain$ in Young's Modulus of material are aligned. However, the scenario is more intricate in financial systems. When introducing positive stress to the financial market, the price difference can manifest as either negative or positive, as illustrated in Figure 7.1. Consequently, to ensure consistent changes in $\Delta stress$ and $\Delta price$, we employ the standard deviation in lieu of the difference to portray the price fluctuations, given its heightened dynamism.

- Subsequently, the amplitude of prices for various indices/individual stocks can exhibit significant disparities. To standardize the price within a controlled range, we opt for normalized prices instead of raw prices. This normalization is achieved by dividing the mean of the selected price segment when computing $\Delta price$, taking into account the absence of outliers in financial data.

- Considering that the time series of prices is not normally distributed, being influenced by the data trend, we apply the logarithm to both the denominator and numerator in the Young's Modulus formula to transform its distribution into a bell shape (i.e., Gaussian distribution). Consequently, the modified Young's Modulus for Finance (YMF) is then expressed as Equation 7.1. It represents the ratio of the standard deviation of the overall financial stress in the market to the standard deviation of the normalized price of the target index/individual stock. Here, financial stress is quantified by the reciprocal of the Modified Multivariate Multiscale Sample Entropy (Mod-MMSE).

$$YMF = \frac{\Delta stress}{\Delta price} = \frac{std[log(stress)]}{std[log(normalizedprice)]} = \frac{std[log(\frac{1}{Mod-MMSE})]}{std[log(normalizedprice)]} \tag{7.1}$$

7.4 Young's Modulus Results

7.4.1 Window selection of Young's Modulus for Finance

The only adjustable parameter in the Young's Modulus for Finance (YMF) formula is the selection of the window length. The choice of this parameter plays a crucial role in the stability and sensitivity of the estimation. Opting for a small, narrow window may result in an unstable estimation, as it becomes highly sensitive to the dynamic changes in the price signals. On the other hand, a larger window setting may smooth out the trend of robustness changes, failing to provide a detailed resolution of their responses to crises. In this section, we explore and compare the results of YMF with window lengths set at 6 months, 1 year, 2 years, and 4 years. The analyses are conducted on a large index (SNP), an insurance company (AIG), and the price of gold (AU), as presented in Figure 7.2 to Figure 7.4. Each figure includes three panels: the upper panel displays the variance of the normalized price for the current window setting, the middle panel illustrates the change of stress within various window lengths, and the bottom panel presents the final YMF curves. This comparative analysis provides insights into the impact of window length selection on the stability and responsiveness of the YMF formula across different financial entities.

In Figure 7.2, the small fluctuations observed with the 6-month and 1-year window lengths obscure the distinct responses to 'normal' and 'abnormal' situations. However, with the

2-year window setting (in yellow), the YMF index in the last panel reveals increasing trends during three crises and a relatively flat line during normal situations, exhibiting a smooth change. In contrast, the 4-year window case (in red) provides a coarse resolution of the change in robustness, leading to delayed estimations where the peaks of crises are shifted backward. For instance, the strong response to the Internet Bubble Burst is depicted around 2000 in yellow but is delayed to 2002 in the red curve. This comparison highlights the importance of choosing an appropriate window length in capturing nuanced variations in the YMF index, especially during critical financial events.

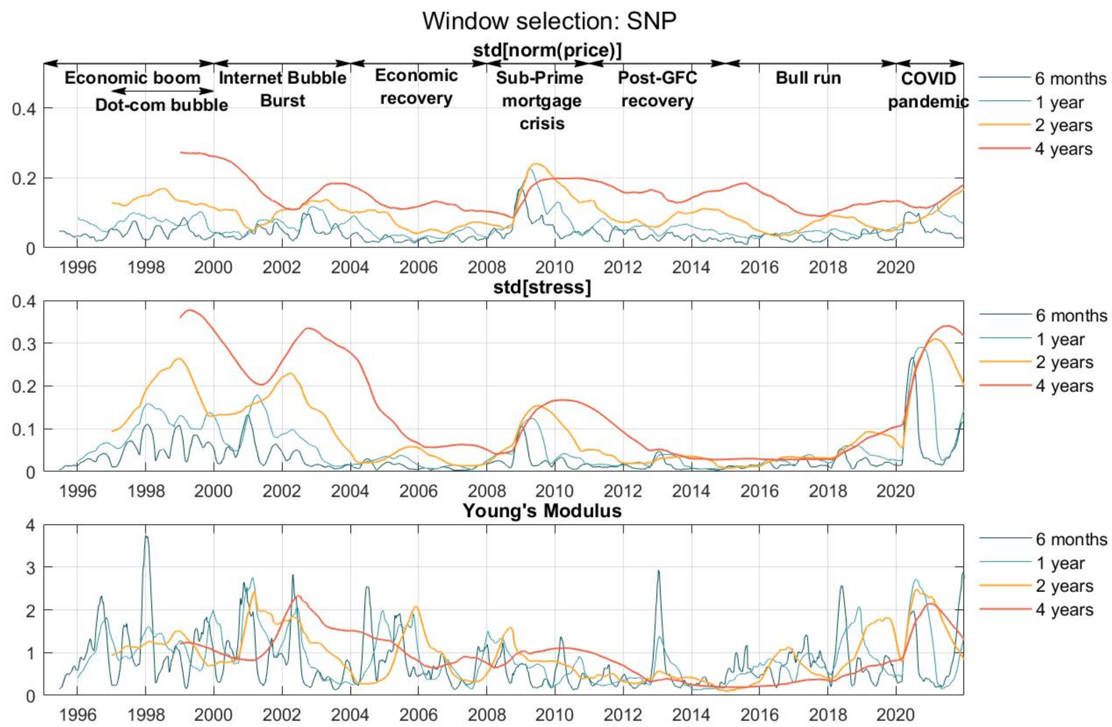


Figure 7.2: Performance of YMF measurement with different window settings on analysis of SNP.

The same analysis was conducted on AIG stock and is illustrated in Figure 7.3. The price signals of AIG, as depicted in Figure 7.6, exhibit a pronounced breakdown point during the Sub-Prime mortgage crisis around 2009. In the bottom panel of Figure 7.3, both the 2-year and 4-year window cases accurately capture the decline in performance around 2009. However, the responses to preceding changes before the breakdown point were smoothed and delayed when a 4-year window was applied. This observation underscores the trade-off between capturing fine-resolution variations and avoiding delays in detecting critical events, emphasizing the importance of selecting an appropriate window length for robust financial analysis.

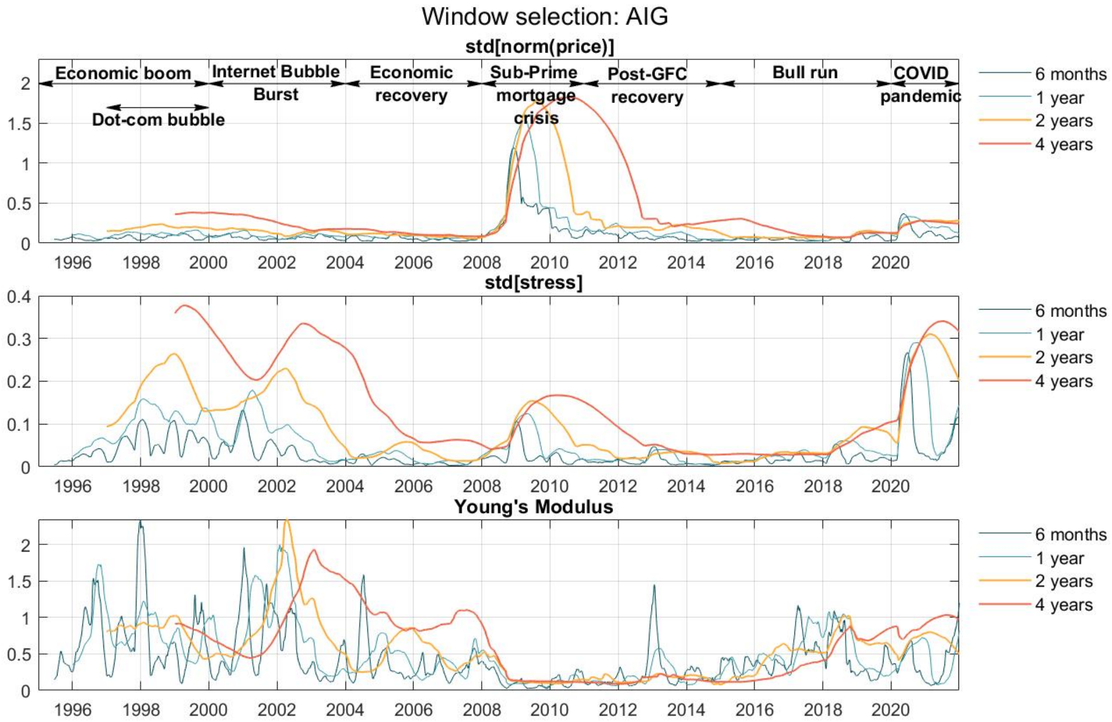


Figure 7.3: Performance of YMF measurement with different window settings on analysis of AIG.

The final example, based on the price of gold, is presented in Figure 7.4. The change in the YMF index of gold price in response to the COVID-19 pandemic in 2020 is depicted for window lengths of 6 months, 1 year, and 2 years, with the 2-year window providing the most distinct and informative results. In general, the 2-year window emerges as the optimal length for the algorithm, effectively balancing the trade-off between time delay and unstable fluctuations. Consequently, the subsequent analyses of other indices and individual stocks in this chapter were conducted using a fixed 2-year window length.

7.4.2 Young's Modulus on single index

According to the proposed formula, Figure 7.5 illustrates the performance of Young's Modulus for Finance (YMF) alongside the step signals for S&P 500. The window length selected for YMF calculation is 2 years (522 data points). The top panel in Figure

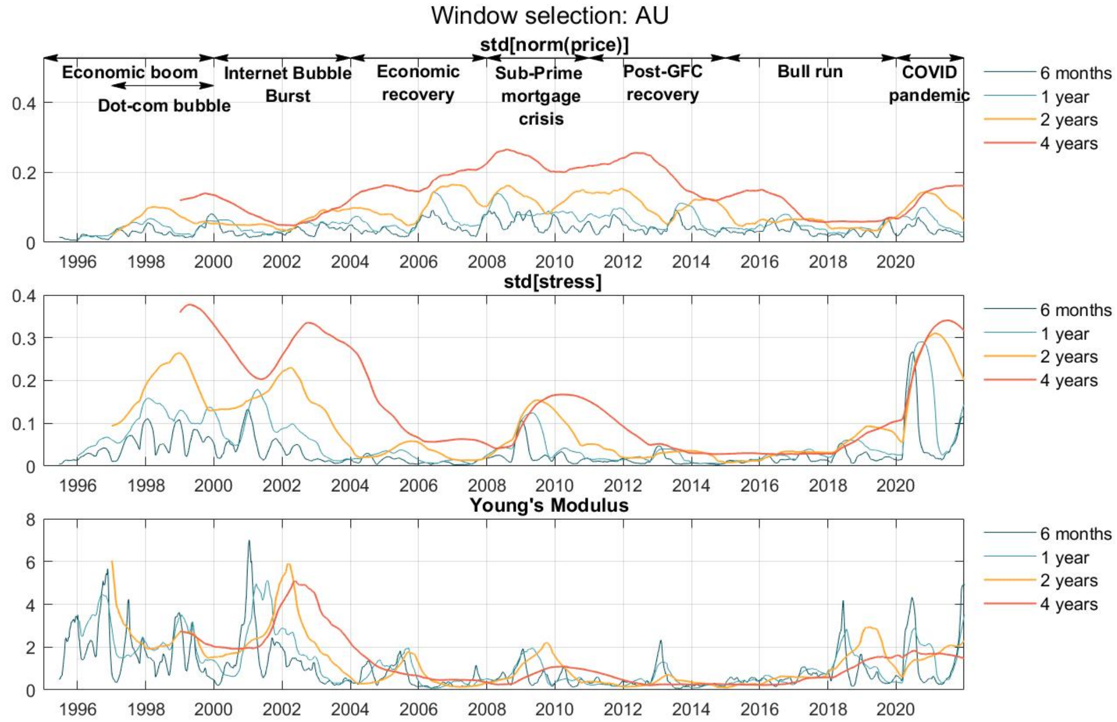


Figure 7.4: Performance of YMF measurement with different window settings on analysis of gold price (AU).

7.5 displays the raw price data and the mean of the price within the selected window from 1995 to 2021. The second panel presents the overall financial stress of the market using the reciprocal of Mod-MMSE and its mean values within the set window. The third panel showcases the change in the logarithm of stress, which is the numerator in the YMF formula. The fourth graph illustrates the change in the logarithm of the normalized price, the denominator in the YMF formula. The last panel displays the result of the final YMF index, with the red dashed line in the bottom figure indicating the level of the median value of the YMF curve.

It is important to note that the window length for calculating YMF was set at 2 years as discussed, equivalent to 522 samples. Consequently, the curves representing mean price and mean stress in the first and second panels, along with those in the last three graphs, were computed based on data from the past 2 years, starting in 1997. In the last three graphs of Figure 7.5, five segments highlighted in different colours represent specific 2-year periods, including three global crises over the last 31 years: the Internet Bubble Burst (2000-2002), Economic Recovery (2005-2007), the Sub-Prime mortgage crisis (2008-2010), the Bull Run (2015-2017), and the COVID pandemic (2020-2022). Recognized as one of the most stable and robust indices, the S&P 500 serves as a leading index in the financial market. A low YMF index conveys different implications in normal and abnormal situations. During a crisis, a low YMF signifies a weak response to applied stress, indicating low resilience, where the index/stock is significantly impacted by the crisis. Conversely, during normal circumstances after a period of intense stress, the price is expected to behave in accordance with the 'return to normal' assumption [179]. Consequently, the YMF is anticipated to be at a low level due to the relaxed environment and assumption. As illustrated in Figure 7.5, during a crisis (in blue, yellow, and green segments), the YMF index of S&P increases accordingly, underscoring the index's robustness to all three crises. The curve exhibits a notable 'return to normal' behaviour, with YMF dropping after the crisis.

As an illustrative example within the financial sector, we conduct an analysis focused on the American International Group (AIG) insurance company, as depicted in Figure 7.6. The raw price data of AIG (shown in the first panel) reveals a pronounced impact from the Sub-Prime mortgage crisis compared to the other two crises. Examining the YMF index in the last panel, it is notable that before the onset of the Sub-Prime mortgage crisis (in 2009), the YMF during the Internet Bubble Burst (in blue) indicates that the stock exhibited responsiveness to the crisis. However, it demonstrates remarkably low resilience during the Sub-Prime mortgage crisis (in the yellow segment), signaling the stock's vulnerability and substantial damage from that particular crisis. Despite the apparent price disruption shown in the first panel, the standard deviation of the normalized price suggests that, following a period of volatile change, the stock managed to recover from a highly dynamic stable situation to a less dynamic stable state, as depicted in the green segment during the COVID pandemic in the YMF panel. Therefore, the YMF index provides a comprehensive depiction, indicating that AIG exhibits relative robustness to the Internet Bubble Burst and the COVID pandemic but vulnerability to the Sub-Prime mortgage crisis.

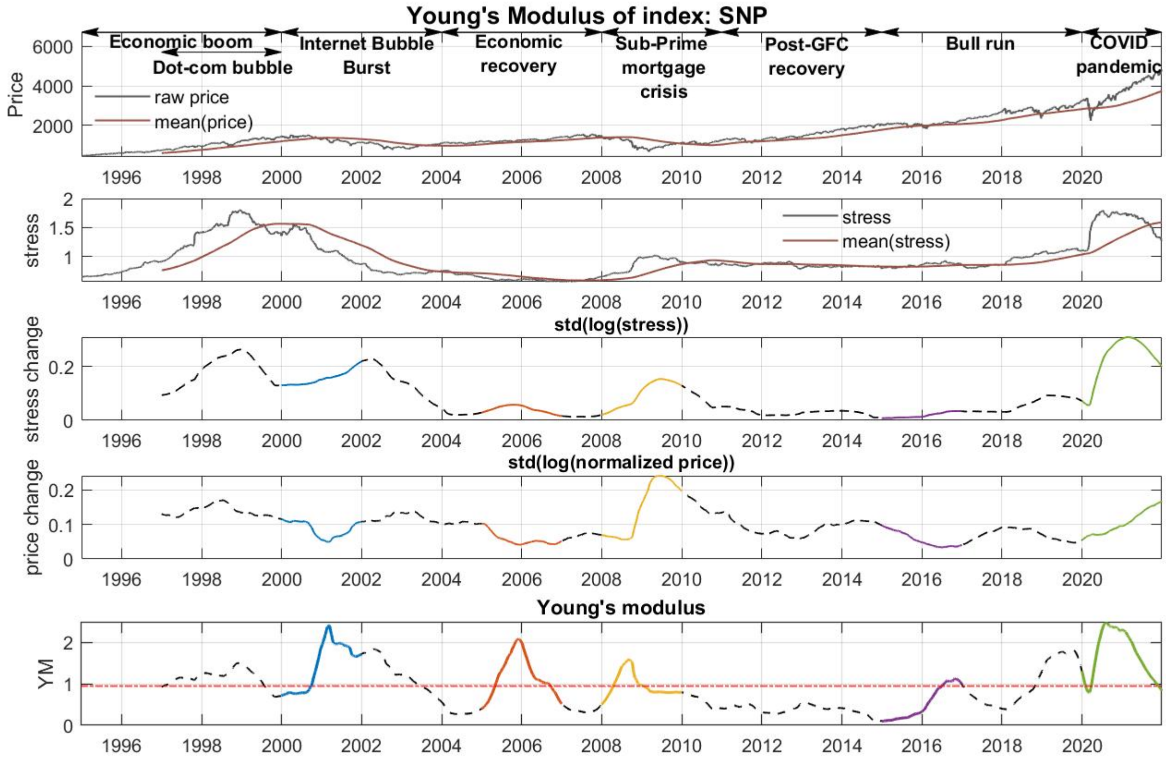


Figure 7.5: Raw price, overall financial stress and Young's Modulus for Finance (YMF) analysis of SNP index. In the last three panels, five distinct segments are identified, with blue, yellow, and green indicating three crises (the Internet Bubble Burst, the Sub-Prime mortgage crisis, and the COVID pandemic), while amber and purple denote two normal situations between crises. The red dashed line in the final graph indicates the level of the median value of the entire curve.

7.4.3 Young's Modulus on Finance of all indices

The YMF curves for 10 different indices/stocks are depicted in Figure 7.7 across three panels with a consistent y-scale for straightforward comparison. It is essential to recall that the YMF signifies the robustness of an index/stock in response to a crisis. Consequently, a higher YMF index during a crisis indicates a more resilient characteristic of the respective index/stock. Conversely, in normal situations with minimal disturbances, YMF curves appear smoother and more flattened compared to their counterparts in 'abnormal' conditions.

In Figure 7.7, the top panel displays the outcomes for the first group, presenting the results of four leading indices (DOW, NAS, RUS, and SNP). Generally, the four lines exhibit similar trends over time, showcasing the indices' adaptive responses to both 'nor-

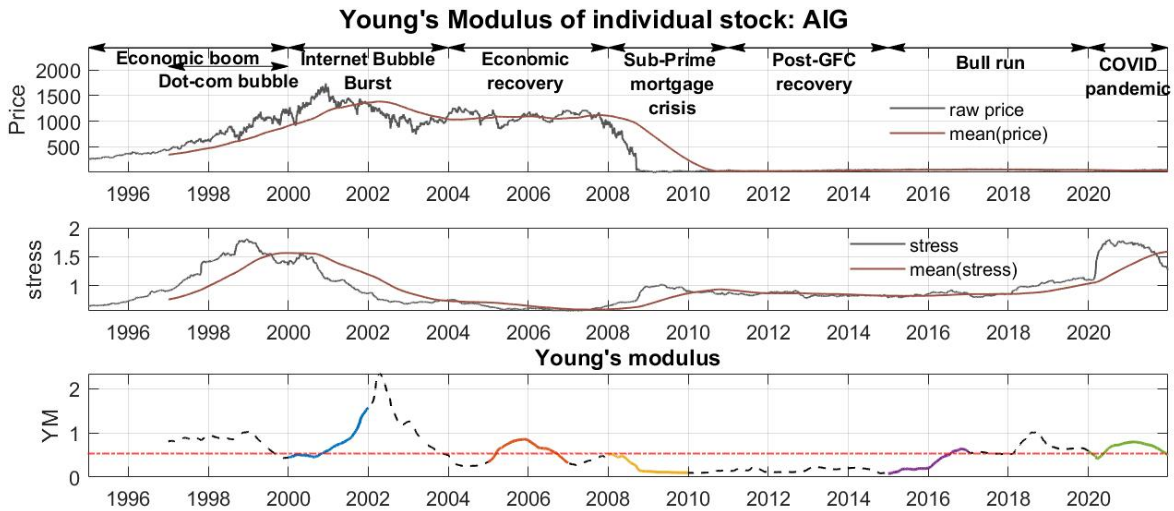


Figure 7.6: Raw price, overall financial stress and Young's Modulus for finance analysis of AIG stock. Five distinct segments are identified, with blue, yellow, and green indicating three crises (the Internet Bubble Burst, the Sub-Prime mortgage crisis, and the COVID pandemic) in the last panel, while amber and purple denote two normal situations between crises. The red dashed line in the final graph indicates the level of the median value of the entire curve.

mal' and 'abnormal' situations. Notably, during the Internet Bubble Burst (in 2002) and the COVID pandemic (in 2021), the YMF values for each index are higher than those during the Sub-Prime mortgage crisis (in 2009), implying that the overall financial market demonstrates greater robustness to the former two abnormal occurrences than the Sub-Prime mortgage crisis. In other words, the Sub-Prime mortgage crisis had a more severe impact on the financial market. Furthermore, the Internet Bubble Burst exhibited a global post-influence with the peak response occurring in 2006. When examining the four indices individually, DOW and SNP (in red and blue, respectively), encompassing diversified stocks and covering all major sectors, demonstrate higher and more stable indices than NAS and RUS (in yellow and sky blue, respectively). Given their large capitalization stocks and sector diversification, DOW and SNP exhibit stronger robustness, generally reflected in their higher YMF measures. Notably, during crises, DOW and SNP promptly react, as evidenced by their YMF curves initially increasing at the onset of crises (e.g., in 2001). In contrast, NASDAQ, primarily representing technology, telecommunications, and other high-growth stocks, displays lower robustness during the Internet Bubble Burst, a crisis specific to the technology sector, as anticipated in the yellow segment around 2002.

The second group in the middle panel comprises leading equities, representing the largest several stocks in the DOW, SNP, and NAS indices. These prominent equities include Apple Inc., Amazon.com, and Microsoft.

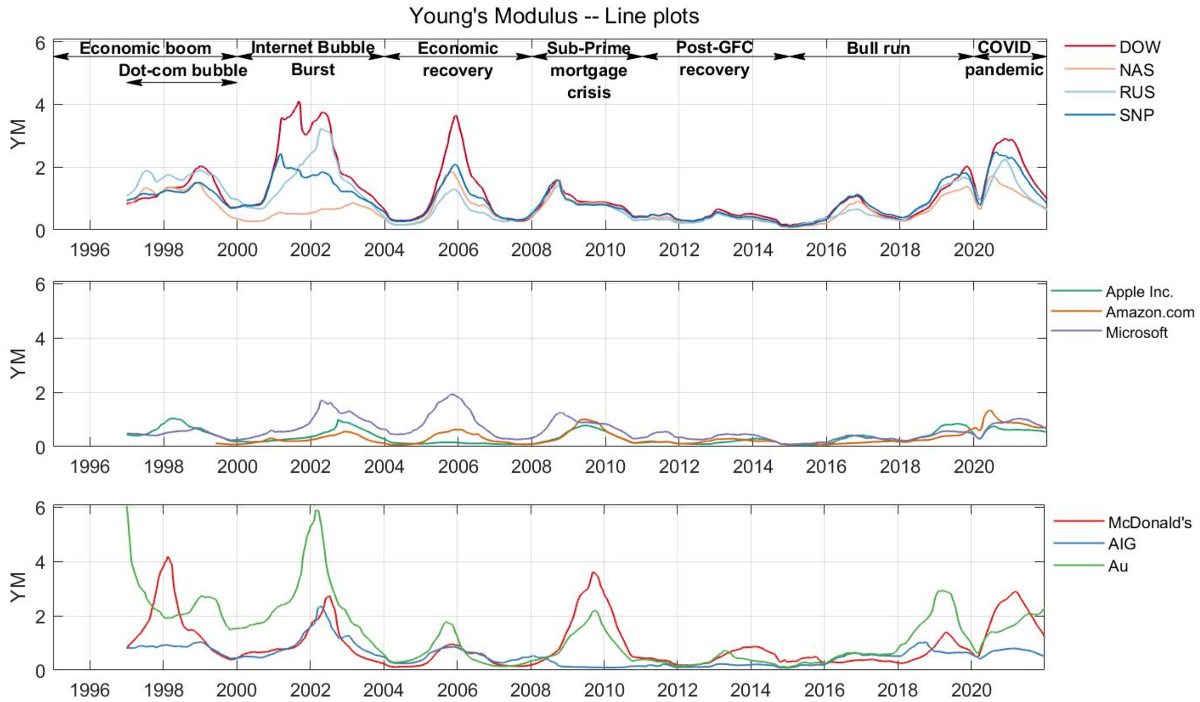


Figure 7.7: Line plots of Young's Modulus for Finance of 10 indices/individual stocks over 1997-2021. The initial set comprises indices, namely DJIA, NASDAQ, Russell 2000, and S&P 500. The second set encompasses technology giants, such as Apple Inc., Amazon.com, and Microsoft. The third cluster includes entities from diverse industries: McDonald's representing the fast-food sector, American International Group (AIG) from the insurance industry, and the price of gold (Au).

Owing to their substantial capitalization, the YMFs exhibit similarities to those of indices in the first panel but on a lower scale between 0 and 2. Throughout the Internet Bubble Burst and Sub-Prime mortgage crisis, Microsoft outperformed Apple Inc. and Amazon.com. The YMF of Microsoft in purple represents the first curve reacting to stress and attains the highest level during the crises. Conversely, Amazon.com in amber underwent early development from a bookstore to an 'everything store' during the Internet Bubble Burst, resulting in lower robustness. However, during the COVID pandemic, Amazon.com demonstrated stronger robustness than the other two companies due to its expanded delivery business and the introduction of food delivery services. Its size and the ability to undercut most smaller businesses on price positioned it as a leader in its sector.

The third group encompasses well-known brands from various industries and precious metals, as depicted in the bottom panel of Figure 7.7. These include representatives from the fast-food industry (McDonald's), the insurance sector (AIG), and a precious metal (Au). Notably, the gold price (in green) exhibits a discernible increase during widespread crises, ranging between 0 and 6. McDonald's, as a player in the basic needs industry, demonstrates commendable performance, especially during the Sub-Prime mortgage crisis and the COVID pandemic. Although its YMF is lower compared to the gold price, McDonald's resilience is noteworthy. The company, belonging to the consumer staples sector, also offers takeout and delivery services, which gained popularity during the COVID lockdown. Conversely, AIG, as detailed in Figure 7.6, suffered severe repercussions from the Sub-Prime mortgage crisis, resulting in a persistently low YMF level since 2008.

We conducted a horizontal comparison of YMFs across all indices and individual stocks over time. To facilitate a vertical assessment of the robustness of the ten entities in response to the same crises or periods, we introduced box plots of YMFs in Figure 7.8. Three distinct crises within a 2-year span were selected for detailed analysis: the Internet Bubble Burst (2000-2002), the Sub-Prime mortgage crisis (2008-2010), and the COVID pandemic (2020-2022). The left side of each graph enumerates the four indices, while the right side showcases the six individual stocks. This arrangement allows for a comprehensive comparison of their performance during these critical periods.

It is noteworthy that throughout all three periods analysed, the four major indices (DOW, NAS, RUS, and SNP), along with McDonald's (McD) and the gold price (Au), consistently demonstrated higher levels of robustness compared to the remaining individual stocks. During the Internet Bubble Burst, depicted in the top panel, the three technology giants (Apple Inc., Amazon.com, and Microsoft), and the technology-focused NASDAQ index, second from the left, exhibited increased vulnerability, indicative of weak robustness due to the specific nature of the crisis. The Sub-Prime mortgage crisis, illustrated in the middle panel, had the most profound impact on the overall financial system among the three crises, with most YMFs registering lower than 2, except for McDonald's and the precious metal gold. AIG, as anticipated, experienced the most significant influence and displayed the lowest response. Finally, during the COVID pandemic in the bottom panel, individual stocks on the right-hand side, such as Apple Inc., Amazon.com, Microsoft, and AIG, were adversely affected by the global pandemic, while indices on the left-hand side, McDonald's, and gold prices, showcased relatively stronger robustness.

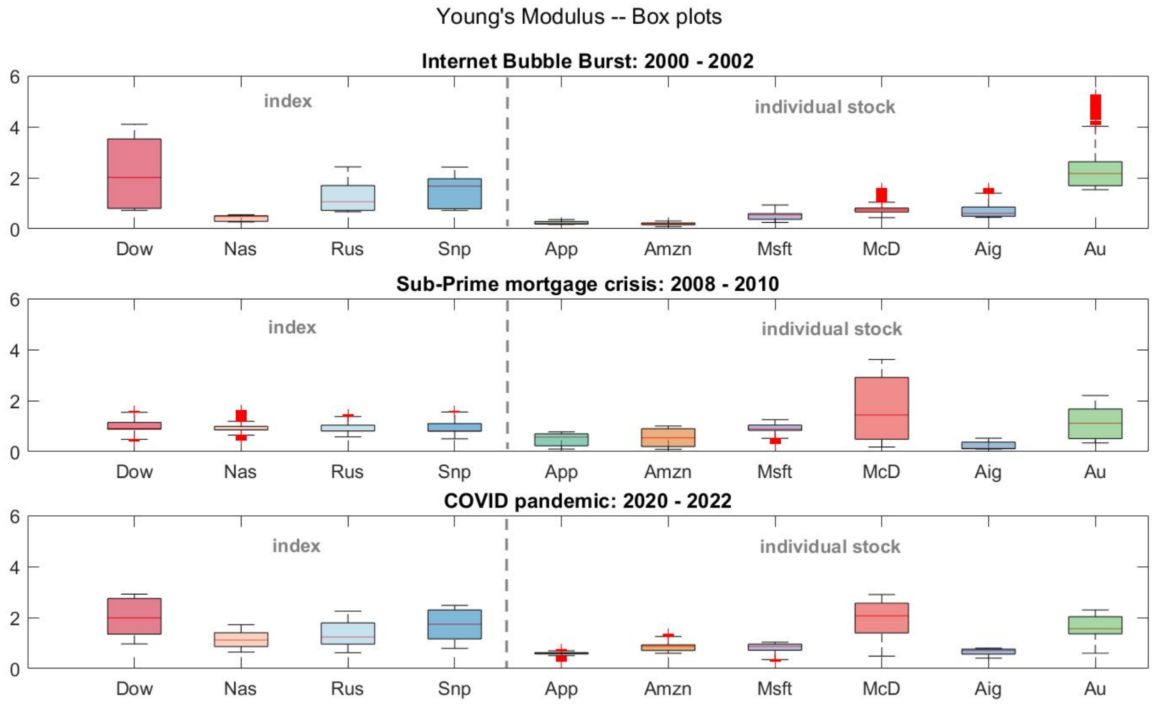


Figure 7.8: Box plots of Young's Modulus for Finance of 10 indices/individual stocks in response to three crises (Internet Bubble Burst, Sub-Prime mortgage crisis and COVID pandemic from top to bottom).

7.5 Conclusion

We have extended the concept of Young's Modulus from physics to financial analysis, so-called Young's Modulus for Finance (YMF), and have evaluated it over four financial indices and six individual stocks. We have found that: i) the YMF during a crisis is generally higher than during normal situations; ii) in normal circumstances, the YMF curves are flatter than that during a crisis; iii) different stresses/crises exhibit different levels of YMF. We have also found that the behaviours of YMFs of different indices/stocks to a certain crisis/stress are different; for example, AIG is more robust to the Internet Bubble Burst than the Sub-Prime mortgage crisis, while the YMFs of Apple and Amazon behave opposite to AIG. As a collection of hundreds of stocks, an index has been shown to have a higher YMF than individual stocks, that is, an index exhibits a more stable and robust behaviour than individual stocks.

Given the analogous behaviours of financial systems and bio-systems, future work of Young's Modulus may further borrow from physiological systems, for example, by further exploiting various sample entropies and other complexity science metrics.

Chapter 8

Conclusions and Future Work

The PhD research endeavoured to formulate computational methodologies for quantifying the structural complexity inherent in biological and financial systems. The key contributions and primary findings of this thesis are delineated in Figure 1.3 on page 9. Extensively exploring the foundational tenets of complexity science, this thesis has delved into the fundamental principles of complexity science, applying its insights to stress analysis in real-world systems, spanning biological systems in humans to economic systems in finance. Through the achievement of the three research objectives listed on page 8, this study has effectively achieved its overarching goal across both domains.

To enhance the reliability and stability of stress estimation in biological systems, a thorough investigation was conducted into established entropy-based methodologies. This encompassed an exploration of Multiscale Sample Entropy, Multiscale Fuzzy Entropy, Multiscale Cosine Similarity Entropy, Phase Entropy, Gridded Distribution Entropy, Multivariate Multiscale Sample Entropy, and Multivariate Multiscale Fuzzy Entropy. Identifying the limitations of existing entropy algorithms, we introduced novel contributions:

- 'ClassA Entropy': A novel univariate entropy methodology integrating the recently introduced ClassA Framework with classical Shannon Entropy. To evaluate its efficacy, this method was applied to an EEG data set comprising recordings from 10 brain death and 10 coma subjects, and its performance was benchmarked against other graph-based entropy techniques. The innovation showcased superior performance compared to existing algorithms, showcasing a computational edge through the elimination of distance calculations among embedding vectors. Moreover, the minimal tunable parameters required for ClassA Entropy render it both feasible and adaptable for widespread application across diverse systems, facilitating further investigations.

- 'Variational Embedding Multiscale Sample Entropy' (veMSE): A multivariate approach that combines Sample Entropy with an innovative fusion method. This method integrates a novel fusion technique that adjusts embedding dimensions across channels dynamically, aiming to mitigate the impact of signals in sub-channels with lower quality. The efficacy of veMSE was assessed using simulated data sets, evaluating aspects such as noise robustness, directionality of channels, and computational complexity. Subsequently, veMSE was applied to non-biological wind data and bio-signals based on HRT, and its performance was compared to that of the standard MMSE. Results demonstrated veMSE's capability to discern complex system features across large scales and with higher embedding dimensions.

- 'Multivariate Multiscale Cosine Similarity Entropy': A novel multivariate entropy method that incorporates Cosine Similarity Entropy. This approach merges angular distance-based algorithms with multivariate entropy, showcasing its effectiveness in discerning the circularity and structure of bivariate and trivariate systems. Through comprehensive demonstrations, the method exhibited a significant enhancement in detecting correlation degrees within multivariate systems, thereby contributing to a more nuanced understanding of complex interactions.

These novel approaches were rigorously evaluated using physiological and physical signals, as detailed in Chapters 3, 4, and 5.

In the realm of economic systems, an innovative exploration unfolded in Chapter 6.

Complexity analysis, including Moving-average Multivariate Multiscale Sample Entropy, Recurrence Quantification Analysis, and ALIS, showcased the potential for evaluating financial indices. Notably, entropy-based algorithms emerged as the optimal choice for stress quantification across various indices. Through these methods, the stress triggered by different crises, along with the distinctive properties of each index/individual stock, was effectively monitored using historical price data. Leveraging entropy methodologies, Catastrophe Plots were employed to visually depict the performance of each financial index under both normal and abnormal circumstances, thereby capturing critical breaking points indicative of market crashes.

In Chapter 7, a novel index titled 'Young's Modulus for Finance' took centre stage. Acting as a quantitative measure, YMF effectively monitored the behaviours of equities, encompassing four leading indices and six individual stocks, over the past 31 years, with a particular focus on periods of crises. Notably, it was observed that YMF values tend to be higher during crises compared to normal situations, while the YMF curves exhibited a flatter trajectory during normal circumstances, underscoring its significance as an analytical quantification tool within financial modelling.

Future work on structural complexity analysis in bio-systems will concentrate on the development of multivariate multiscale entropy. On the one hand, ClassA Entropy, as detailed in Chapter 3, has thus far been applied to a single EEG data set in a univariate context, its extension to multivariate scenarios holds promise for examining diverse signal types and a range of scenarios. On the other hand, the exploration of multiple features, such as randomness and structural richness, will be crucial for a comprehensive understanding of the nonlinear properties inherent in these systems. For example, the integration of a two-dimensional feature matrix employing both SampEn-based and CSE-based algorithms could offer a potential avenue for characterising the complex properties of the studied system in greater detail.

In the realm of financial systems, the present analysis conducted in this thesis relies on daily price data with a 4-year window for entropy calculations. Future investigations should further explore methods to ensure stable estimations with shorter data lengths. Additionally, while most of the indices and individual stocks examined in this thesis are US-based, future research could broaden its scope to include leading indices from around the world, facilitating a more comprehensive understanding of external stress dynamics for both catastrophe plots and the YMF index. Comparing financial indices across different countries or regions could also offer valuable insights. Furthermore, while the entropy techniques employed in the two chapters primarily focus on amplitude-based Sample Entropy, future studies should consider incorporating angular distance-based algorithms, such as Cosine Similarity Entropy, due to their reduced requirements on data length and other advantages, as discussed in Chapter 5. Additionally, future research should address the comparison between complexity-based analyses and traditional model-based financial research methodologies. Lastly, forthcoming endeavours should explore high-frequency data analysis to evaluate the reliability of short data length-based analyses and explore big data applications for the development of practical prediction models.

Bibliography

[1] T. Nakamura, T. Adjei, Y. Alqurashi, D. Looney, M. J. Morrell, and D. P. Mandic, “Complexity science for sleep stage classification from EEG,” in *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 4387–4394, 2017.

[2] W. Er and D. P. Mandic, “Dynamical complexity analysis of multivariate financial data,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8732–8736, 2013.

[3] W. von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, “Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV,” *Frontiers in Physiology*, vol. 8, article ID 360, 2017.

[4] N. F. Johnson, P. Jefferies, and M. H. Pak, *Financial market complexity.* Oxford: Oxford University Press, 2003.

[5] N. Johnson, *Simply complexity: A clear guide to complexity theory.* Simon and Schuster, 2009.

[6] S. M. Manson, “Simplifying complexity: A review of complexity theory,” *Geoforum*, vol. 32, no. 3, pp. 405–414, 2001.

[7] R. Clausius, “On a modified form of the second fundamental theorem in the mechanical theory of heat,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 12, no. 77, pp. 81–98, 1856.

[8] L. Glass, “Synchronization and rhythmic processes in physiology,” *Nature*, vol. 410, no. 6825, pp. 277–284, 2001.

[9] L. E. V. Silva, B. C. T. Cabella, U. P. da Costa Neves, and L. O. M. Junior, “Multiscale entropy-based methods for heart rate variability complexity analysis,” *Physica A: Statistical Mechanics and its Applications*, vol. 422, pp. 143–152, 2015.

[10] L. Faes, G. Nollo, and A. Porta, “Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique,” *Physical Review E*, vol. 83, no. 5, article ID 051112, 2011.

[11] M. Baranger, “Chaos, complexity, and entropy,” *New England Complex Systems Institute, Cambridge*, vol. 17, 2000.

[12] M. Mitchell and A. Complexity, *Complexity: A guided tour.* Oxford University Press, 2009.

[13] M. D. Costa, C. K. Peng, and A. L. Goldberger, “Multiscale analysis of heart rate dynamics: Entropy and time irreversibility measures,” *Cardiovascular Engineering*, vol. 8, no. 2, pp. 88–93, 2008.

[14] L. Han, F. Escolano, E. R. Hancock, and R. C. Wilson, “Graph characterizations from von Neumann entropy,” *Pattern Recognition Letters*, vol. 33, no. 15, pp. 1958–1967, 2012.

[15] T. Chanwimalueang and D. P. Mandic, “Cosine similarity entropy: Self-correlation-based complexity analysis of dynamical systems,” *Entropy*, vol. 19, no. 12, article ID 652, 2017.

[16] S. Ramdani, F. Bouchara, and A. Lesne, "Probabilistic analysis of recurrence plots generated by fractional Gaussian noise," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 8, 2018.

[17] C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and M. López, "Complexity analysis of the magnetoencephalogram background activity in Alzheimer's disease patients," *Medical Engineering & Physics*, vol. 28, no. 9, pp. 851–859, 2006.

[18] D. O'Hora, S. Schinkel, M. J. Hogan, L. Kilmartin, M. Keane, R. Lai, and N. Upton, "Age-related task sensitivity of frontal EEG entropy during encoding predicts retrieval," *Brain Topography*, vol. 26, no. 4, pp. 547–557, 2013.

[19] S. Pincus, "Approximate entropy as an irregularity measure for financial data," *Econometric Reviews*, vol. 27, no. 4-6, pp. 329–362, 2008.

[20] A. Rohila and A. Sharma, "Phase entropy: A new complexity measure for heart rate variability," *Physiological Measurement*, vol. 40, no. 10, article ID 105006, 2019.

[21] T. S. Parker and L. Chua, *Practical numerical algorithms for chaotic systems*. Springer Science & Business Media, 2012.

[22] C. L. Webber Jr and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol. 76, no. 2, pp. 965–973, 1994.

[23] J. M. Hausdorff, "Gait dynamics in Parkinson's disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 2, article ID 026113, 2009.

[24] T. Gautama, D. P. Mandic, and M. M. Van Hulle, "The delay vector variance method for detecting determinism and nonlinearity in time series," *Physica D: Nonlinear Phenomena*, vol. 190, no. 3-4, pp. 167–176, 2004.

[25] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[26] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[27] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277–283, 1988.

[28] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, p. 174102, 2002.

[29] J. M. Hausdorff, S. L. Mitchell, R. Firtion, C. K. Peng, M. E. Cudkowicz, J. Y. Wei, and A. L. Goldberger, "Altered fractal dynamics of gait: Reduced stride-interval correlations with aging and huntington's disease," *Journal of Applied Physiology*, vol. 82, no. 1, pp. 262–269, 1997.

[30] T. Gautama, D. P. Mandic, and M. M. Van Hulle, "Indications of nonlinear structures in brain electrical activity," *Physical Review E*, vol. 67, no. 4, article ID 046204, 2003.

[31] X. Wang, S. Si, and Y. Li, "Multiscale diversity entropy: A novel dynamical measure for fault diagnosis of rotating machinery," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5419–5429, 2020.

[32] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, "Escaping the curse of dimensionality in estimating multivariate transfer entropy," *Physical Review Letters*, vol. 108, no. 25, article ID 258701, 2012.

[33] L. C. Amarantidis and D. Abásolo, "Interpretation of entropy algorithms in the context of biomedical signal analysis and their application to EEG analysis in epilepsy," *Entropy*, vol. 21, no. 9, p. 840, 2019.

[34] M. Costa, C. K. Peng, A. L. Goldberger, and J. M. Hausdorff, “Multiscale entropy analysis of human gait dynamics,” *Physica A: Statistical Mechanics and its Applications*, vol. 330, no. 1-2, pp. 53–60, 2003.

[35] H. Xiao, T. Chanwimalueang, and D. P. Mandic, “Multivariate multiscale cosine similarity entropy,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5997–6001, 2022.

[36] H. Xiao, T. Chanwimalueang, and D. P. Mandic, “Multivariate multiscale cosine similarity entropy and its application to examine circularity properties in division algebras,” *Entropy*, vol. 24, no. 9, article ID 1287, 2022.

[37] F. C. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito, and I. Palamara, “Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer’s disease EEG,” *Entropy*, vol. 14, no. 7, pp. 1186–1202, 2012.

[38] A. L. Goldberger, L. A. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley, “Fractal dynamics in physiology: Alterations with disease and aging,” *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 1, pp. 2466–2472, 2002.

[39] T. Chanwimalueang, L. Aufegger, T. Adjei, D. Wasley, C. Cruder, D. P. Mandic, and A. Williamon, “Stage call: Cardiovascular reactivity to audition stress in musicians,” *PloS one*, vol. 12, no. 4, article ID e0176023, 2017.

[40] K. K. Dames, A. J. Lopes, and P. L. De Melo, “Airflow pattern complexity during resting breathing in patients with COPD: Effect of airway obstruction,” *Respiratory Physiology & Neurobiology*, vol. 192, pp. 39–47, 2014.

[41] C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and M. López, “Complexity analysis of the magnetoencephalogram background activity in Alzheimer’s disease patients,” *Medical Engineering & Physics*, vol. 28, no. 9, pp. 851–859, 2006.

[42] D. Abásolo, R. Hornero, C. Gómez, M. García, and M. López, "Analysis of EEG background activity in Alzheimer's disease patients with Lempel–Ziv complexity and central tendency measure," *Medical Engineering & Physics*, vol. 28, no. 4, pp. 315–322, 2006.

[43] E. Shumbayawonda, P. D. Tosun, A. Fernández, M. P. Hughes, and D. Abásolo, "Complexity changes in brain activity in healthy ageing: A permutation Lempel-Ziv complexity study of magnetoencephalograms," *Entropy*, vol. 20, no. 7, p. 506, 2018.

[44] P. D. Tosun, D. J. Dijk, R. Winsky-Sommerer, and D. Abasolo, "Effects of ageing and sex on complexity in the human sleep EEG: A comparison of three symbolic dynamic analysis methods," *Complexity*, vol. 2019, 2019.

[45] L. A. Lipsitz and A. L. Goldberger, "Loss of 'complexity' and aging: Potential applications of fractals and chaos theory to senescence," *Journal of American Medical Association*, vol. 267, no. 13, pp. 1806–1809, 1992.

[46] R. Menayo, A. Manzanares, and F. Segado, "Complexity, regularity and non-linear behavior in human eye movements: Analyzing the dynamics of gaze in virtual sailing programs," *Symmetry*, vol. 10, no. 10, article ID 528, 2018.

[47] B. Manor and L. A. Lipsitz, "Physiologic complexity and aging: Implications for physical function and rehabilitation," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45, pp. 287–293, 2013.

[48] R. Hari and M. V. Kujala, "Brain basis of human social interaction: From concepts to brain imaging," *Physiological Reviews*, vol. 89, no. 2, pp. 453–479, 2009.

[49] L. Ni, J. Cao, and R. Wang, "Analyzing EEG of quasi-brain-death based on dynamic sample entropy measures," *Computational and Mathematical Methods in Medicine*, vol. 2013, article ID 618743, 2013.

[50] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, "Looking at pictures: Affective, facial, visceral, and behavioral reactions," *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993.

[51] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, no. 3, pp. 354–381, 1996.

[52] G. G. Berntson, J. T. Cacioppo, and K. S. Quigley, "Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint," *Psychological Review*, vol. 98, no. 4, p. 459, 1991.

[53] A. Hemakom, T. Chanwimalueang, A. Carrión, L. Aufegger, A. G. Constantinides, and D. P. Mandic, "Financial stress through complexity science," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 6, pp. 1112–1126, 2016.

[54] C. S. Hakkio and W. R. Keeton, "Financial stress: What is it, how can it be measured, and why does it matter," *Economic Review*, vol. 94, no. 2, pp. 5–50, 2009.

[55] D. Wesselbaum, "Catastrophe theory and the financial crisis," *Scottish Journal of Political Economy*, vol. 64, no. 4, pp. 376–391, 2017.

[56] C. Wu, C. Kao, and C. F. Lee, "Time-series properties of financial series and implications for modeling," *Journal of Accounting, Auditing & Finance*, vol. 11, no. 2, pp. 277–303, 1996.

[57] D. E. Allen, M. McAleer, R. Powell, and A. K. Singh, "A non-parametric and entropy based analysis of the relationship between the VIX and S&P 500," *Journal of Risk and Financial Management*, vol. 6, no. 1, pp. 6–30, 2013.

[58] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Con-*

tributions to the Theory of Statistics, vol. 4, pp. 547–562, University of California Press, 1961.

[59] Y. Li, X. Wang, Z. Liu, X. Liang, and S. Si, "The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review," *IEEE Access*, vol. 6, pp. 66723–66741, 2018.

[60] C. Beck, "Generalised information and entropy measures in physics," *Contemporary Physics*, vol. 50, no. 4, pp. 495–510, 2009.

[61] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, pp. 479–487, 1988.

[62] Z. Liang, Y. Wang, X. Sun, D. Li, L. J. Voss, J. W. Sleigh, S. Hagihira, and X. Li, "EEG entropy measures in anesthesia," *Frontiers in Computational Neuroscience*, vol. 9, p. 16, 2015.

[63] A. N. Kolmogorov, "A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces," *Trudy Matematicheskogo Instituta imeni VA Steklova*, vol. 169, pp. 94–98, 1985.

[64] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, no. 2, article ID 021906, 2005.

[65] J. P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," *The Theory of Chaotic Attractors*, pp. 273–312, 1985.

[66] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, pp. 366–381, Springer, 1981.

[67] E. R. Deyle and G. Sugihara, "Generalized theorems for nonlinear state space reconstruction," *Plos One*, vol. 6, no. 3, p. e18295, 2011.

[68] F. Edition, A. Papoulis, and S. U. Pillai, "Probability, random variables, and stochastic processes," 2002.

[69] A. M. Unakafov and K. Keller, "Conditional entropy of ordinal patterns," *Physica D: Nonlinear Phenomena*, vol. 269, pp. 94–102, 2014.

[70] R. E. Blahut, *Principles and practice of information theory*. Addison-Wesley Longman Publishing Co., Inc., 1987.

[71] A. Velichko and H. Heidari, "A method for estimating the entropy of time series using artificial neural networks," *Entropy*, vol. 23, no. 11, p. 1432, 2021.

[72] S. Chen, S. Feng, W. Fu, and Y. Zhang, "Logistic map: Stability and entrance to chaos," in *Journal of Physics: Conference Series*, vol. 2014, p. 012009, IOP Publishing, 2021.

[73] M. Rostaghi and H. Azami, "Dispersion entropy: A measure for time-series analysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 610–614, 2016.

[74] K. Zografos and S. Nadarajah, "Expressions for rényi and Shannon entropies for multivariate distributions," *Statistics & Probability Letters*, vol. 71, no. 1, pp. 71–84, 2005.

[75] M. Riedl, A. Müller, and N. Wessel, "Practical considerations of permutation entropy: A tutorial review," *The European Physical Journal Special Topics*, vol. 222, no. 2, pp. 249–262, 2013.

[76] S. Raubitzek and T. Neubauer, "Combining measures of signal complexity and machine learning for time series analysis: A review," *Entropy*, vol. 23, no. 12, p. 1672, 2021.

[77] J. Li, J. Yan, X. Liu, and G. Ouyang, "Using permutation entropy to measure the changes in EEG signals during absence seizures," *Entropy*, vol. 16, no. 6, pp. 3049–3061, 2014.

[78] D. Li, X. Li, Z. Liang, L. J. Voss, and J. W. Sleigh, "Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia," *Journal of Neural Engineering*, vol. 7, no. 4, p. 046010, 2010.

[79] E. Ferlazzo, N. Mammone, V. Cianci, S. Gasparini, A. Gambardella, A. Labate, M. A. Latella, V. Sofia, M. Elia, and F. C. Morabito, "Permutation entropy of scalp EEG: A tool to investigate epilepsies: Suggestions from absence epilepsies," *Clinical Neurophysiology*, vol. 125, no. 1, pp. 13–20, 2014.

[80] A. Humeau-Heurtier, C. W. Wu, and S. D. Wu, "Refined composite multiscale permutation entropy to overcome multiscale permutation entropy length dependence," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2364–2367, 2015.

[81] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: A review," *Entropy*, vol. 14, no. 8, pp. 1553–1577, 2012.

[82] L. Xiao-Feng and W. Yue, "Fine-grained permutation entropy as a measure of natural complexity for time series," *Chinese Physics B*, vol. 18, no. 7, p. 2690, 2009.

[83] C. Bian, C. Qin, Q. D. Y. Ma, and Q. Shen, "Modified permutation-entropy analysis of heartbeat dynamics," *Physical Review E*, vol. 85, no. 2, p. 021906, 2012.

[84] Y. Yin and P. Shang, "Weighted permutation entropy based on different symbolic approaches for financial time series," *Physica A: Statistical Mechanics and its Applications*, vol. 443, pp. 137–148, 2016.

[85] H. Azami and J. Escudero, "Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation," *Computer Methods and Programs in Biomedicine*, vol. 128, pp. 40–51, 2016.

[86] Z. Chen, Y. Li, H. Liang, and J. Yu, "Improved permutation entropy for measuring complexity of time series under noisy condition," *Complexity*, vol. 2019, 2019.

[87] Z. Huo, Y. Zhang, L. Shu, and X. Liao, "Edge permutation entropy: An improved entropy measure for time-series analysis," in *Annual Conference of the IEEE Industrial Electronics Society (IECON)*, vol. 1, pp. 5998–6003, IEEE, 2019.

[88] P. Li, C. Liu, K. Li, D. Zheng, C. Liu, and Y. Hou, "Assessing the complexity of short-term heartbeat interval series by distribution entropy," *Medical & Biological Engineering & Computing*, vol. 53, pp. 77–87, 2015.

[89] D. Y. Lee and Y. S. Choi, "Multiscale distribution entropy analysis of heart rate variability using differential inter-beat intervals," *IEEE Access*, vol. 8, pp. 48761–48773, 2020.

[90] W. Fu, J. Tan, Y. Xu, K. Wang, and T. Chen, "Fault diagnosis for rolling bearings based on fine-sorted dispersion entropy and SVM optimized with mutation SCA-PSO," *Entropy*, vol. 21, no. 4, p. 404, 2019.

[91] Y. Li, X. Gao, and L. Wang, "Reverse dispersion entropy: A new complexity measure for sensor signal," *Sensors*, vol. 19, no. 23, p. 5203, 2019.

[92] H. Azami and J. Escudero, "Amplitude-and fluctuation-based dispersion entropy," *Entropy*, vol. 20, no. 3, article ID 210, 2018.

[93] X. Liu, A. Jiang, N. Xu, and J. Xue, "Increment entropy as a measure of complexity for time series," *Entropy*, vol. 18, no. 1, p. 22, 2016.

[94] Y. Li, Y. Yang, G. Li, M. Xu, and W. Huang, "A fault diagnosis scheme for planetary gearboxes using modified multi-scale symbolic dynamic entropy and mRMR feature selection," *Mechanical Systems and Signal Processing*, vol. 91, pp. 295–312, 2017.

[95] V. Rajagopalan and A. Ray, "Wavelet space partitioning for symbolic time series analysis," *Chinese Physics Letters*, vol. 23, no. 7, p. 1951, 2006.

[96] C. F. Hsu, S. Y. Wei, H. P. Huang, L. Hsu, S. Chi, and C. K. Peng, "Entropy of entropy: Measurement of dynamical complexity for biological systems," *Entropy*, vol. 19, no. 10, p. 550, 2017.

[97] M. A. Woo, W. G. Stevenson, D. K. Moser, R. B. Trelease, and R. M. Harper, "Patterns of beat-to-beat heart rate variability in advanced heart failure," *American Heart Journal*, vol. 123, no. 3, pp. 704–710, 1992.

[98] C. Yan, P. Li, C. Liu, X. Wang, C. Yin, and L. Yao, "Novel gridded descriptors of Poincaré plot for analyzing heartbeat interval time-series," *Computers in Biology and Medicine*, vol. 109, pp. 280–289, 2019.

[99] M. E. Cohen, D. L. Hudson, and P. C. Deedwania, "Applying continuous chaotic modeling to cardiac signal analysis," *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 5, pp. 97–102, 1996.

[100] D. Cuesta-Frau, "Slope entropy: A new time series complexity estimator based on both symbolic patterns and amplitude information," *Entropy*, vol. 21, no. 12, p. 1167, 2019.

[101] J. Yang, G. I. Choudhary, S. Rahardja, and P. Franti, "Classification of interbeat interval time-series using attention entropy," *IEEE Transactions on Affective Computing*, 2020.

[102] T. Gutjahr and K. Keller, "Ordinal pattern based entropies and the Kolmogorov–Sinai entropy: An update," *Entropy*, vol. 22, no. 1, p. 63, 2020.

[103] L. Faes, G. Nollo, and A. Porta, "Mechanisms of causal interaction between short-term RR interval and systolic arterial pressure oscillations during orthostatic challenge," *Journal of Applied Physiology*, vol. 114, no. 12, pp. 1657–1667, 2013.

[104] L. Faes, G. Nollo, and A. Porta, "Non-uniform multivariate embedding to assess the information transfer in cardiovascular and cardiorespiratory variability series," *Computers in Biology and Medicine*, vol. 42, no. 3, pp. 290–297, 2012.

[105] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[106] D. Abásolo, C. J. James, and R. Hornero, "Non-linear analysis of intracranial electroencephalogram recordings with approximate entropy and Lempel-Ziv complexity for epileptic seizure detection," in *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1953–1956, IEEE, 2007.

[107] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.

[108] D. Abásolo, R. Hornero, P. Espino, D. Alvarez, and J. Poza, "Entropy analysis of the EEG background activity in Alzheimer's disease patients," *Physiological Measurement*, vol. 27, no. 3, p. 241, 2006.

[109] R. Hornero, D. Abásolo, J. Escudero, and C. Gómez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1887, pp. 317–336, 2009.

[110] R. Alcaraz, D. Abásolo, R. Hornero, and J. Rieta, "Study of Sample Entropy ideal computational parameters in the estimation of atrial fibrillation organization from the ECG," in *Proceedings of the IEEE Conference on Computing in Cardiology*, pp. 1027–1030, 2010.

[111] T. Chanwimalueang, L. Aufegger, W. von Rosenberg, and D. P. Mandic, "Modelling stress in public speaking: Evolution of stress levels during conference presentations," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 814–818, 2016.

[112] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 2, pp. 266–272, 2007.

[113] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, “The pupil as a measure of emotional arousal and autonomic activation,” *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.

[114] S. Simons, P. Espino, and D. Abásolo, “Fuzzy entropy analysis of the electroencephalogram in patients with Alzheimer’s disease: Is the method superior to sample entropy?,” *Entropy*, vol. 20, no. 1, p. 21, 2018.

[115] G. Manis, M. D. Aktaruzzaman, and R. Sassi, “Bubble entropy: An entropy almost free of parameters,” *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 11, pp. 2711–2718, 2017.

[116] A. J. Mandell and M. F. Schlesinger, “Lost choices: Parallelism and topological entropy decrements in neurobiological aging,” *The Ubiquity of Chaos*, pp. 35–46, 1990.

[117] E. Pereda, A. Gamundi, R. Rial, and J. González, “Non-linear behaviour of human EEG: Fractal exponent versus correlation dimension in awake and sleep stages,” *Neuroscience Letters*, vol. 250, no. 2, pp. 91–94, 1998.

[118] M. Lo Giudice, G. Varone, C. Ieracitano, N. Mammone, G. G. Tripodi, E. Ferlazzo, S. Gasparini, U. Aguglia, and F. C. Morabito, “Permutation Entropy-based interpretability of convolutional neural network models for interictal EEG discrimination of subjects with epileptic seizures vs. psychogenic non-epileptic seizures,” *Entropy*, vol. 24, no. 1, p. 102, 2022.

[119] Y. C. Zhang, “Complexity and 1/f noise. A phase space approach,” *Journal de Physique I*, vol. 1, no. 7, pp. 971–977, 1991.

[120] M. Costa, A. L. Goldberger, and C. K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, no. 6, article ID 068102, 2002.

[121] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, p. 1091, 2006.

[122] M. D. Costa and A. L. Goldberger, "Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series," *Entropy*, vol. 17, no. 3, pp. 1197–1203, 2015.

[123] H. Azami, D. Abásolo, S. Simons, and J. Escudero, "Univariate and multivariate generalized multiscale entropy to characterise EEG signals in Alzheimer's disease," *Entropy*, vol. 19, no. 1, p. 31, 2017.

[124] H. Azami, A. Fernández, and J. Escudero, "Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis," *Medical & Biological Engineering & Computing*, vol. 55, pp. 2037–2052, 2017.

[125] A. Humeau-Heurtier, "The multiscale entropy algorithm and its variants: A review," *Entropy*, vol. 17, no. 5, pp. 3110–3123, 2015.

[126] W. J. Bosl, T. Loddenkemper, and S. Vieluf, "Coarse-graining and the Haar wavelet transform for multiscale analysis," *Bioelectronic Medicine*, vol. 8, no. 1, pp. 1–7, 2022.

[127] S. D. Wu, C. W. Wu, K. Y. Lee, and S. G. Lin, "Modified multiscale entropy for short-term time series analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 23, pp. 5865–5873, 2013.

[128] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[129] Y. Jiang, C. K. Peng, and Y. Xu, "Hierarchical entropy analysis for biological signals," *Journal of Computational and Applied Mathematics*, vol. 236, no. 5, pp. 728–742, 2011.

[130] Y. Lv, R. Yuan, T. Wang, H. Li, and G. Song, "Health degradation monitoring and early fault diagnosis of a rolling bearing based on CEEMDAN and improved MMSE," *Materials*, vol. 11, no. 6, article ID 1009, 2018.

[131] M. X. Xiao, H. C. Wei, Y. J. Xu, H. T. Wu, and C. K. Sun, "Combination of RR interval and crest time in assessing complexity using multiscale cross-approximate entropy in normal and diabetic subjects," *Entropy*, vol. 20, no. 7, p. 497, 2018.

[132] M. U. Ahmed, N. Rehman, D. Looney, T. M. Rutkowski, and D. P. Mandic, "Dynamical complexity of human responses: A multivariate data-adaptive framework," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, pp. 433–445, 2012.

[133] Y. Li, G. Li, Y. Yang, X. Liang, and M. Xu, "A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy," *Mechanical Systems and Signal Processing*, vol. 105, pp. 319–337, 2018.

[134] Y. Li, Y. Yang, X. Wang, B. Liu, and X. Liang, "Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine," *Journal of Sound and Vibration*, vol. 428, pp. 72–86, 2018.

[135] Y. Li, M. Xu, H. Zhao, and W. Huang, "Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis," *Mechanism and Machine Theory*, vol. 98, pp. 114–132, 2016.

[136] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy: A tool for complexity analysis of multichannel data," *Physical Review E*, vol. 84, no. 6, p. 061918, 2011.

[137] M. U. Ahmed, T. Chanwimalueang, S. Thayyil, and D. P. Mandic, "A multivariate multiscale fuzzy entropy algorithm with application to uterine EMG complexity analysis," *Entropy*, vol. 19, no. 1, article ID 2, 2017.

[138] S. T. Aung and Y. Wongsawat, "Prediction of epileptic seizures based on multivariate multiscale modified-distribution entropy," *PeerJ Computer Science*, vol. 7, article ID e744, 2021.

[139] H. Azami, A. Fernández, and J. Escudero, "Multivariate multiscale dispersion entropy of biomedical times series," *Entropy*, vol. 21, no. 9, article ID 913, 2019.

[140] X. Wang, S. Si, and Y. Li, "Variational embedding multiscale diversity entropy for fault diagnosis of large-scale machinery," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 3109–3119, 2021.

[141] A. Thul, J. Lechinger, J. Donis, G. Michitsch, G. Pichler, E. F. Kochs, D. Jordan, R. Ilg, and M. Schabus, "EEG entropy measures indicate decrease of cortical information processing in disorders of consciousness," *Clinical Neurophysiology*, vol. 127, no. 2, pp. 1419–1427, 2016.

[142] T. Gautama, D. P. Mandic, and M. M. Van Hulle, "Signal nonlinearity in fMRI: A comparison between bold and mion," *IEEE Transactions on Medical Imaging*, vol. 22, no. 5, pp. 636–644, 2003.

[143] M. U. Ahmed, L. Li, J. Cao, and D. P. Mandic, "Multivariate multiscale entropy for brain consciousness analysis," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 810–813, 2011.

[144] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy analysis," *IEEE Signal Processing Letters*, vol. 19, no. 2, pp. 91–94, 2011.

[145] H. Xiao and D. P. Mandic, "Variational embedding multiscale sample entropy: A tool for complexity analysis of multichannel systems," *Entropy*, vol. 24, no. 1, article ID 26, 2022.

[146] P. Guzik, J. Piskorski, T. Krauze, A. Wykretowicz, and H. Wysocki, "Heart rate asymmetry by Poincaré plots of RR intervals," *Biomedizinische Technik. Biomedical Engineering*, vol. 51, no. 4, pp. 272–275, 2006.

[147] C. K. Karmakar, A. H. Khandoker, and M. Palaniswami, "Phase asymmetry of heart rate variability signal," *Physiological Measurement*, vol. 36, no. 2, article ID 303, 2015.

[148] A. Porta, K. R. Casali, A. G. Casali, T. Gnecchi-Ruscone, E. Tobaldini, N. Montano, S. Lange, D. Geue, D. Cysarz, and P. Van Leeuwen, "Temporal asymmetries of short-term heart period variability are linked to autonomic regulation," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 295, no. 2, pp. R550–R557, 2008.

[149] C. Yan, P. Li, L. Ji, L. Yao, C. Karmakar, and C. Liu, "Area asymmetry of heart rate variability signal," *Biomedical Engineering Online*, vol. 16, no. 1, pp. 1–14, 2017.

[150] M. Brennan, M. Palaniswami, and P. R. Kamen, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 283, no. 5, pp. H1873–H1886, 2002.

[151] P. W. Kamen and A. M. Tonkin, "Application of the Poincaré plot to heart rate variability: A new measure of functional status in heart failure," *Australian and New Zealand Journal of Medicine*, vol. 25, no. 1, pp. 18–26, 1995.

[152] T. Adjei, W. von Rosenberg, T. Nakamura, T. Chanwimalueang, and D. P. Mandic, "The classA framework: HRV based assessment of SNS and PNS dynamics without

LF-HF controversies," *Frontiers in Physiology*, vol. 10, pp. 1–15, article ID 505, 2019.

[153] L. Li, Y. Xia, B. Jelfs, J. Cao, and D. P. Mandic, "Modelling of brain consciousness based on collaborative adaptive filters," *Neurocomputing*, vol. 76, no. 1, pp. 36–43, 2012.

[154] Y. Xia, L. Li, J. Cao, M. Golz, and D. P. Mandic, "A collaborative filtering approach for quasi-brain-death EEG analysis," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 182–196, 2011.

[155] Y. Li, X. Wang, S. Si, and S. Huang, "Entropy based fault classification using the Case Western Reserve University data: A benchmark study," *IEEE Transactions on Reliability*, vol. 69, no. 2, pp. 754–767, 2019.

[156] Y. Zhang, J. Li, and J. Wang, "Exploring stability of entropy analysis for signal with different trends," *Physica A: Statistical Mechanics and its Applications*, vol. 470, pp. 60–67, 2017.

[157] J. F. Valencia, A. Porta, M. Vallverdu, F. Claria, R. Baranowski, E. Orlowska-Baranowska, and P. Caminal, "Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 9, pp. 2202–2213, 2009.

[158] S. D. Wu, C. W. Wu, S. G. Lin, C. C. Wang, and K. Y. Lee, "Time series analysis using composite multiscale entropy," *Entropy*, vol. 15, no. 3, pp. 1069–1084, 2013.

[159] S. D. Wu, C. W. Wu, S. G. Lin, K. Y. Lee, and C. K. Peng, "Analysis of complex time series using refined composite multiscale entropy," *Physics Letters A*, vol. 378, no. 20, pp. 1369–1374, 2014.

[160] A. Porta, L. Faes, V. Bari, A. Marchi, T. Bassani, G. Nollo, N. M. Perseguini, J. Milan, V. Minatel, A. Borghi-Silva, A. C. M. Takahashi, and A. M. Catai, "Effect of age on complexity and causality of the cardiovascular control: Comparison

between model-based and model-free approaches," *PloS One*, vol. 9, no. 2, article ID e89463, 2014.

[161] I. Vlachos and D. Kugiumtzis, "Nonuniform state-space reconstruction and coupling detection," *Physical Review E*, vol. 82, no. 1, article ID 016207, 2010.

[162] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos*, vol. 5, no. 1, pp. 82–87, 1995.

[163] S. Makeig, C. Kothe, T. Mullen, N. Bigdely-Shamlo, Z. Zhang, and K. Kreutz-Delgado, "Evolving signal processing for brain-computer interfaces," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1567–1584, 2012.

[164] R. B. Govindan and H. Kantz, "Long-term correlations and multifractality in surface wind speed," *Europhysics Letters*, vol. 68, no. 2, article ID 184, 2004.

[165] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "Physiobank, Physiotoolkit, and Physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[166] N. Iyengar, C. K. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, vol. 271, no. 4, pp. R1078–R1084, 1996.

[167] D. C. Marinescu, *Classical and quantum information.* Academic Press, 2011.

[168] R. B. Govindan, J. D. Wilson, H. Eswaran, C. L. Lowery, and H. Preißl, "Revisiting sample entropy analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 376, pp. 158–164, 2007.

[169] D. Looney, T. Adjei, and D. P. Mandic, “A novel multivariate sample entropy algorithm for modeling time series synchronization,” *Entropy*, vol. 20, no. 2, article ID 82, 2018.

[170] A. Hemakom, V. Goverdovsky, D. Looney, and D. P. Mandic, “Adaptive-projection intrinsically transformed multivariate empirical mode decomposition in cooperative brain–computer interface applications,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, article ID, 20150199, 2016.

[171] L. Zhao, S. Wei, H. Tang, and C. Liu, “Multivariable fuzzy measure entropy analysis for heart rate variability and heart sound amplitude variability,” *Entropy*, vol. 18, no. 12, p. 430, 2016.

[172] J. Zheng, D. Tu, H. Pan, X. Hu, T. Liu, and Q. Liu, “A refined composite multivariate multiscale fuzzy entropy and Laplacian score-based fault diagnosis method for rolling bearings,” *Entropy*, vol. 19, no. 11, p. 585, 2017.

[173] H. Amoud, H. Snoussi, D. Hewson, M. Doussot, and J. Duchene, “Intrinsic mode entropy for nonlinear discriminant analysis,” *IEEE Signal Processing Letters*, vol. 14, no. 5, pp. 297–300, 2007.

[174] T. Gautama, D. P. Mandic, and M. M. Van Hulle, “A differential entropy based method for determining the optimal embedding parameters of a signal,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 6, pp. VI–29, IEEE, 2003.

[175] M. Paluš, “Testing for nonlinearity using redundancies: Quantitative and qualitative aspects,” *Physica D: Nonlinear Phenomena*, vol. 80, no. 1-2, pp. 186–205, 1995.

[176] P. Berens, “Circstat: A matlab toolbox for circular statistics,” *Journal of Statistical Software*, vol. 31, pp. 1–21, 2009.

[177] A. S. Soofi and L. Cao, "Nonlinear forecasting of noisy financial data," *Modelling and Forecasting Financial Data: Techniques of Nonlinear Dynamics*, pp. 455–465, 2002.

[178] M. Zanin, "Forbidden patterns in financial time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 18, no. 1, 2008.

[179] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[180] A. C. García, T. Chanwimalueang, G. G. Calvi, A. Hemakom, R. M. Ricós, and D. P. Mandic, "Modelling economic stress through financial systemic balance index," in *2016 IEEE International Conference on Digital Signal Processing (DSP)*, pp. 565–569, IEEE, 2016.

[181] R. McCarty, "The fight-or-flight response: A cornerstone of stress research," in *Stress: Concepts, Cognition, Emotion, and Behavior*, pp. 33–37, Elsevier, 2016.

[182] S. Pincus and R. E. Kalman, "Irregularity, volatility, risk, and financial market time series," *Proceedings of the National Academy of Sciences*, vol. 101, no. 38, pp. 13709–13714, 2004.

[183] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5-6, pp. 237–329, 2007.

[184] Y. Wu, P. Shang, and Y. Li, "Multiscale sample entropy and cross-sample entropy based on symbolic representation and similarity of stock markets," *Communications in Nonlinear Science and Numerical Simulation*, vol. 56, pp. 49–61, 2018.

[185] Y. Zhang and P. Shang, "Multivariate multiscale distribution entropy of financial time series," *Physica A: Statistical Mechanics and its Applications*, vol. 515, pp. 72–80, 2019.

[186] Y. Yin and P. Shang, "Asymmetric asynchrony of financial time series based on asymmetric multiscale cross-sample entropy," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 3, 2015.

[187] Y. Zhang and P. Shang, "Refined composite multiscale weighted-permutation entropy of financial time series," *Physica A: Statistical Mechanics and its Applications*, vol. 496, pp. 189–199, 2018.

[188] H. Niu and J. Wang, "Quantifying complexity of financial short-term time series by composite multiscale entropy measure," *Communications in Nonlinear Science and Numerical Simulation*, vol. 22, no. 1-3, pp. 375–382, 2015.

[189] H. Yang, "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 339–347, 2010.

[190] Y. Cao, W. W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," *Physical Review E*, vol. 70, no. 4, p. 046217, 2004.

[191] F. Strozzi, E. Gutierrez, C. Noè, T. Rossi, M. Serati, and J. M. Zaldívar, "Measuring volatility in the nordic spot electricity market using recurrence quantification analysis," *The European Physical Journal Special Topics*, vol. 164, pp. 105–115, 2008.

[192] M. C. Ruiz, A. Guillamón, and A. Gabaldón, "A new approach to measure volatility in energy markets," *Entropy*, vol. 14, no. 1, pp. 74–91, 2012.

[193] R. Thom, "Stabilité structurelle et morphogénèse [structural stability and morphogenesis]," *Paris, France: Interéditions*, 1972.

[194] R. Birău, "The implications of catastrophe theory for stock market forecasting," *ARPN Journal of Science and Technology*, vol. 2, no. 5, pp. 360–363, 2013.

[195] D. Yang, D. Mu, and M. Yao, "Analysis of financial crisis: A perspective based on catastrophe theory," in *International Conference on Management and Service Science*, pp. 1–4, IEEE, 2009.

[196] J. Baruník and M. Vosvrda, "Can a stochastic cusp catastrophe model explain stock market crashes?," *Journal of Economic Dynamics and Control*, vol. 33, no. 10, pp. 1824–1836, 2009.

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

[198] 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.

[199] 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.

[200] E. Ofek and M. Richardson, "Dotcom mania: The rise and fall of internet stock prices," *The Journal of Finance*, vol. 58, no. 3, pp. 1113–1137, 2003.

[201] Y. Demyanyk and O. Van Hemert, "Understanding the subprime mortgage crisis," *The Review of Financial Studies*, vol. 24, no. 6, pp. 1848–1880, 2011.

[202] T. T. Y. Alabdullah, E. R. Ahmed, and M. I. Nor, "The world declining economy and coronavirus pandemic: Systems should be continued," *Russian Journal of Agricultural and Socio-Economic Sciences*, vol. 102, no. 6, pp. 89–96, 2020.

[203] R. G. Hawtrey, *The art of central banking.* Routledge, 2012.

[204] I. N. Stewart and P. L. Peregoy, "Catastrophe theory modeling in psychology.," *Psychological Bulletin*, vol. 94, no. 2, pp. 336–362, 1983.

[205] E. C. Zeeman, "Catastrophe theory structural stability in physics," 1979.

[206] I. M. McNally, “Contrasting concepts of competitive state-anxiety in sport: Multidimensional anxiety and catastrophe theories,” *Athletic Insight: The Online Journal of Sport Psychology*, vol. 4, no. 2, pp. 10–22, 2002.

[207] E. C. Zeeman, “Catastrophe theory in brain modelling,” *International Journal of Neuroscience*, vol. 6, no. 1, pp. 39–41, 1973.

[208] K. S. Suslick, “Encyclopedia of physical science and technology,” *Sonoluminescence and Sonochemistry*, pp. 1–20, 2001.

[209] A. Vaidya and K. Pathak, “Mechanical stability of dental materials,” in *Applications of Nanocomposite Materials in Dentistry*, pp. 285–305, Elsevier, 2019.

[210] H. Xiao, Y. L. Xu, and D. P. Mandic, “Complexity-based financial stress evaluation,” *arXiv preprint arXiv:2212.02281*, 2022.

[211] Y. Yuan, Y. Li, and D. P. Mandic, “A comparison analysis of embedding dimensions between normal and epileptic EEG time series,” *The Journal of Physiological Sciences*, vol. 58, no. 4, pp. 239–247, 2008.

[212] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki, “A comparative study of synchrony measures for the early diagnosis of Alzheimer’s disease based on EEG,” *NeuroImage*, vol. 49, no. 1, pp. 668–693, 2010.

Appendix A

Case Study: Self-recorded 41-years data set

Owing to the declining functionality of biological systems due to ageing and disease, Complexity Loss Theory (CLT) and the modified structural complexity theory suggest an increase in the correlation and deterministic components of the system. This results in a reduced degree of randomness and a loss of stochastic elements [5, 15, 38, 39, 45]. Numerous studies have successfully explored the potential for changes in complexity in humans as an early indicator of physical and psychological diseases, including Alzheimer's disease, epilepsy, and Huntington's disease [29, 37, 108, 173, 211, 212]. However, proving and investigating the variance in complexity due to ageing remains challenging. While the complexity differences based on EEG and MEG signals in ageing have been mentioned and presented in a few studies, the existing data set relies on groups of individuals of varying age groups, introducing individual complexity distinctions within subjects for comparison [43, 44, 106]. To this end, in this case study, we present a visualization of complexity changes correlated with age, utilizing a distinctive data set of self-recorded databases that spans over 40 years and includes multi-channel variables.

The data were self-measured by a clinically healthy male subject, with an average recording frequency of five times daily. Raw data collection spanned from 1967 to 2011, encompassing the subject's early 20s to his late 60s.

Three rapidly changing variables were selected for the analysis, including Heart Rate (HRT) in beats/min, Systolic Blood Pressure (SBP) in mmHg, and Diastolic Blood Pressure (DBP) in mmHg. This section primarily focuses on HRT as the main analytical subject.

Prior to complexity analysis, several preprocessing steps were implemented. Invalid recordings were excluded, and the raw data was truncated, removing the initial three years due to monthly gaps. Subsequently, daily average values were calculated to standardize and harmonize the frequency for all variables. Following that, interpolation was applied to the scaled signal using the built-in function, interp1, in MATLAB to address days with missing recordings. Consequently, the newly scaled data, covering the timeline from May 1970 to June 2011 (i.e., 41 years), comprises 15,028 data points with a frequency of one data point per day.

The subsequent section will unveil the findings from the preliminary complexity analysis, encompassing univariate single-scale entropy, univariate multiscale entropy, and multivariate single-scale entropy. This will be succeeded by a concise discussion of the Complexity-Entropy profile applied to the present data set, along with insights into future research directions.

A.1 Univariate Single-scale Analysis

The first and second panels of Figure A.1 showcase features in the time domain, specifically the mean and standard deviation, while the third and fourth panels depict features in the information domain, such as Sample Entropy and Cosine Similarity Entropy. In each graph, the red line represents the trend value of the feature, and the blue bar illustrates the dynamics of the feature with a step of 1 week (equivalent to 7 data points).

To ensure reliable estimations from the Entropy-based calculation and to mitigate the impact of unknown dynamical changes resulting from mild diseases, a window length of N = 2000 (approximately 5.5 years) was selected.

Figure A.1 illustrates the trends in HRT features over time. The mean of HRT exhibits a gradual and mild increase during middle age, followed by a noticeable decrease in senior age. Simultaneously, the second panel depicting the standard deviation shows a peak before 2006, succeeded by a decline to a lower level. In the third panel, Sample Entropy (SampEn) characterizes the degree of randomness and visualizes the nonlinear feature of the biological system based on HRT data [26]. The bottom panel illustrates Cosine Similarity Entropy (CSE), which quantifies the self-correlation of the system [15]. Broadly, as age advances, the randomness of HRT decreases, while correlation increases, signifying the loss of adaptivity and self-organization. Notably, both SampEn and CSE curves identify a peak before 2006, with CSE showing the earliest rise.

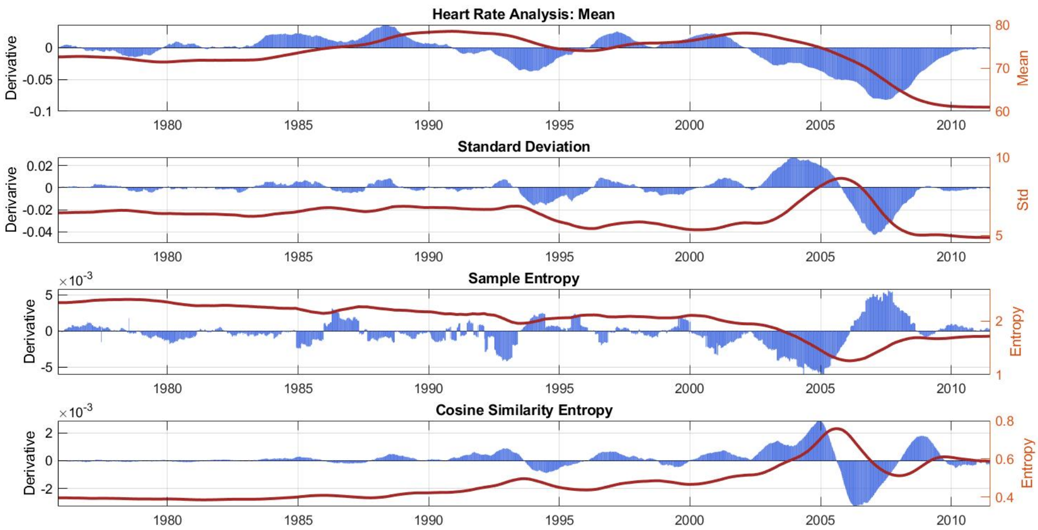
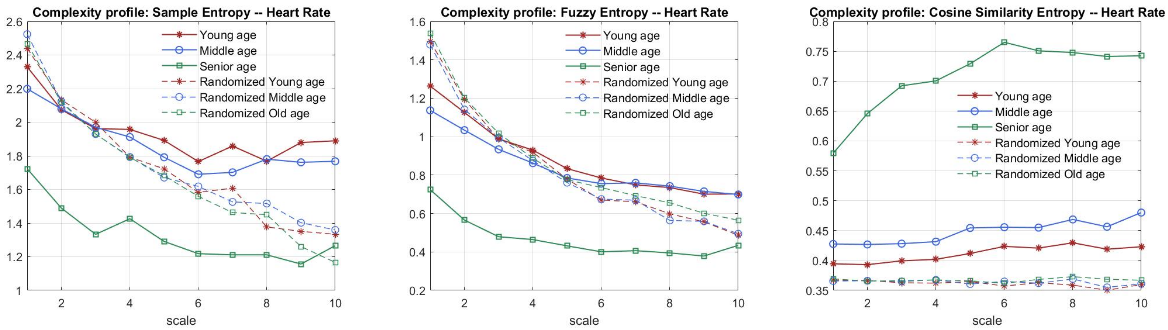


Figure A.1: Line plots depicting features of the scaled Heart Rate (HRT) signal from 1976 to 2011, utilizing a window length of 2000 and a step of 7 days. The top two panels illustrate the mean and standard deviation separately, while the bottom two panels showcase the Sample Entropy and Cosine Similarity Entropy, respectively. In each figure, the red line represents the feature values, with an axis label on the right-hand side, while the blue bars depict the dynamics of the features, accompanied by an axis label on the left side.

A.2 Univariate Multiscale Analysis

To elucidate the concealed pattern of structural complexity across diverse scales, we applied Multiscale Sample Entropy (MSE), Multiscale Fuzzy Entropy (MFE), and Multiscale Cosine Similarity Entropy (MCSE) across three distinct age-hoods: young age (1971 - 1975, i.e., 24-28 years old), middle age (1986-1991, i.e., 39-44 years old), and senior age (2006-2010, i.e., 59-63 years old). The algorithms utilized scales ranging from 1 to 10, and the complexity profiles are depicted in Figure A.2. Within each graph, red lines illustrate the curves of young age, blue lines represent the states of middle age, and green lines portray the changes in senior ages. Solid curves depict the complexity profile of recorded data, while dashed lines delineate features of the shuffled data set.



(**a**) MSE (**b**) MFE (**c**) MCSE

Figure A.2: Complexity profiles for three distinct age-hood data sets depicted through (a) Multiscale Sample Entropy, (b) Multiscale Fuzzy Entropy, and (c) Multiscale Cosine Similarity Entropy. Default parameters were configured as m = 2, N = 2000, l = 1, r = 0.25 for MSE/MFE, and r = 0.08 for MCSE.

Upon observation, MSE in the left panel effectively distinguishes the curve of senior age from the other two; however, it inaccurately attributes a lower degree of randomness to the real HRT data in the solid green line compared to the uncorrelated shuffled data in dashed lines. MFE in the middle panel illustrates a smoother profile than MSE but provides limited enhancement in complexity estimation performance. MCSE in the right panel produces optimal results: 1) the curves depicting self-correlation in solid lines for all recorded data surpass those of shuffled data in dashed lines; 2) the levels of structural complexity for different ages are distinctly separated, with senior age exhibiting the highest correlation, followed by middle and young ages. 3) the gap between curves stabilizes when the scale reaches 7.

A.3 Multivariate Single-Scale Analysis

Taking into account the available multivariate recordings, we implemented Multivariate Entropy algorithms based on HRT in conjunction with Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP). This approach explores correlations and patterns among variables from multiple channels, aiming for a more effective and reliable estimation despite the short data length [36, 136].

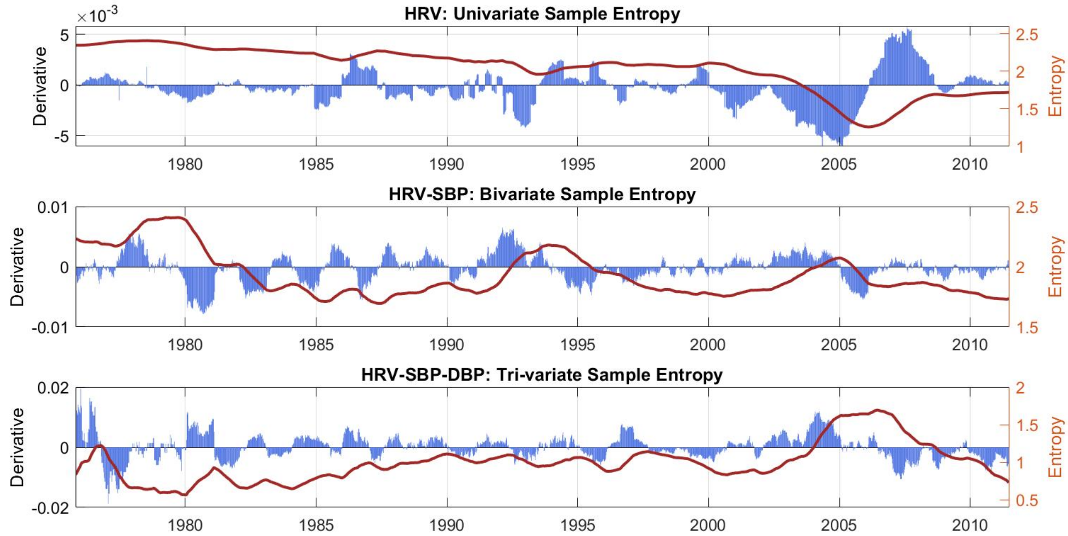


Figure A.3: Complexity estimation of the current biological system based on Heart Rate, Systolic Blood Pressure, and Diastolic Blood Pressure, processed by Univariate Sample Entropy in the top panel, Bivariate Sample Entropy in the middle panel, and Trivariate Sample Entropy in the bottom panel. Each figure depicts the red line representing the feature values with an axis label on the right-hand side, while the blue bars illustrate the dynamics of features with an axis label on the left side. The window length is set at 2000 with a step of 7 days. Default parameters for the standardized segment were configured as m = 2, l = 1, r = 0.25.

Multivariate Sample Entropy (Mv-SE) and Multivariate Cosine Similarity Entropy (Mv-CSE) were implemented, and the results are presented in Figure A.3 and A.4, respectively. The top panel in Figure A.3 features Univariate Sample Entropy, quantifying the degree of randomness, which is generally observed to decrease with ageing. In contrast, Multivariate Sample Entropy in the same figure provides insight into the structural complexity, revealing the determinism of the system. The bottom panel successfully unveils the increasing trend of correlation due to the loss of adaptivity. Notably, despite the partly common information shared between Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP), the addition of the third channel significantly enhances structural complexity estimation, as demonstrated in the transition from the middle to the bottom panel based on Sample Entropy.

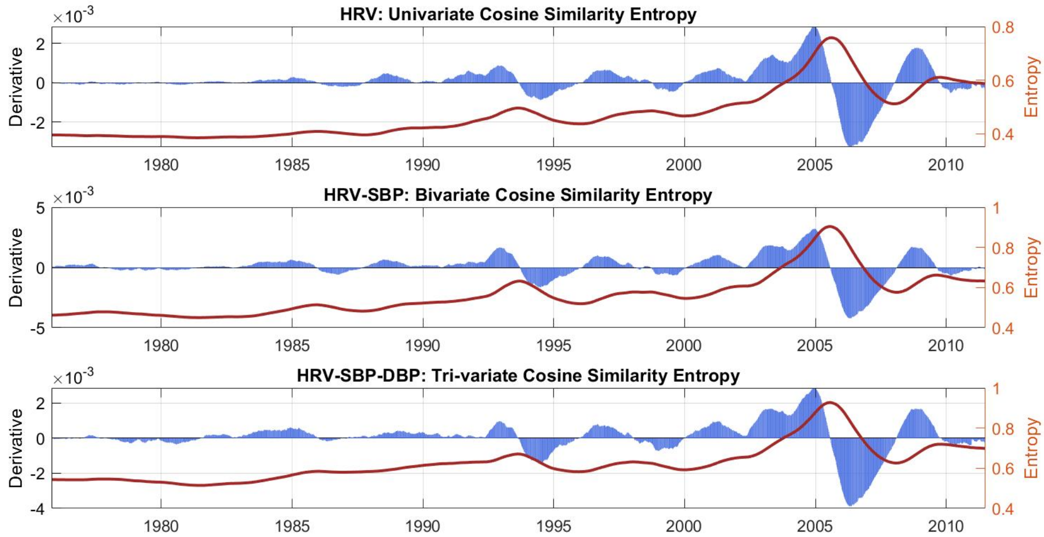


Figure A.4: Complexity estimation of the current biological system based on Heart Rate, Systolic Blood Pressure, and Diastolic Blood Pressure, processed by Univariate Cosine Similarity Entropy in the top panel, Bivariate Cosine Similarity Entropy in the middle panel, and Trivariate Cosine Similarity Entropy in the bottom panel. Each figure illustrates the red line representing the feature values with an axis label on the right-hand side, while the blue bars describe the dynamics of features with an axis label on the left side. The window length is set at 2000 with a step of 7 days. Default parameters for Univariate analysis were configured as m = 2, l = 1, r = 0.08; for Bivariate analysis, r = 0.225; and for Trivariate analysis, r = 0.287.

Results provided by Multivariate Cosine Similarity Entropy are illustrated in Figure A.4. Cosine Similarity Entropy is known for revealing the self-correlation of a complex system, indicative of structural complexity. Univariate analysis in the top panel demonstrates good performance with a rising trend as age increases. Notice that incorporating the second channel (SBP in the middle panel) enhances the variation of entropy values with ageing, as indicated by the larger y-axis scale on the right-hand side. However, the inclusion of the third channel (DBP) contributes to limited improvement due to the correlated information between the two blood pressure measures, as depicted in the bottom panel.

A.4 Nonlinar Features in Humans

Drawing inspiration from the well-known Complexity-Entropy profile and its extended version in biological systems, as presented in Figure 1.1 & 1.2(a) in Chapter 1, we utilized the entropy results derived from the HRT data set. Subsequently, we graphed the Complexity (representing self-correlation/determinism, given by Univariate Cosine Similarity Entropy) VS Entropy (representing randomness/stochasticity, given by Univariate Sample Entropy) profile in Figure A.5.

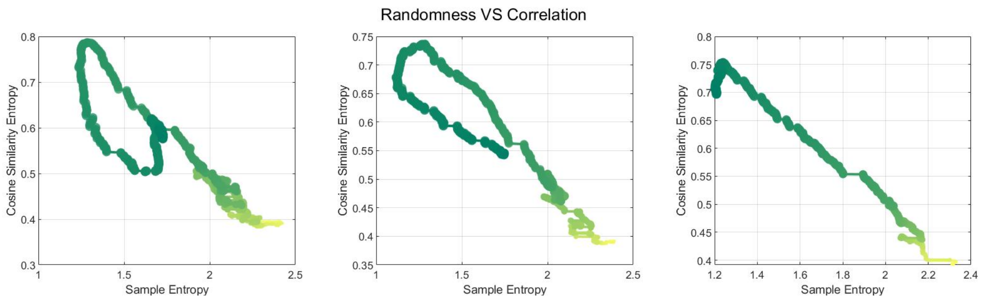


Figure A.5: Scatter plots illustrating the relationship of randomness and correlation based on the provided HRT data set for Univariate Sample Entropy and Univariate Cosine Similarity Entropy. Each graph is colour-coded, ranging from the lightest to the darkest shade, covering the period from 1976 to 2011. Different window lengths for entropy algorithms were applied from left to right graphs, with values of N = 2000, 3000, and 5000, respectively.

In Figure A.5, three scatter plots depict the relationship between Randomness and correlation of the HRT data using entropy algorithms. The left plot is based on a window length of 2000, representing more than 5 years of data. The middle graph is calculated with a length of 3000 (around 8 years), and the right one is given a length of 5000, reflecting 13 years of data. Consistent with the predictions of Complexity Loss Theory [45] and self-correlation theory [15], the reciprocal relationship of randomness and correlation is evident. With increasing age, the loss of randomness and the increase in correlation manifest from the right bottom corner to the left top corner in each graph, delineated by colours ranging from light to dark.

Future work will concentrate on several key aspects: 1) identifying the significance of nonlinear features among young, middle, and senior ages using the non-segmented data set; 2) exploring the complex distinctions of states in various time frames (morning/afternoon/evening) across different age-hoods; and 3) implementing enhanced Multivariate Multiscale Entropy in the current analysis to achieve a higher level of reliability in complexity estimation.