Investigation of Causality Pattern Structure for the Exploration of Dynamic Time-Varying Behaviour

Leo Carlos-Sandberg

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Supervisor: Professor Christopher D. Clack, Department of Computer Science

Department of Computer Science

University College London

July 13, 2023

I, Leo Carlos-Sandberg, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

The analysis of time-varying interactions within multivariate systems has seen a great deal of interest within the last decade, with the international oil market being an archetypal and important system that demonstrates this behaviour. However, unlike work on static systems, research on time-varying systems rarely leverages specific information on the inter-system interactions for understanding the systems temporal dynamics. This thesis utilises this information to present methodologies for new descriptions of these systems, focussing on the international oil market. This is achieved via three experiments.

The first experiment expands upon the state-of-the-art methodologies for investigating these systems; complex networks. Presenting a novel complex network approach that encodes the transitional behaviour of the dynamic interactions. The work introduces: two transition metrics, a complex network, and various metrics and properties of this network. Using this approach it is shown that for the international oil market the evolution favours staying in similar causality patterns before switching to a new group of similar patterns.

The second experiment puts forth two novel paradigms for the evolution of a dynamic multivariate system, and from these paradigms the principle features that drive the systems dynamics. It is also shown demonstrated that a p-value representation of causality can improve the description of the dynamics. Through dimensional reduction based on these paradigms and prediction of the systems future states on the reduced system, that the international oil market dynamics are well captured by the total change in causality of the system.

The third experiment further explores and validates a hypothesis of the inter-

national oil markets dynamics based on the findings of the first two experiments. Proposing a approach for the formal definition of such system dynamics, and applying this to the proposed hypothesis. This hypothesis is then validated via a novel clustering approaches to determine that the international oil markets state is primarily contained within clusters that slightly vary around central causality patterns, and that the system does not follow a repeated gradual change when transitioning between these clusters.

This work allows for a more detailed and alternative description of a system's dynamic behaviour than those given by other current methodologies.

Impact statement

The major contribution of this thesis is novel for descriptions of the behaviours of dynamically changing Granger causality networks in multivariate systems generally, and specifically for the international oil market, increasing the understanding of this system past its current state. To produce these findings this thesis also contributes methodologies and frameworks for the investigation, description, and validation of the dynamic behaviour of a multivariate system from the perspective of its time-varying interaction structure.

These methodologies present 1) a complex network representation, amendable to analysis, that has descriptive power in regards to the dynamic change in the system's interaction structure; 2) a framework for the further analysis and dimensionality reduction of the system's dynamic behaviour, paradigms for the evolution of a system, and the prediction of future causality patterns of that system, based on information gathered from the presented complex network; 3) a framework for the construction and investigation of specific descriptions of the dynamic behaviour of a system in terms of the interaction structure of the system. Thus, by offering time-varying behavioural descriptions and methodologies that include information from the interaction structure, this work extends past current literature.

The interest in the dynamic behaviour of multivariate systems has seen a dramatic increase in the last decade across a wide range of disciplines in both academia and industry, ranging from neuroscience to climatology and economics. In particular significant interest has been shown in finance by investors, policymakers, regulators, and risk analysts. Due to the breadth of this interest, the work presented here can be significantly impactful in a wide range of areas and to various practitioners.

The work provides several benefits within academia. Firstly, this work provides a foundation for the use of dynamic behaviour of the interaction structure of a system for a more detailed description of a system's behaviour, a relatively unexplored academic area. Secondly, the work presents frameworks and methodologies that can be expanded to conduct further/targeted analysis, allowing for application to many systems and investigation of specific behaviours; proposals for some such extensions are provided. Thirdly, this work contributes explicitly to the body of literature investigating the international oil market by providing novel findings and descriptions for the dynamic behaviour within sample data of this market.

Outside of academia, there has been growing interest in complex system analysis, particularly in finance, where traditional analysis of markets has become widespread and accessible to many, so the search to find new information has become focused on more complex behaviours. This has lead to the usage of increasingly complex methods, such as machine learning approaches. The work presented here can be considered an addition to the suite of methodologies applied in this area, with an example use case being that of feature selection. However, more directly, this work can be applied in many areas such as risk management and policy design; proposals for some such use cases are provided. The work presented here generally can facilitate analysis to provide a more detailed description of complex time-varying systems. With the inherent complexity and interconnectedness of real-world systems and an ever-increasing interest in these systems, this work's application and usage can be seen as interdisciplinary. As such, the dissemination of this work is achieved through the submission of literature to an interdisciplinary scholarly peer-reviewed journal.

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Chris Clack, for his guidance and support throughout the PhD. I would also like to thank Prof. Tomaso Aste, for his kindness in reading this thesis and the advice that he gave. In addition I would like to acknowledge the academic support provided by the Financial Computing and Analytics Centre for Doctoral Training and, more broadly, University College London. I would also like to recognise the financial support of the Engineering and Physical Sciences Research Council (EPSRC).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation for this research

Nowadays it is commonplace for real-world systems to generate and store continuous data steams. It is often the case that this data is naturally indexed by time; being common in fields such as (i) finance, e.g. daily spot prices; (ii) biology, e.g. monitory techniques such as magnetic resonance imaging (MRI); and (iii) meteorology, with surface-air temperatures. This data is often naturally expressible as time series, a sequence of data points indexed by the time of observations, which occur at successive and, in many cases, equally spaced points in time. Hence time series data can be seen to have an embedded discrete temporal ordering. This natural ordering makes analysis of time series data distinct from non-temporal data, with time series analysis considering the series of data points, rather than individual values at different instances.

A single time series, not under the effect of others, is a univariate system and has been the focus of much research, solidifying it as a mature field. This research has lead to data mining tools for the discovery of hidden patterns and the forecasting of the series [5, 6, 7, 8].

Though there is significant interest in univariate data, most real-world data comes from multivariate system's, as discussed above, where confounding variables and interactions between variables occur. With the maturing of the field of univariate research, there has been an increased interest in the analysis of multivari-

ate systems, spurred on by ever-increasing sources of complex datasets [9]. Work on multivariate systems is a broad field and represents the high level area of interest of this thesis, and understanding phenomena from this complex data often crucially depends on the analysis of multivariate time series [10]. This analysis takes many forms in the literature, with a notable separation between approaches for bi-variant systems (those with only two variables) and multivariate systems with greater than two variables.

When considering a multivariate system, a defining feature is the connections/interactions between the variables generating the time series, with these interactions allowing for variables to drive and influence each other. Furthermore, these interactions can cause emergent phenomena and behaviours to occur within the systems. Hence understanding these interaction dynamics allows for a deeper understanding of the system and desirable analysis such as forecasting. However, the interactions and connections are not known for many real-world systems, particularly those with a high level of complexity. A natural approach to discovering these connections is to perturb a variable and observe the reaction of another; in reality, this approach is often unfeasible due to physical or legal limitations. Due to this, researchers often rely on statistical measures to determine the connections between time series. A popular method, particularly in the field of economics, is the Granger causality test, which presents a statistical measure of causality between variables [11, 12, 13, 14], with a network of variables interacting via Granger causality being referred to as a ”causality pattern” [3]. Following this the area of interest for this thesis can be further specified to multivariate systems and where the variables are interacting via Granger causality.

When these causal interactions (which for this thesis will be referred to interchangeably as causality) are considered, it is predominantly as static measures between the variables. There is a sizeable body of research focused on the static causality structure within a multivariate system [15, 16, 17]. However, it is well accepted that the dynamics governing these interactions are seldom genuinely static and often are time-varying. Furthermore, the behaviours of these interactions tend

to be nonlinear in nature [9]. Therefore the temporal dimension of this data can be seen as crucial for their understanding. The area of interest for this thesis can hence be given as multivariate systems interacting via Granger causality, where the causal interactions change throughout time.

This property of many real-world systems has led to an increasing interest in time-varying interaction dynamics within multivariate systems. Interest in this field ranges from predicting stocks, feature selection for models, risk management, regulation, and early warning of instability, with scholars, investors, and regulators, to name a few, all benefiting from advances in the field. This interest is expressed in two influential papers on network instability by Bardoscia et al. [18] and May [19], using spectral theory to investigate network complexity. These papers exemplify a core reason why the study of these systems and their behaviour is important, demonstrating that instability in a multivariate system of stable components can evolve (as an emergent property) and that this instability is connected to the size and topology of the connectivity of the system. Due to the nature of this research, diverse fields like finance, climatology, fluid dynamics, neurophysiology, engineering, and economics, have demonstrated the potential of this area of research [20, 9]. However, characterising complicated dynamics from multivariate time series data can still be considered a fundamental problem of continuing interest [21].

A specific system that has received significant attention in this field is the international oil market, with several authors demonstrating the presence of time-varying nonlinear behaviour within the interaction dynamics of this market, illustrating the complexity apparent within the comprehension of the market's behaviour and prediction of future aspects of it [3, 22, 23, 24, 25, 26, 4, 27, 28]. Understanding the nature of the interactions in the international oil market can be seen as necessary for portfolio managers, investors, and speculators to construct diversified portfolios, make good risk management decisions and enhance returns. It is also important for policymakers and regulators, who are interested in avoiding disturbances in the stock market caused by oil price shocks, particularly during periods of financial turmoil, and in the implementation of policies [22].

The body of research focusing on the time-varying interaction dynamics of a multivariate system has seen a significant increase in contribution over the last decade (discussed further in Chapter 2). This work has taken a number of forms, including comparing causality patterns of specific time segments (see Long et al. [29]) and leveraging the established field of complex networks, which has successfully investigated nonlinear dynamics in a variety of applications from a broad range of disciplines [30, 31, 32, 33, 9]. These applications of complex networks to time-varying interaction in a multivariate system can be seen as the state-of-the-art within this field [34, 35, 36], with these networks being able to be analysed in-depth to provide insight into the original time series [33]. Of particular note is the work of Qi et al. [1], who use a mode based network, Yu et al. [37], who use a visibility graph based network, and Jiang et al. [3], who use a transition order based network (discussed further in Chapter 2).

Though the recent research into this field has been insightful, it is still an active area of research, with many open questions and ample scope for new developments [9, 33, 38].

The work applying complex networks to this problem can be broadly summarised as creating a complex network of complex networks. This can be achieved by producing a new network, where the nodes are representations of each causality pattern that the system visits throughout time, the edges are the directed temporal ordering of these causality patterns (i.e. two nodes are connected if their causality patterns occur consecutively in time), and the edges are weighted by the frequency of the temporal ordering throughout the history of the system.

In general, these complex network methods construct a form for the data that is amenable to description, with the goal of uncovering behaviour of the original system of time series. This descriptive approach often employs classic network metrics, properties, and comparison (discussed further in Chapter 3) to draw conclusions on the dynamics of the system. Compared to many static (non-time-varying) descriptions of the interactions of multivariate systems, the time-varying descriptions achieved via the complex network approach can be seen as abstracted or at a

"higher-level". This higher level can be seen to lose information on the underlying transformation of the system, specifically the actual change that the causality pattern undergoes (i.e. the change of the links in the causality pattern). In broad terms, the static analysis is comparable to a network of interactions, and the dynamic analysis is comparable to a network of static networks. Therefore the interaction structure can be seen as "low-level" behaviour, and the movement of this structure can be seen as "high-level" behaviour. As mentioned, current methods focus on this high-level behaviour and hence miss information contained in the low-level behaviour and how this changes through the movement of the structure. By taking this high level view a number of implicit assumptions become embedded in the analysis and conclusions of these types of approaches. Specifically this network construction embeds the assumptions that the important network features are the connectivity of the nodes and the frequency of the edges. Though these features may be important these assumptions add limitations for systems that break them, such as a system where poorly connected nodes or less travelled edges are of great importance (e.g. a market going from high to low stability, this edge and node may be rarely visited but of great impact). As such the causality pattern can be see as a key feature of a system, and hence its change can be seen as a key feature of the systems evolution. This is the gap in prior research that this thesis aims to fill, by investigating methodologies for constructing time-varying descriptions of a multivariate system's behaviour that include information about the changes in the low-level behaviour. This nuance of the systems dynamics is thereby captured in investigation and behavioural descriptions.

1.2 Research objectives

Given the current interest in the behaviour of time-varying interactions within multivariate systems and their wide interdisciplinary impact, this thesis investigates an aspect of this domain. There are many questions open to investigate, however this thesis puts forth the hypothesis that information describing how the structure of the interactions in the multivariate system change from one time step to another can

be used to give new insight into the behviour of the system. This can hence be stated as the following research question *" how can behaviours in the structure of the interactions be harnessed to gain a richer description of the system's dynamic behaviour?"*. Hence other research questions are out of the scope of this work.

To answer this research question one can not trivially "add" this information to the system, there are a number of factors that must be considered. The current complex network approaches produce a network of relative simplicity that facilitates its usability (these networks are widely applicable to current research methods), yet through adding additional information to these networks they can easily lose this usability. Therefore, a solution most overcome these competing aspects. An approach that could be taken to mitigate this is creating specific solutions for the exact problem and data set at hand, however this approach is time consuming, non-scalable, and creates analysis and conclusions that are challenging to compare to other problem specific solutions. Thus a more general approach is needed. However, to create a general approach it is challenging not to include implicit assumptions about the key features, leading to conclusions being driven by these assumptions and not the actual underlying data (these can be seen as simply re-introducing the research question). Finally, the analysis of complex systems (especially dynamic complex systems) needs be explainable. This explainability is needed to correctly draw conclusions on the underlying system, and to make the analysis more amenable to a wider audience, which is important in the multidisciplinary field of complex dynamic systems. It is not immediately clear how a solution to this research question can be found while also retaining the current level of explainability.

Based on these challenges the aim of the work presented in this thesis is to provide tools and frameworks for discovering new descriptions of a system's dynamic behaviour, while maintaining explainability and usability.

As a case study, this work investigates the international oil market, as mentioned in the previous section and expanded upon in Chapter 2.4. This hypothesis is investigated through three different conceptual approaches to utilising this low-level information; complex networks, time series, and configuration space.

This split aligns with the statements by Takeishi and Yairi [39], and Li, Pedrycz, and Jamal [40], that investigations of multivariate time series can be often broken into three conceptual approaches: (i) graph-based (network) methods, (ii) transformation-based (dimension reduction) methods, and (iii) generative model-based (state-space) methods. These approaches are most suited for different forms of investigation of a dynamic system: behaviour discovery, analysis of dynamics, and behaviour description (each of these discussed further below) respectively. This thesis draws on each of these approaches sequentially through three logical groupings of contributions. First, a new complex network is proposed to directly address the research question and allow for behavioural discovery. Second, a new question is asked "are there archetypal important behaviours in this space?"; to answer this question two new paradigms of behavioural evolution are proposed. Third, an answer is sought for the question "given a set of independent conclusions drawn from multiple analyses, can their combined hypothesis be validated?"; to do this, a new methodology for exploring hypotheses of the dynamic causality behaviour is proposed. These three contributions are presented as the following research questions:

1. **Can time-varying behaviour of a system be more fully explored via the inclusion of information on the transition properties between causality patterns?**

   The meaningful information that best describes the behaviour of a system of interacting components throughout time is predominantly viewed in literature as the transition order and frequency of interaction (causality) patterns. Therefore when transforming these systems into a form more readily analysed (such as a complex network) the order and frequency of transitions is the information maintained and brought to the forefront. This thesis posits that though the order and frequency of transition is very informative, so is the nature of the transition (i.e. what change in the system occurred in this transition). This information allows for describing the temporal dynamics of the system in terms of its interaction dynamics.

This type of description of the systems dynamics allows for a deeper understanding and context that is important for many use cases, such as investigating instability and changes in the financial markets [41, 42].

The type of system form and analysis needed to generate this kind of description are difficult to achieve due to the complex nature of the systems in question. To investigate the dynamics of these systems some level of abstraction is required, this abstraction reducing the information explicitly expressed.

To investigate this research question a novel network representation is proposed that incorporates two novel metrics embodying the nature of the transition between temporally consecutive causality patterns.

The introduction of this network also addresses an important aspect of the analysis of time-varying interacting system. Currently, after transformation into the amenable complex network form the analysis techniques applied are typically network specific, but non-specific to the underlying system (do not contain knowledge of what the nodes and edges represent), there by limiting the meaningful information that can be extracted from the network.

To allow for a more complete analysis of the novel network, five novel properties/metrics are proposed based on knowledge of the underlying system and the information of it captured in the network.

This network and analysis is validated on a synthetic data set and employed to find novel information on sample data from the international oil market.

2. **Can a meaningful one-dimensional time series representation of a time-varying system be constructed based upon that system's overall transition behaviour?**

   An advantage of a one-dimensional time series representation of a system's history is its exact rendering of the temporal transition order, especially for the recent history of the system. This allows for an understanding of where in the system's dynamic trajectory the current state is and can be important for predicting future states of the system. The process of reducing a

multi-dimensional time series (a system of interacting components) to a one-dimensional representation inherently loses/obscures information. This loss of information is an unavoidable aspect of dimensionality reduction, however what information is lost and what information is brought forward is dependent on the methods used. This thesis posits that knowledge of the transition behaviour between causality patterns of the system can be used to reduce the dimensionality of the system in such a way as to maintain information relevant to the time-varying nature of the system, and hence produce a one-dimensional representation that exhibits similar dynamic properties to the original system. Reversely the demonstration of the dimensionality reduction successfully capturing the systems dynamics demonstrates the accuracy of the description used to reduce it.

To investigate this hypothesis a novel analysis, and dimensionality reduction methodology are proposed. These are based on two novel paradigms for the time-varying dynamics of a system, with the predominate paradigm for a system determining the dimensionality reduction applied.

The proposed analysis and dimensionality reduction is applied to sample data from the international oil market to create a one-dimensional time series representation. This representation is then investigated through a proposed prediction algorithm, as well as three other alternative approaches.

3. **Can a configuration space be utilised to test a hypothesis of specific system's dynamic behaviour, described at both low- and high-level?**

   When investigating a system using multiple separate analyses, a singular idea of the system's dynamics is often reached through combining hypotheses from each. However, it can difficult, due to the breadth of approaches used for analysis, to formalise these many hypotheses into a single model hypothesis that can be tested. This thesis posits that a configuration space can offer an environment well suited to investigating complex hypotheses on the dynamics of a time-varying system.

To investigate this a novel methodology is proposed. First, a novel representation is proposed for the temporal trajectory of a time-varying system in a multi-dimensional configuration space. Second, using sample data from the international oil market as a case study a hypothesis for its dynamics is proposed and formalised using the configuration space. Third, a novel clustering algorithm is created to see the system's alignment to the proposed hypothesis model. This methodology results in a novel conclusion on the dynamic behaviour of the international oil market.

The above research questions give novel insight into the dynamics of the international oil market, as well as providing a framework that can be more generally applied to investigate the temporal dynamics any complex system providing a novel description. These new descriptions can be utilised for complex and important use cases, such as early warning systems of instability or market changes. This work expands upon the literature through a focus on the dynamics contained within the transitions between causality patterns, in three core areas: behaviour discovery, analysis of dynamics, and behaviour description.

1.3 Thesis structure

The structure of this thesis is as follows:

Chapter 2 covers the background and current literature on characterizing the dynamics of multivariate systems. Going into detail on interaction measures and the cutting edge complex network representations of the dynamics of these systems. Chapter 2 also introduces the international oil market as a focus of the analysis in this thesis. With a sample data set of this market being analysed and transformed to be amenable to investigation throughout this thesis.

Chapter 3 presents novel analysis of the interactional oil market by introducing a methodology for the inclusion of information from the causality pattern transitions in a complex network, and proposing several properties and metrics for a system under this methodology. A validation of this methodology and metrics is also given on synthetic data.

Chapter 4 presents a novel time series representation for the international oil market, and demonstrates this representation for the prediction of interaction states. This is done through the introduction of a methodology for the construction of a time series based upon paradigms of the dynamics of the underlying causality patterns.

Chapter 5 presents and tests a formal definition of the temporal behaviour of the international oil market. This is achieved through the introduction of a configuration space methodology for the formal construction of temporal behaviour within multivariate interacting systems. This methodology is validated against a synthetic data set.

Chapter 6 summarises the presented methodologies and analysis results of the research and outlines potential future extensions.

1.4 Related publication

The following publication is related to this thesis:

Leo Carlos-Sandberg and Christopher D. Clack. Incorporation of causality structures to complex network analysis of time-varying behaviour of multivariate time series. *Scientific Reports* **11**, 18880 (2021).

https://doi.org/10.1038/s41598-021-97741-2

Chapter 2

Background, Literature Review, and Data

This chapter introduces the necessary background and literature to give context to this research and notes specific work that this thesis is expanding upon, giving motivation and an overview to research focussed on the dynamic behaviour of interacting systems, paying attention to the recent adoption of network representations. In addition, presented is a general overview of interaction measures, and a discussion on Granger causality and its application in time-varying multivariate systems. This chapter concludes by introducing the international oil market and some initial analysis on it. The aim of this chapter is to contextualise this thesis and give an idea of where it falls within this branch of research. Therefore more specific literature on content presented later in this work will be discussed in the chapter in which it appears.

2.1 Characterizing behaviour from time series

Times series data describes many real-world complex systems of interest; due to this, the investigation of these systems often focuses on the analysis of these time series. This has led to a large body of research building up around the analysis of time series.

One branch of this research focuses on complex systems where multiple time series describe interacting components. Within this further research focuses on

these time series systems where the interactions are non-static and change with time [43, 44].

The work relevant to this thesis is concerned with analysing dynamic interaction relationships between time series in a time-varying multivariate system. Due to the breadth of the literature, this chapter does not intend to give an exhaustive list of every approach but rather an overview of the state of the field and the research directions it has taken. For further insight, readers are directed to the recent and comprehensive literature reviews by Silva *at al* [33], and Zou *et al.* [9], which cover aspects of this field in more depth.

Univariate systems

Much of the work on multivariate time series is inspired by or expands on methodologies for univariate time series (where the dynamic behaviour of one time series is investigated). This univariate analysis takes many different conceptual approaches to this problem, however, a large section of this work focuses on comparison/interaction between regions of the time series [45, 46]. To investigate the dynamics within a univariate time series, an approach that has become mainstream is mapping the time series to a complex network representation. This allows for the structure of the network to be assessed to understand the dynamic characteristics of the original time series. This methodology coincides with a decade of rapid development within the study of complex networks [47, 48, 48, 49], leading to the emergence of multidisciplinary methodologies for the use of complex networks for characterising these dynamic systems. These complex networks are briefly discussed below, with the reader directed to Gao *et al.* [49] who present a review of the significant work in this field.

Significant literature exists covering multiple methodologies for the transformation of a time series to a complex network, such as using phase spaces [21, 50], visibility graphs [51, 52, 53, 54, 55], network patterns [56, 57, 58], multiple resolutions [59], and other specific approaches[60, 61]. Of particular note are [23]:

- Visibility graphs [62, 63] (discussed further in Section 2.3), with a number of versions of this approach existing: visibility graph [64], horizontal visibil-

ity graph (HVG) [65], and multiscale limited penetrable horizontal visibility graph (MLPHVG) [54].

- Phase space reconstruction [50, 66, 38] (discussed further in Chapter 5).

Multivariate systems

As mentioned above, this work on univariate time series has inspired research on dynamic multivariate time series. This move to multivariate time series comes from questioning the assumption in univariate work that the time series can be considered independently from other factors. This is untrue for many real-world cases, with confounding variables and complex multivariate behaviour being frequent and important. In particular, multivariate systems approaches leverage the univariate work on complex network representations. These complex network representations gain their power from the new perspective they give to dynamic systems, allowing for a potential reduction in complexity and amenability for analysis on dynamic behaviours.

The most intuitive usage of complex networks in multivariate systems is to apply them to a static system; generally, these representations take the approach of mapping components (variables/time series) in the system to nodes in the network and interactions/relationships between components to edges within the network [67, 10, 68, 69, 70, 71, 72]. This approach has been expanded in various ways to construct methodologies to investigate time-varying multivariate systems and the dynamic behaviour contained therein, through approaches such as statistical comparisons of networks and time series [73, 74, 75], investigation of network features [76, 77, 78, 79, 80], modelling of systems and processes [81, 40, 82, 83, 84, 85], dimensionality reduction [86], visibility graphs [10, 87, 88], and application of further or specialised networks [89, 90, 91].

This literature contains numerous problem-specific and focussed approaches and breadth of more interdisciplinary solutions. An important example of this is the literature focused on instability within these systems, such as the landmark paper by May [19] that employs spectral theory to examine how a systems stability

reacts to its complexity in size and connectivity, finding that systems can quickly move from stable to unstable at a certain complexity. More recently Bardoscia et al. [18] investigated how the interaction topology of the system influences its susceptibility to instability, finding that highly connected components can act as propagators increasing the instability in the system. These two papers represent significant contributions to the study of stability in complex systems and networks. May's work focuses on the general stability of complex systems, while Bardoscia et al.'s research specifically examines the pathways towards instability in financial networks. Both papers highlight the importance of considering the interconnections, dynamics, and feedback effects within complex systems when assessing their stability. They contribute to the broader literature on the spectral theory of networks and systems' stability by providing insights and frameworks for understanding the stability and potential vulnerabilities of complex systems in different domains.

Three conceptual approaches that represent the state of the literature for investigating time varying multivariate systems, that do not impose an investigation of specific behaviours, build upon the view that the relationships between time series may exhibit short term variations that, taken together, can describe dynamic features of the system. These three approaches are introduced below.

The first conceptual approach comes from considering what constitutes the "short" term? This leads into a discussion of scale and what behaviour the system exhibits at different scales. This work constitutes one of the conceptual approaches to time-varying multivariate system analysis, where a system is investigated at multiple scales. Work of this type is often combined with some form of frequency analysis and is focussed on the comparison of behaviour and properties exhibited by the complex networks at different scales [28, 22, 92, 93, 94, 95] (this is mentioned for completeness and is not within the scope of this thesis).

The second conceptual approach comes from considering what short-term part of the original system should be investigated? This question naturally leads to the concept of changes in the short term behaviour at different time intervals or segments of the original system. Therefore analysis can be achieved by compar-

ing complex networks constructed from different segments of the original system [29, 87, 96]. More targeted forms of this analysis select known segments of a system, e.g. those corresponding to specific world events, or known regions of turbulence or calm behaviour, allowing for analysis of the reaction of the systems interaction dynamics to qualitative events [67, 97]. This comparison of the system at different time intervals can easily be conceptually expanded to investigate the dynamic change of the system over time. A temporal list of complex networks can be found by splitting the original system into consecutive segments and then constructing a complex network for each of these segments. The tracking of the dynamic change of features of these complex networks through time can give a method for investigating the dynamics of a time-varying multivariate system [98, 99, 100, 101, 102].

The third conceptual approach to the investigation of time-varying multivariate systems discussed here utilises the series of complex networks produced in the above method and, inspired by approaches for univariate time series, a complex network of this series of complex networks [3, 27, 56, 26, 103, 23, 104, 105, 1]. These methodologies for complex network construction define a transformation to nodes and edges from a time-varying multivariate system; this is covered in general by three primary methodologies discussed further in Section 2.3. Therefore these approaches can be seen as creating a complex network encoding the "high level" dynamic behaviour of the system (where the "low level" behaviour is related to a complex network of the interactions between the variables of the system, the "high level" behaviour is hence related to the total movement between these structures). This type of complex network has significant power due to the representation of individual system states (interaction structures of the system) as separate nodes, making it particularly amenable to analysis of recurring complex non-linear dynamic behaviour, with methods such as clustering being popular. This type of approach abstracts away from the actual structure of the low-level behaviour when constructing a complex network; therefore of note are recent works that, to some level, reincorporate this low-level behaviour into the analysis [2, 37].

This section gave an overview of the current state of the literature, specifically

in terms of the research direction and inspiration of this thesis. From this, the state-of-the-art can be seen as the usage of complex networks to capture the dynamic time-varying behaviour of a system (discussed further in Section 2.3), with initial work utilising purely high-level information and current work starting to incorporate low level and high-level information for analysis of the dynamic behaviour.

2.2 Interaction measures

The research discussed above depends on making some form of measure related to the interaction/connection of two or more different time series. Before discussing the state-of-the-art, mentioned above, in more detail, first a discussion is given of how interactions might be measured. The field of research devoted to the discovery of interaction dynamics between systems has been active for a significant period, and hence a vast collection of literature exists on this subject [106, 107, 108, 109]. From this literature, two important conceptual approaches can emerge; those based on perturbation [110] and those based on statistical measures [106, 107]. Perturbation measures generally look at systems where one component can be altered to see if it affects a change in another. In contrast, statistical measures are based on the concept that the system can not be interacted with and hence a relationship must be determined based purely on historical data. For many systems, a perturbation is not possible, either because the system can not be interacted with in that manner or because it would be immoral or illegal to do so. This is the case for many popular research areas in the field of time-varying dynamical systems, such as the financial markets or neuroscience. Therefore the predominant interest falls on statistical measures of relationships between time series data.

This section gives a general overview of three often-employed statistical relationship measures within the field of time-vary dynamical systems. Many versions and formulations of these measures exist, often designed for specific problems (e.g. for non-linear or multivariate systems). However for the purpose of contextualising this work and comparing the conceptual approach of these measures the overview given here will focus on basic forms of these measures as applied to a general lin-

ear bivariate system. These measures are also discussed only in terms of their classic non-temporal forms, used for finding static interactions between variables. This discussion in justified as the temporal aspect of these interactions is handled through the complex network and a rolling window, allowing for multiple consecutive static measures to be taken, giving an overall dynamic behaviour that can be investigated. This is a popular, and as discussed above state of the art, approach for investigating temporal correlation and causality often referred to as dynamic correlation networks [111]. There also exists a number of other methods for estimating temporal correlations causality such as: detrended cross-correlation analysis, which looks at segements of a time series [112]; Dynamic Conditional Correlation, which uses generalized autoregressive conditional heteroskedasticity (GARCH) models for the variables and estimates the dynamic correlation from the residuals [20]; State Space Models, often focussed on dynamic volatility, which aims to incorporate latent state space variables to detect the time varying aspects of these systems [113]; and Bayesian Approaches, which allow for the incorporation of prior beliefs on the correlation and can use approaches such as Markov chain Monte Carlo (MCMC) to estimate time varying correlation [114]. However, these methods do not afford the same amenability to investigating the general dynamic behaviours of a multivariate system, and hence this thesis chooses to focus on dynamic networks.

Correlation

Correlation can in general be summarised as the degree to which two variables are linearly related. This measure gives a continuous result between -1 for strongly anti-correlated and 1 for strongly correlated, being formally defined for the Pearson correlation coefficient as [115]:

$$\rho_{X,Y} = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

With $\rho_{X,Y}$ being the correlation between variables X and Y, $cov()$ being the covariance, σ being the standard deviation, and μ being the expected value.

This method is helpful in its continuous measure of relationships; however, it

lacks any direction to this (with $corr(X,Y) = corr(Y,X)$) and of course cannot be readily seen as a measure of causality within a system.

Transfer entropy

Transfer entropy is generally summarised as the amount of directed information transfer between two variables. In other words, transfer entropy measures the amount of uncertainty reduction in the prediction of values from variable Y given the past values of variable X, and in both cases using the past values of variable Y [116]. More formally, this can be expressed as:

$$T_{X\to Y} = H(Y_t|Y_{t-K:t-L}) - H(Y_t|Y_{t-K:t-L}, X_{t-1:t-L})$$

With $T_{X\to Y}$ being the transfer of entropy of variable X to variable Y, $H()$ being Shannon entropy [117], and K and L being lags.

This method contains direction and can infer causality, at least in a statistical sense, to the passage of information.

Granger Causality

The third statistical relationship measure discussed is Granger causality, which can be described as a test to determine if one variable causes another, i.e. does the information contained in one variable improve the prediction of another over the information already contained in that other variable [11]. In detail the Granger causality test can be described as follows: for two variables X and Y, where X causes Y (written as $X\to Y$) an unrestricted regression model is created:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j x_{t-j} + \varepsilon_t, \tag{2.1}$$

where ε_t denotes the residual error for time t, $i_{1..p}$ and $j_{1..q}$ denote the lag intervals, with α and β being free variables that are chosen via least squares regression [118]. This model is then compared via a hypothesis test to a restricted model:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t, \tag{2.2}$$

which takes the position of the null hypothesis in the test. Therefore for a Granger causal link to be detected the null hypothesis must be rejected.

A popular choice in Granger analysis is the F test, defined as follows [3, 119, 120]:

$$F = \frac{(RSS_r - RSS_u)/q}{RSS_u/(T-p-q-1)} \tag{2.3}$$

Where RSS_r and RSS_u are the residual sum of squares for the restricted and unrestricted model respectively, and T denotes the sample size [3]. From the F-statistic, the p-value can be calculated, and it is then considered that $X \rightarrow Y$ if the p-value is less than or equal to a predefined alpha level (often set as 0.05). A causal link is then denoted as 1 and a non-causal link as 0.

This method determines if one variable causes another in a statistical sense and has seen great popularity in many fields such as economics and neuroscience.

Comparison of measures

The presented methods represent only a small selection of those present in research but illustrate the general approaches to determining interaction dynamics for simple cases. The work in this thesis is interested in capturing the specific dynamics of the interactions; hence a method is needed that determines directionality. When comparing transfer of entropy and Granger causality, it has been shown that in specific linear systems, they are equivalent [121]. One downside of transfer entropy against Granger causality is that it has been shown to require more data [122]; in general cases, data size does not often present a problem, however, in the analysis of interest, windowing is used that artificially reduces data size (discussed below). Therefore for this thesis, Granger causality is selected as the measure of interaction dynamics to be used going forward. However, it should be noted that much of the work is within reason agnostic to the underlying method.

2.2.1 Granger causality in detail

Given the selection of Granger causality for this work a more detailed discussion of it is given here, as well as how this will be applied going forward.

Linearity

Granger causality was initially introduced for linear interactions [11] and though there have since been expansions to include non-linear interactions [123] much usage still employs the original linear approach. For this thesis the classic linear Granger causality analysis is used with the following justifications. The novel work presented in this thesis utilises the results of the Granger analysis, but is invariant to the manner in which these results are achieved (e.g. through a linear or non-linear analysis) and hence either is appropriate for usage. The system that thesis investigates to determine a novel understanding of and to demonstrate the presented work is a sample of the international oil market (discussed further in Section 2.4), this data set has been previously investigated using linear approaches and has had time-varying interactions discovered (prompting its usage in this thesis) [3]. Therefore, it is assumed here that a linear version of Granger causality is appropriate for usage on this system, though it is acknowledge that systems exhibiting linear behaviour may also contain additional non-linear behaviour. The use of linear Granger causality, allows for the complexity of this thesis to be more precisely focussed on the introduced novel work, instead of the underlying causality measure (which is well known and used across a many of fields), which is being employed but not expanded upon. Additionally the novel work presented in this thesis is designed to capture some level of non-linearity in the form of temporal dynamics of the causality, so even with the use of classic Granger causality the non-linear behaviour of interest can still be captured.

Statistical validation

Granger causality, similar to many measure of this type, contains noise in its analysis and hence statistical validation is needed when drawing any meaningful findings from its results. In this thesis a common approach for the statistical validation of these results is employed, as discussed in detail in section 2.2. The F-statistic is used as an estimator, due to its suitability to validation of nested models. For the choice of significance level (alpha level) at which point to considered the p-value produced by the F-statistics as significant, the common practice is again taken at

5%.

Multivariate Granger causality

The classical approach to Granger causality is restricted to a bivariate system where either A causes B, B causes A or both. However, more than two variables will be present within the systems universe for many real systems of interest. A naive approach to this problem is to apply Granger causality in a bivariate approach to each possible pair of variables within the system; however, this approach can lead to cases of spurious causality. This is due to the other variables in the system acting as confounding variables. For a trivariate system, this can be demonstrated in two models: indirect interaction and driver interaction. For an indirect interaction, the causal system may be $C \rightarrow B \rightarrow A$; however, when analysis via a bivariate Granger approach, one may also find the relation $C \not\rightarrow A$ (where $\not\rightarrow$ represents a spurious link). For a driver case, the causality may be $C \rightarrow B$, $C \rightarrow A$, however during a bivariate analysis; one may also find $B \not\rightarrow A$. To account for this conditional Granger causality can be employed [124]. This expansion of Granger causality can be defined as follows (building upon the bivariate description given above):

First, a bivariate Granger causality analysis is conducted; for this discussion, assume that this analysis produces the results $C \rightarrow B$, $C \rightarrow A$, and $B \rightarrow A$. In this case, potential spurious causal links could be $C \rightarrow A$ (in an indirect case) or $B \rightarrow A$ (in a driver case). Therefore these two links will be tested using a conditional Granger causality method. This approach follows a similar concept to the bivariate case; however, the restricted and unrestricted models change here. Taking the link $C \rightarrow A$ as an example, construct an unrestricted model as follows (For details on the model, see above):

$$A_t = \sum_{i=1}^{p} \alpha_i A_{t-i} + \sum_{j=1}^{q} \beta_j B_{t-j} + \sum_{k=1}^{r} \gamma_j C_{t-j} + \varepsilon_t$$

A restricted model can now also be constructed as follows:

$$A_t = \sum_{i=1}^{p} \alpha_i A_{t-i} + \sum_{j=1}^{q} \beta_j B_{t-j} + \varepsilon_t$$

These two models can now be compared following the standard hypothesis approach; if the restricted model is rejected, the link $C \rightarrow A$ can be considered non-spurious. For the discussed system, this process must also be taken for the link $B \rightarrow A$ with an unrestricted model:

$$A_t = \sum_{i=1}^{p} \alpha_i A_{t-i} + \sum_{k=1}^{r} \gamma_j C_{t-j} + \varepsilon_t$$

If this model is rejected over the unrestricted model, then the link $B \rightarrow A$ is non-spurious.

This conditional Granger causality can be seen as comparing the additional predictive power of a variable whose link is in question to the ensemble of all other predictive variables. Therefore to expand this approach to a multivariate case where more than one alternative predictor exists, one needs to test against the ensemble of other predictors. Therefore the above case can be generalised such that the unrestricted model contains the complete set of variables and the restricted contains the complete set minus the one in question [15].

Windowing

When applying the Granger causality method (as well as other causality methods) to time series that are dynamic and whose interaction behaviour change over time, one cannot simply take the entire length of the time series and apply the method to it, as this would obscure the dynamics of the interactions. To maintain the dynamic nature of the interactions often a rolling window is employed, this breaks the time series into consecutive subset where the causality method can be applied to each subset.

When implementing a rolling window two parameters need be set, the length of the window and the number of time steps between the start of consecutive windows (with different combinations leading either overlapping, or non-overlapping windows). When considering the choice of these parameters the goal is to select a time series long enough to allow for robust statistical estimation, and short enough to capture the time-varying dynamics of the interactions [125]. The the parameters

chosen for use in this thesis are given and justified in Section 2.4.

For clarity, an overview of this approach is as follows: taking a system of N time series, each of length T, and selecting a window size of w and a distance between the start of the windows as l, a series of causality patterns can be constructed. First, the N time series are each split into segments of length w, where each segment starts l time steps after the first, i.e. the first segment is in the range $[0, w]$, the next $[l, (l+w)]$, then $[2l, (2l+w)]$, and so on till the step $T-1$ is reached. For each window, a causality pattern can be constructed from the segment of the N time series. This process can be repeated for each segment to generate a series of these patterns.

Limitations and assumptions

Granger causality is a statistical method and hence has a number of assumptions and limitations associated with it that should be considered when using its results, these are [11]:

- Granger causality assumes a linear relationship between two components, and may not detect or incorrectly detect non-linear relationships.
- Granger causality requires a lag selection for the number of time steps in the past to consider causality to come from, hence the measure of causality is sensitive to the lag length selected.
- Granger causality assumes that the data it is applied to is stationary.
- Granger causality assumes that there are no confounding variables, and that all relevant variables are included in the analysis.
- The causality detected by Granger causality is a correlation of data, and not actual true causal connect.

2.3 Network representations for multivariate time series analysis

In this section a more detailed discussion of the state-of-the-art is given along with a brief discussion of complex networks which this work is based on.

2.3.1 Complex Networks

As previously discussed in this chapter the state-of-the-art in dynamic interactions in multivariate systems utilise complex networks. Like classic networks they are composed of nodes and edges (with these edges either being directed or non-directed), however they expand upon classical networks (such as lattices and random graphs) through their non-trivial topological features [126, 127] (such as non-regular nor random patterns). Two often discussed archetypes of complex networks are scale-free networks [128] and small-world networks [129].

2.3.2 Archetypical Approaches

This section discusses in detail the three methodologies (mentioned in Section 2.1) that represent the majority of approaches to the construction of a complex network from a series of complex networks (these methodologies are not presented as a complete list, more as a selection of prominent approaches of particular relevance to this thesis). For each of these methodologies, a research paper is selected as a case study. The methods presented in these papers are discussed in a generalised sense with reference to the papers where appropriate.

These methods all employ a windowing approach to construct the series of complex networks from which they build, this is done similarly or the same as the method described in Section 2.2.1.

Symbol based network approach

The first methodology discussed is an approach very commonly employed for systems where correlation is taken to measure the interaction dynamics (resulting in correlation patterns). The discussion of this approach follows the work of Qi *et al.* and others [1, 105, 104], with a graphical summary shown in Fig. 2.1, this approach can be described as follows:

1. Select a fixed window size, then apply a sliding window approach to the time series data to separate the series into sub-periods.

2. Analyse the interactions between the sub-periods for each time series that correspond to the same time frame (e.g. via a correlation analysis).

3. Assign a symbolic representation to categorise the analysis results (e.g. if the correlation is ≥ 0.8, assign the symbol P). This will produce a single time series of symbols representing the interactions throughout the initial time series.

4. Select a fixed window size, then apply a sliding window approach to the time series of symbols. Each of these windows represents a "mode" and is composed of an ordered series of symbols. If more than one window contains the same ordered series of symbols, their assigned mode is the same.

5. Construct a complex network, where the nodes are the found modes, and the edges exist for consecutive modes. With the weighting being the number of times these modes appear consecutively (the frequency) (i.e. if mode$_i$ transitions to mode$_j$ n times in the series, then a directed edge would exist from mode$_i$ to mode$_j$ weighted by n).

This approach has great potential for determining group transition behaviours, i.e. behaviours concerning the modes, seeing usage in areas such as the five day trading week and allowing for behaviours relating to the mode of each week to be uncovered [1]. However, this method does face drawbacks where behaviours over shorter scales may be obscured, e.g. where the change between individual trading days becomes important.

Visibility graph based network approach

The second method discussed is an expansion of an approach commonly used for univariate time series, the Limited Penetrable Visibility Graph (LPVG) [51]. Here the work of Yu *et al.* [2] is followed, who uses a directed multivariate version of LPVG called Directed Limited Penetrable Visibility Graph (DLPVG): this is employed via a network approach they title Multivariate Time Series-Dynamic Association Network (MTS-DAN). A graphical summary of this approach is shown in Fig. 2.2, and can be described as follows:

1. Select a fixed window size, then apply a sliding window approach to the time series data to separate the series into sub-periods.

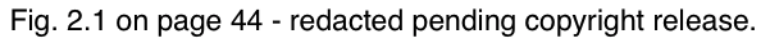

Figure 2.1: *Figure from the work of Qi et al [1], detailing their method of taking a bivariate system through several stages to produce a complex network. These stages are: 1) applying a Pearson correlation analysis to a series of windows of the original time series, 2) turning this series of correlation results into a symbolic series (they call this "symbolization"), 3) turning this series of symbols into a series of modes, using windowing, 4) turning the series of modes into a complex network.*

2. Analyse the interactions between the sub-periods for each time series corresponding to the same time frame (e.g. via transfer entropy) to produce an association pattern for each sub-period.

3. Apply the Principle Component Analysis (PCA) algorithm to each pattern to reduce it to a one-dimensional vector, resulting in a time series of these one-dimensional vectors.

4. Construct a complex network by applying the Directed Limited Penetrable Visibility Graph (DLPVG) approach to this time series of one-dimensional vectors, DLPVG can be summarised as follows (and can be seen in Fig. 2.3):

   (a) Transfer the time series to a graphical representation, where data is plotted in order as columns representing the value of the individual data at each time step.

   (b) Select a limited penetrable distance D, this defines the number of columns a time step can see through (explained in the following).

   (c) For each column connect it to any column it can see from its peak that is not blocked by another column (it may be able to ignore D other columns to do this). This can be more formally defined as, two data points ((t_i, y_i) and (t_j, y_j)) will have visibility and be connected to each other as long as $\leq D$ data points between them (t_c, y_c) satisfy the following: $y_c \geq y_b + (y_a - y_b)\frac{t_b - t_c}{t_b - t_a}$ [51].

   (d) The directed nature of this approach means that all connections will be directed forward in time; also, no weights are associated with the connections.

   (e) A complex network can be constructed, with nodes as the data points and edges as the links discovered above.

This method incorporates some knowledge of the underlying behaviour of each pattern into its complex network through its use of PCA. However, this approach has a particular linkage method, which may not be appropriate depending on the system, e.g. if one does not desire connections between non-consecutive patterns.

Figure 2.2: *Figure from the work of Yu et al [2], detailing their method of taking a multivariate system through a number of stages to produce a complex network. These stages are: 1) apply a windowing approach to the data, 2) find the association patterns for each window via transfer entropy, 3) reduce the association patterns to one-dimension via PCA, 4) apply MTS-DAN method to this series to construct a complex network.*

Transition order based network

The third method captures the historic probablity of transition between system states. To demonstrate this approach the work of Jiang *et al.* [3] is followed, with a graphical summary shown in Fig. 2.4 the method can be seen as follows:

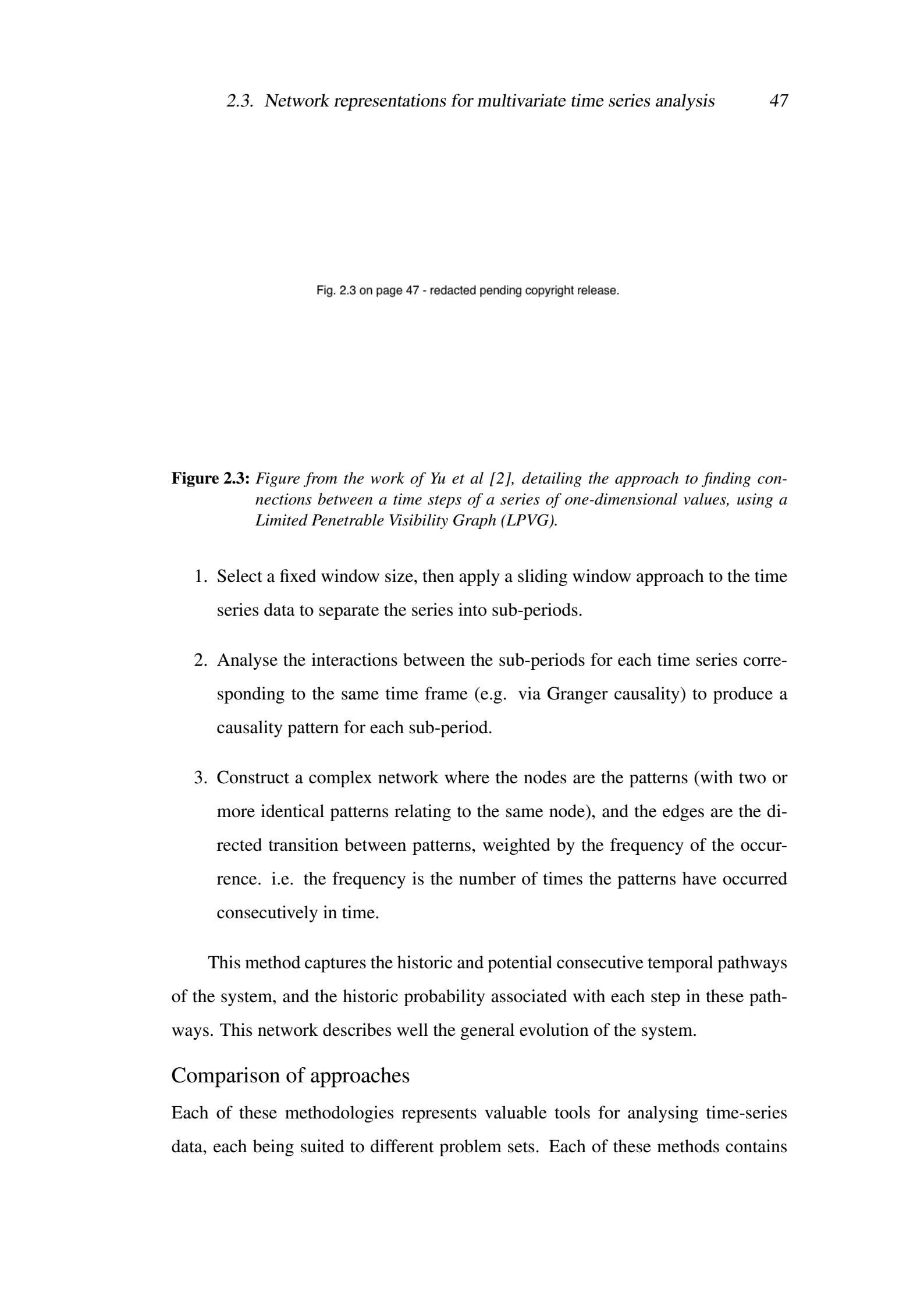

Figure 2.3: *Figure from the work of Yu et al [2], detailing the approach to finding connections between a time steps of a series of one-dimensional values, using a Limited Penetrable Visibility Graph (LPVG).*

1. Select a fixed window size, then apply a sliding window approach to the time series data to separate the series into sub-periods.

2. Analyse the interactions between the sub-periods for each time series corresponding to the same time frame (e.g. via Granger causality) to produce a causality pattern for each sub-period.

3. Construct a complex network where the nodes are the patterns (with two or more identical patterns relating to the same node), and the edges are the directed transition between patterns, weighted by the frequency of the occurrence. i.e. the frequency is the number of times the patterns have occurred consecutively in time.

This method captures the historic and potential consecutive temporal pathways of the system, and the historic probability associated with each step in these pathways. This network describes well the general evolution of the system.

Comparison of approaches

Each of these methodologies represents valuable tools for analysing time-series data, each being suited to different problem sets. Each of these methods contains

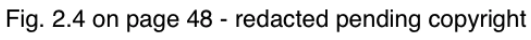

Figure 2.4: *Figure from the work of Jiang et al. [3], detailing their method of taking a multivariate system through several stages to produce a complex network. These stages are: 1) apply a windowing approach to the data, 2) determining the causality pattern for each window, 3) turning the series of these patterns into a complex network by taking the patterns as nodes and the transitions between them as edges.*

similarities and differences. In particular, the edge selection is similar between the work of Jiang *et al.* [3] and Qi *et al.* [1], whereas the node selection is similar between Jiang *et al.* [3] and Yu *et al.* [2]. For application to general systems composed of time-varying interacting components, the work of Qi *et al.* and Yu *et al.* impose assumptions that can not always be assumed. For Qi *et al.* the assumption that mode like structure exist in the high-level time-varying behaviour of the system (as well as that needed frequency of these modes is accurately determined), and that even if it does, that this is the behaviour of merit, makes it more sorted

for investigating this specific behaviour than to apply to general systems. For Yu *et al.* the assumption that the principle component of the interaction pattern is a subtle linkage factor, assumes that stronger interacting patterns are of more importance, and have merit in being linked, than weaker interacting patterns. Though this may be the case in some system, and for analysis, this is certainty not the case in all system, where weakly interacting states may be very important in the temporal ordering, making this approach better suited to systems with known behaviours for their interacting patterns. The work of Jiang *et al.* does not impose assumptions on the importance of one structure over another, and treats all transitions and structures the same. Due to this, this network is the most general being able to be used without the assumptions associated with the other two approaches.

Inclusive of limitations these approaches allow for a breadth of analysis to be applied to systems of time-varying interacting components, however an important source of information on these systems dynamics that is excluded from the above is the the nature of change to the system state when a transition occurs, captured primarily only in direction and frequency. This gap is tackled in Chapter 3 of this thesis, owing to its flexibility as an approach, the network representation presented by Jiang *et al.* is chosen as a starting point for filling this gap.

2.4 Causality in the international oil market

When adding new research and results to those presented in this chapter it is desirable to have a real-world data set on which to investigate.

There are a few key requirements to consider when selecting this data; the data should be stationary for the Granger causality test to be applicable, the data should be known to exhibit time-varying behaviour, and it would be desirable for the data to be from a topic of genuine interest within the field.

There are many types of data known to satisfy the first requirement of stationarity. In particular, for many financial variables, though their raw time series are non-stationary, they often exhibit stationary in the mean for their returns [130]. Therefore the financial markets are selected as the area from which the case study

shall be drawn.

A region of the financial markets that has received much attention for its time-varying behaviour, especially within the last decade, is the international oil market [22, 23, 24, 25, 26, 3, 4, 27, 28]. In particular, this market has been shown to containing time-varying Granger causality relationships [3] and has been noted as an example system for time-varying behaviour [3]. Therefore the international oil market can be seen to fulfil the second requirement.

The research in [3] was conducted due to interest in this market by investors, regulators, policymakers, and economists. Driving this interest and establishing the international oil market as an important area of research is the evidence that fluctuations in this market are linked to stock returns and factors of market stability [22, 131, 132, 133, 134, 135, 136, 137]. This connection demonstrates that the international oil market satisfies the third requirement, with time-varying behaviour in this market being of interest to many practitioners.

Hence this market is selected as the area from which the data used within this thesis is drawn.

Due to the international oil market being the focus of much work, many different data streams have been utilised. This research can be roughly categorised into two groups: those that look at the relationships between the international oil market and other markets or factors [22, 23], and those that look at relationships within the international oil market [24, 25, 26, 3, 4, 27, 28]. For the first of these, the international oil market is predominately represented by a single data source, e.g. the spot price of one crude oil source, that acts as a proxy for the market as a whole. For the second group, where the international oil market is investigated internally, more than one oil price source is required, and hence this research often selects several sources that are seen as benchmarks to act as the proxies for regions of the oil market. A data set where the sources are justifiably linked is presented by Jia *et al.* [3], who select four spot price time-series: China-Daqing and its three reference benchmark oil prices Brent, Dubai, and Minas. This selection can be seen as authoritative for the international oil market, with Daqing representing the crude oil

price in China and the reference benchmarks being selected to represent the three important regions of the North Sea, the Middle East, and the Asia Pacific. Therefore in this thesis, these four sources were selected for use, with the specific time series being the same as those used by Jia *et al.* [4]. These time series are for the period of December 27^{th}, 2001 - October 31^{31st}, 2011, and are shown in Fig.2.5.

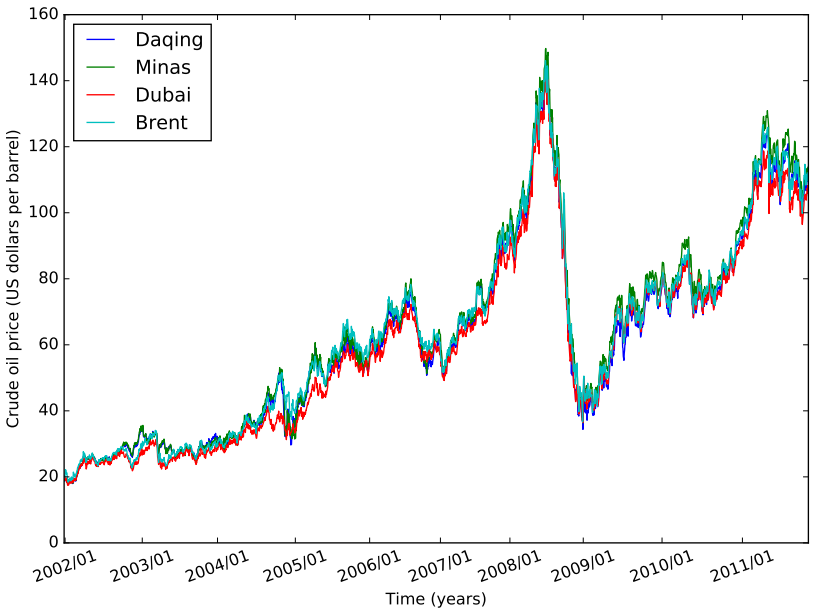

Figure 2.5: *The daily price in US dollars per barrel of Daqing, Minas, Dubai, and Brent, over the period December 27^{th}, 2001 - October 31^{31st}, 2011, data sourced from Jia et al [4].*

Data transformation and analysis

For usage this data is transformed to its returns, to fulfil the stationarity condition. The returns are defined as the difference of natural logarithms of price, formally as $r_t = ln(P_t) - ln(P_{t-1})$ [130]. The returns of this data can be seen in Fig.2.6. These returns are then tested for stationarity using a Augmented Dickey-Fuller (ADF) test, results shown in Table.2.1, indicating that all return series are stationary.

These returns can now be analysed using the previously described multivariate Granger causality tests. It should be noted that self causality is not investigated (as it is not relevant when considering the interaction between returns), and hence all self causality links will be automatically assumed not to exist (therefore set at 0).

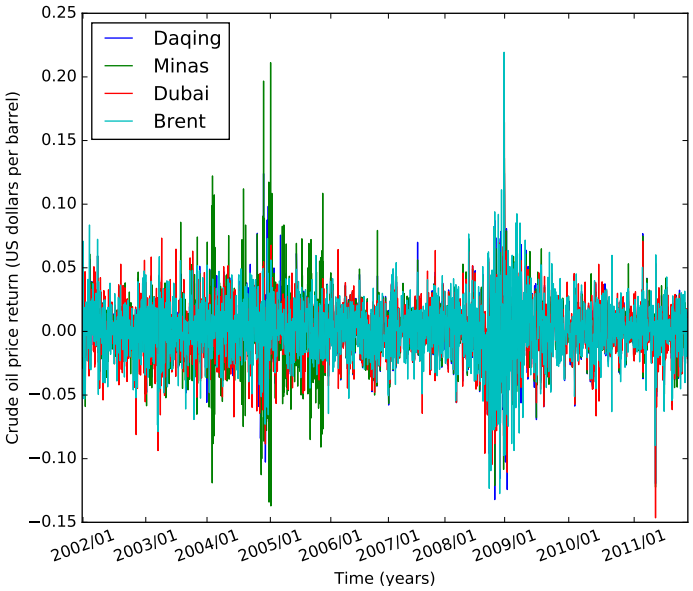

Figure 2.6: *The daily price returns in US dollars per barrel of Daqing, Minas, Dubai, and Brent, from the daily price shown in Fig.2.5.*

Table 2.1: Results of stationary tests using a Augmented Dickey-Fuller (ADF) test. A p-value$<$0.01 indicates the rejection of the null hypothesis for the test at a 1% level.

	ADF-Statistic (3 sf)	p-value
Daqing	-11.7	0.001
Minas	-20.3	0.001
Dubai	-51.8	0.001
Brent	-9.79	0.001

Here a static analysis of the whole series is conducted, the collection of causality links discovered will be referred to as a causality pattern (following the convention of [3]). The static causality pattern, representing the causality calculated from the full data length, can be seen in Fig.2.7

To create a complex network representation of this data, the method used by Jiang *et al.* (discussed in Section 2.3.2) is employed. This method requires windowing to separate the time-series data into sub-periods. When selecting this window length, two important considerations are; (i) the length of the window should fit the goal of the analysis, i.e. for analysis of short term behaviours, smaller sizes are appropriate, and for longer-term behaviours, larger sizes are, (ii) the diversity of the

	Daqing	Minas	Dubai	Brent
Daqing	0	1	1	0
Minas	0	0	0	0
Dubai	1	1	0	0
Brent	1	1	1	0

Figure 2.7: *The static causality pattern of the sample data returns, 1 represents a Granger casual link from left to top and 0 represents no link. The Granger tests were done at a significance level of* 0.05.

causality patterns and transitions among them should fit the goal of the analysis, with an increase in window size the results of separate windows will become more similar to the original time series. More generally these considerations can be described from the perspective of statistical estimation. The window size determines which data will be used in the estimations, for dynamic interactions longer window sizes may capture this behaviour and then further behaviour, leading to the estimation potentially obscuring short lived interactions. Therefore, the window size can be seen as determining the causality pattern. The gap between these window starts determines how much similar data is being used to define consecutive causality patterns, for non-overlapping windows completely new data is being analysed, whereas for predominately overlapping windows most data is identical. Therefore, for the former it is possible to have very different consecutive causality patterns discovered, whereas for the latter it is probable that consecutive causality patterns will be very similar. Therefore the window spacing can be seen to determine the transitions between causality patterns.

The analysis in this thesis is focussed on short term variation within the system, looking for distinct short term behaviours within the system; therefore, a window size that maintains a large diversity of causality patterns is most appropriate. For the causality patterns produced further analysis is intended to be employed, hence a time-series of causality patterns of significantly length is desirable. Furthermore, the transition throughout time of the causality patterns of the system are of interest,

hence a small step size between windows that allows for a gradual change in the causality pattern will be of benefit. Due to the these reasons a rolling window increasing by on time step at a time is chosen. To determine the window length the number of nodes and edges in networks produced from different window sizes is investigated and shown in Fig.2.8. Following the above argument and the results of Fig.2.8 a window size of thirty time steps is selected (similar arguments and conclusions have been drawn within the literature [3])

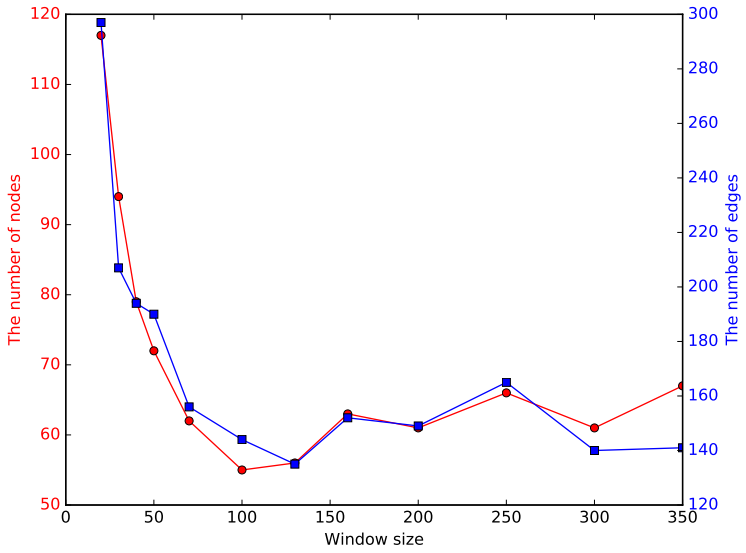

Figure 2.8: *The number of nodes and the number of edges for complex networks generated from window sizes 20, 30, 40, 50, 70, 100, 130, 160, 200, 250, 300, and 350, for the data shown in Fig. 2.6.*

Chapter 3

Properties of Dynamic Causality Patterns in a Complex Network

Chapter 2 introduced three of the main approaches currently employed within the field to transform a time-varying multivariate system of time series to a complex network representation. A discussion was also presented of how in general these methods can be seen as decreasing the level of complexity of the problem by reducing the knowledge of the causality pattern structure (the underlying structure; the system interactions during a time step). The literature review also presented a discussion of the use of this knowledge, with many systems being characterized by their causality pattern structure in other forms of analysis such as static systems [72, 23, 67, 69]. Therefore, it is apparent that within literature a systems causality pattern is considered to hold valuable descriptive information, and that it is also the case that this information is limited in its incorporation into analysis of dynamic causality systems.

This chapter seeks to fill this gap by presenting a novel approach to the incorporation and investigation of time-varying multivariate causality patterns, hence allowing for new and efficient descriptions of these systems. This is achieved by proposing a method to incorporate the knowledge of the change in causality pattern over a time step into a new complex network representation, further a number of properties and metrics are presented to demonstrate the ease of analysis offered by this approach. These properties and metrics are validated via synthetic data, finally

an analysis of the international oil market is conducted using this method.

3.1 Background and related work

This section discusses two aspects of literature: complex networks construction approaches and properties/metrics applied to complex networks.

As mentioned above, this chapter is interested in the use of knowledge from the underlying causality patterns of the system that has dynamic behaviour. Therefore it is important to understand how this causality information is traditionally incorporated into complex network representations, which are the predominate form of analysis for this type of system. In Chapter 2 a literature review of this topic was presented that discussed three of the main approaches to achieving a complex network representation. To summarise, these approaches on the whole reduce the complexity of the data by limiting the knowledge of the causality patterns and their transitions. However it should be noted that some of the methods incorporate some level of knowledge of the underlying causality patterns and their transitions, this is most notable in the work presented by Yu *et al* [2], Multivariate Time Series-Dynamic Association Network (MTS-DAN) (discussed in detail in Chapter 2.3.2). Yu *et al* take a slightly different approach to edge construction than the alternatives. They construct the network's edges via an edited form of limited penetrable visibility graph (LPVG), which is of note due to its requirement for each node to have a value, representing its causality pattern, associated with it. This is achieved via Principle Component Analysis (PCA) to determine a one-dimensional vector representation of the interaction pattern (causality pattern) representing each node. This one-dimensional vector associated with each node is then used to determine the edges between the nodes in the complex network. This approach does something that the other complex network transformation methods do not, and that is to take some knowledge of the interaction pattern for a subcomponent and incorporate that into the constructed complex network. This approach however has a number of short comings for the analysis that is of concern in this chapter: the constructed network is not exclusively temporally ordered and hence edges between nodes do not

exclusively represent consecutive causality patterns, hence they can not be weighted by the frequency of transitions between patterns, and the causality pattern and its transitions are encoded in a manner that obstructs behaviours associated with them from further analysis (a further discussion is given in Chapter 2.3.2).

Complex networks, such as those used to represent time-varying systems, are often examined through characterisation of local and global properties. The field of complex networks is considered mature, with a significant amount of literature existing on the topic [9]. However since the usage of complex networks for time-varying systems is a more recent development the application of characterisation to these cases has received less investigation. Here an overview of a prominent aspects of this work is presented.

Notable work on this type of complex network is presented by Jiang *et al* [3], Yu *et al* [2], Qi *et al* [1], Dong *et al* [56], and Yu *et al* [37] (and talked about in detail in Chapter 2). Here a briefly introduction and discuss of some of the analysis methods that these authors have applied to their networks.

An important metric in a network is the ”degree” of each node within the network. This is the weighted sum of the edges connected to a certain node, and can be seen as a measure of how important that node is to the network at large. Work has been done using this metric to determine the importance of certain causality patterns to the dynamic behviour of the system, this primarily is done by investigating if certain nodes have significantly larger degree values than others. A common technique is to compare the distribution of degree across the network to a power law or a scale free network, to show that a small subset of the causality patterns exert a large amount of importance over the dynamic behviour. Comparison with the average weighted degree is also sometimes used to determine important nodes [3, 2, 1, 56].

Following the metric focused on the nodes, one can apply similar logic to the edges of the network. The distribution of the edges based on the frequency of that edge can be compared to power laws to determine how the transition between causality patterns is spread throughout the behviour, allowing the isolation

of particularly prominent edges. The edges can also be investigated to look at self-transformations/self-loops (i.e. when a causality pattern does not change over a time step). This is done at a global level to determine the total percentage of self-loops in a network, giving an understanding of how stable that network is, but can also be applied to nodes of interest to determine their individual self-loop weighting [3, 56].

A popular network method that has been employed is the closeness of the network, measuring how well connected the entire network is. This is defined as the number of edges within the network divided by the total potential number of edges that could exist in the network [138]. This represents the tightness of transition between different causality patterns, demonstrating the wilingness of causality patterns to transition to any other available pattern [1].

Another measure that has been employed is the betweenness of a node, this represents the number of times a node appears as part of the shortest path between two others. This metric is important for determining causality patterns that are highly important for multi pattern transitions, allowing for the movement from one causality pattern to another [2, 1, 56, 37].

Determining groups of nodes that are more likely to transition internally than externally can be important for understanding the dynamic behviour of the causality pattern. In particular an algorithm by Blondle *et al* based on maximising modularity has been used to make determination of clusters of nodes that can be considered as distinct groupings within the network [139]. These groupings can be used to categorise broad behaviours of the dynamics of the causality pattern, including further analysis on the individual clusters such as average path length [3].

3.2 USIC: a new complex network

The analysis of non-linear behaviours, and structures, of time varying causality interactions within multivariate systems can be challenging due to the complex nature of this information. An approach to make the data more amenable to analysis is to encode the evolutionary information into a network representation [138], as discussed in Chapter 2.

The transfer of a time-varying multivariate system to a complex network to analyse its evolution has been employed in a number of studies [3, 2, 1, 56, 37]. Although the exact methodology for this transfer varies, the general approach can be seen as similar. This thesis will be following the approach presented by Jiang *et al* [3] which employs Granger causality, a comparison of approaches is discussed in Chapter 2.3.2 and in the following section, but for the purpose of this Chapter the method presented by Jiang *et al* can be seen as a good basis for expanding upon. This method can be described as constructing a complex network where the nodes represent causality patterns and the edges represent the frequency of transition between these patterns and is described in Chapter 2.3.2. Below a formal description of this general approach is presented, to provide an unambiguous foundation for the discussion that follows. To the authors knowledge a formal specification of this nature has not been presented elsewhere.

This specification is given by defining the following:

- A set V, which is the set of all labelled time-series variables v_x where x ranges from 1 to n inclusive. Using the notation $[[a,b]]$ to indicate the set of all integers from a to b inclusive, written as:

$$V = \{v_x\} \; \forall x \in [[1,n]]$$

- An individual Granger causality metric $c_{x,y}$ which gives 1 if time-series variable v_x Granger-causes v_y and 0 otherwise. Using the previously-introduced Granger causation arrow, written as:

$$c_{x,y} = \begin{cases} 1, & \textit{if } v_x \rightarrow v_y \\ 0, & \textit{otherwise} \end{cases}$$

- A *time-labelled* individual Granger causality metric $c_{x,y,t}$ from time-series variable v_x to v_y at time t (where t is the time label of the sliding window, as previously described). Using $\mathbb{N}_0$ to indicate the set of natural numbers in-

cluding 0, define $c_{x,y,t}$ to be the 2-tuple $(c_{x,y}, t)$ such that the predicate $t \in \mathbb{N}_0$ holds. Written as:

$$c_{x,y,t} = (c_{x,y}, t) \mid t \in \mathbb{N}_0$$

- A multivariate causality pattern C, which is the set of all individual causality metrics $c_{x,y}$ where x and y both range from 1 to n inclusive. This is written as:

$$C = \{c_{x,y}\} \; \forall x, y \in [[1, n]]$$

- A *time-labelled* multivariate causality pattern C_t, which is the set of all time-labelled individual causality metrics $c_{x,y,t}$ where x and y both range from 1 to n inclusive and t is the end time of the sliding window, as previously described. Thus:

$$C_t = \{c_{x,y,t}\} \; \forall x, y \in [[1, n]]$$

Each time the sliding window moves forward in time a new causality pattern is observed. If observations start at time 0 and end at time T, define a set O to be the set of all observed time-labelled causality patterns C_t where time t ranges from 0 to T inclusive. Thus:

$$O = \{C_t\} \; \forall t \in [[0, T]]$$

Next define a network $\mathscr{N}$ of nodes and edges by $\mathscr{N} = (N, E)$ where nodes in N are representative causality patterns (defined below) and edges in E are weighted directed connections between representative causality patterns (nodes). If two or more observed causality patterns $C_{t_1}, C_{t_2}, \ldots C_{t_k} \in O$ have the same pattern of causality (albeit measured at different times) then they are represented by a single node in $\mathscr{N}$. Thus, first define the equivalence sets of the observed time-labelled causality patterns (each equivalence set contains all observed C_k that share the same underly-

ing pattern of causality) and then define the associated labelled representative node R_l to be the associated underlying causality pattern. This also requires a function to extract the underlying pattern of causality from any such C_k. Therefore define:

- A function $patt()$, which gives the set of all time-independent individual Granger causality metrics that correspond to all of the time-labelled individual Granger causality metrics in a causality pattern C_t. Thus:

$$patt(C_t) = \{c_{x,y}\}\ \forall (c_{x,y,t}) \in C_t$$

- A labelled equivalence set R_l^{set} with label l (a natural number) is a set of all 2-tuples comprising a causality pattern C_k and label l, for all C_k in the set O of observed causality patterns, such that all members of a given R_l^{set} will have the same pattern of causality returned by the function *patt* (i.e a R_l^{set} will exist for each unique found causality pattern and will contain all instances of that pattern observed in O). Thus (using "*and*" to connect multiple quantifiers, and "$\wedge$" for the logical conjunction of predicates):

$$R_l^{set} = \{(C_k, l)\}\ \forall C_k \in O\ \ and\ \ \forall (C_p, l), (C_q, l) \in R_l^{set} \mid l \in \mathbb{N} \wedge patt(C_p) = patt(C_q)$$

- A labelled representative node R_l (node in the network) with label l (a natural number) is a 2-tuple comprising (i) the underlying causality pattern of any C_t in the equivalence set with the label l, and (ii) the label l. The label provides a one-to-one mapping between each R_l and its associated R_l^{set} (so knowing R_l implies knowledge of R_l^{set}):

$$R_l = (patt(C_t), l) \mid (C_t, l) \in R_l^{set} \wedge\ l \in \mathbb{N}$$

- The set N is a set of labelled representative nodes such that for each observed causality pattern C_t there exists at least one representative node in N that

contains the underlying pattern of causality of C_t. Thus:

$$N = \{R_l\}\ \forall C_t \in O \textit{ and } \exists R_k \in N \mid R_k = (patt(C_t), k)$$

- The set E is a set of weighted edges, each being a 3-tuple of the start node (R_{l_1}), the end node (R_{l_2}) and the weight. Initially the weight is the number of transitions from R_{l_1} to R_{l_2} taking place over one time step; this is commonly referred to as the frequency "F". Start by defining the function $freq()$ (where $freq() \in \mathbb{N}_0$) to calculate the desired result — it does this by summing all transitions that exist from a causality pattern C_t in $R_{l_1}^{set}$ to a causality pattern at the next time step C_{t+1} in $R_{l_2}^{set}$ (i.e. the number of times one causality pattern transitions into another), Using the notation $|\{\}|$ to give the cardinality of a set, thus:

  $$freq(R_{l_1}, R_{l_2}) = |\{((C_t, l_1) \in R_{l_1}^{set}, (C_{t+1}, l_2) \in R_{l_2}^{set})\}|\ \ \forall t \in [[0, T]] \textit{ and } f \in \mathbb{N}_0$$

  Now define the set E as follows (with the additional constraint that an edge doesn't exist in E if the calculated frequency is zero):

  $$E = \{(R_{l_1}, R_{l_2}, f)\}\ \forall l_1, l_2 \in [[1, |N|]] \mid f = freq(R_{l_1}, R_{l_2})\ \wedge\ F > 0$$

Following this definition the evolution of the causality between time series variables, V, can be expressed as the complex network $\mathscr{N}$.

3.2.1 Limitations of current network approaches

As discussed earlier in this Chapter and in Chapter 2.3.2 the current network approaches for the analysis of time varying multivariate causality patterns reduce the data in such a way as to abstract away from the actual change occurring in the causality pattern between time steps. This change in causality pattern structure (which variables are causality related to each other in each time step) contains information that can inform and contextualise these transitions. With the topol-

ogy of a network being important to its function, and the change of this topology over time being connected to important in the system it represents [140, 141, 142]. Whereas previous approaches focus on which causality patterns follow each other (in time) but not what these temporal orderings mean for the systems evolution [3, 2, 1, 56, 37]. This lack of information on the transitions of causality patterns leaves gap in the literature for a more informative description of the time varying nature of these systems.

3.2.2 A causality pattern based complex network

Here a novel method for encoding the information on the transition of the causality patterns in a complex network for analysis is presented, this is done through the introduction of two metrics to encode the information and a complex network to carry it. This introduce a network approach that allows for more efficient discover of some behaviours and novel discovery of others.

3.2.2.1 Two New Metrics

As discussed in the previous section, the general approach in construction of a complex network to represent the evolution of causality patterns only encodes information relating to frequency and temporal ordering. This approach produces a network that is easily handled by existing network approaches, and therefore generalised network analysis is often employed [138, 31]. This approach to analysis can yield important results about the system and can be seen as a significant tool in the analysis of time varying causality patterns. However, these existing standard network approaches focus on frequency-weighted edges, and this does not allow for further exploration of the changes occurring between the causality patterns in the start and end nodes of each edge. This is a problem because it removes the potential to uncover complex behaviours taking place related to the change in structure of the underlying causality patterns, which may hide important information about the evolution of those causality patterns.

There do exist complex network mapping methods that take some account of the underlying structure of the system's interactions (referred to earlier). Of note

is the work of Yu *et al* who introduce a Multivariate Time Series-Dynamic Association Network (MTS-DAN) using a Directed Limited Penetrable Visibility graph (DLPVG) approach [2]. The incorporation of the underlying pattern structure in this method is achieved by using Principle Component Analysis (PCA) to produce a one-dimensional representation of each causality pattern, and then based on this representation to add new links between nodes. This approach constructs a complex network that is unweighted, directed forward in time (but not necessary temporally sequential), and contains links not associated directly with transitions. This construction therefore contains some implicit information of the causality patterns in its linkage choice. Though this network construction may be of interest in certain areas of analysis, by definition it does not contain much of the information encoded in Jiang *et al's* construction [3], namely a guaranteed temporal ordering (if two causality patterns are sequential they will be linked) and a frequency weighting of sequential transitions (how many times one causality pattern has transitioned to a specified other).

It is desired to counter these issues and present a novel methodology that is amenable to analysis and incorporates information on the change in causality pattern during evolution. The proposed methodology takes a similar initial approach to that of Jiang *et al's* [3] (discussed earlier), but includes a greater amount of information content through the addition of new edge weights that are specifically constructed to encode select information on the transition. This chapter proposes an extension to the complex network model presented by Jiang *et al* [3] (and discussed earlier), in regard to transitions over one time step, by adding new weights to the edges. These weights correspond to information regarding the start and end nodes, to encode information on the evolution of the causality pattern over the corresponding transition, i.e. how the causality pattern changed over one time step. The manipulation and comparison of the full causality patterns in the start and end nodes of an edge can be unwieldy where there are large numbers of times-series variables. In this chapter it is therefore proposed to use two simple values to encode the structural change in causality pattern, as described below (where either or both

aspects may occur over a single time step).

As a preliminary step, this chapter defines a metric that encodes an important informational aspect of a causality pattern. This is the arithmetic sum of all of the individual causality metrics in the causality pattern of a single node. This is labelled as the "total causality" of a node, and it is calculated by the auxiliary function $total()$, also used is the auxiliary function $fst()$, which returns the first item of a 2-tuple:

$$total(R_l) = \sum_{a,b}(c_{a,b} \in fst(R_l))$$

When a causality pattern changes, the overall strength of causation (the "total causality") may or may not change. Therefore, in this chapter two new metrics to encode these two characteristics of a change in causality pattern are proposed:

1. Total Difference in Causality (α): this captures the changes in individual causality metrics, regardless of the "total causality", and define it as the sum of the differences between corresponding individual causality metrics (each difference is squared and rooted, to make it a positive number independent of direction). This is calculated with the function $alpha()$ as follows (where $alpha() \in \mathbb{Z}$):

$$alpha(R_{l_1}, R_{l_2}) = \sum_{a,b} \sqrt{((c_{a,b} \in fst(R_{l_2})) - (c_{a,b} \in fst(R_{l_1})))^2}$$

2. Net Change in Causality (β): this captures the overall change in "total causality", regardless of any change in which causality links do and do not occur and define it as the difference of the "total causality" metrics for the start and end nodes. This is calculated with the function $beta()$ (where $beta() \in \mathbb{Z}$):

$$beta(R_{l_1}, R_{l_2}) = total(R_{l_2}) - total(R_{l_1})$$

3.2.2.2 USIC-Network

Following the introduction of these two metrics this chapter proposes a new network called Underlying Structural Information Consideration Network (USIC-Network) that makes use of these metrics. This network construction takes each unique representative causality pattern as a node and takes edges between them to exist if the two causality patterns appear sequentially in the evolution. Each edge is weighted with three quantities; the frequency of transition (F), the Total Difference in Causality (α), and the Net Change in Causality (β). The layout of this transformation is shown in Fig. 3.1. By expanding on the earlier formal definition of a complex network representation, the USIC-Network can be defined such that the set E of weighted edges is modified to have a weight that is a 3-tuple containing F, α and β (defined above). For convenience define a labelled edge $e_{i,j}$ as follows (with an example of a edge of this form being shown in Fig. 3.2):

$$\begin{aligned} e_{i,j} =& (R_i, R_j, (f, \alpha, \beta)) \mid i, j \in \mathbb{N} \wedge F > 0 \\ & where\ f = freq(R_i, R_j) \\ & and \quad \alpha = alpha(R_i, R_j) \\ & and \quad \beta = beta(R_i, R_j) \end{aligned}$$

The definition of E can now be rewritten as:

$$E = \{e_{l_1,l_2}\} \ \forall l_1, l_2 \in [[1, |N|]]$$

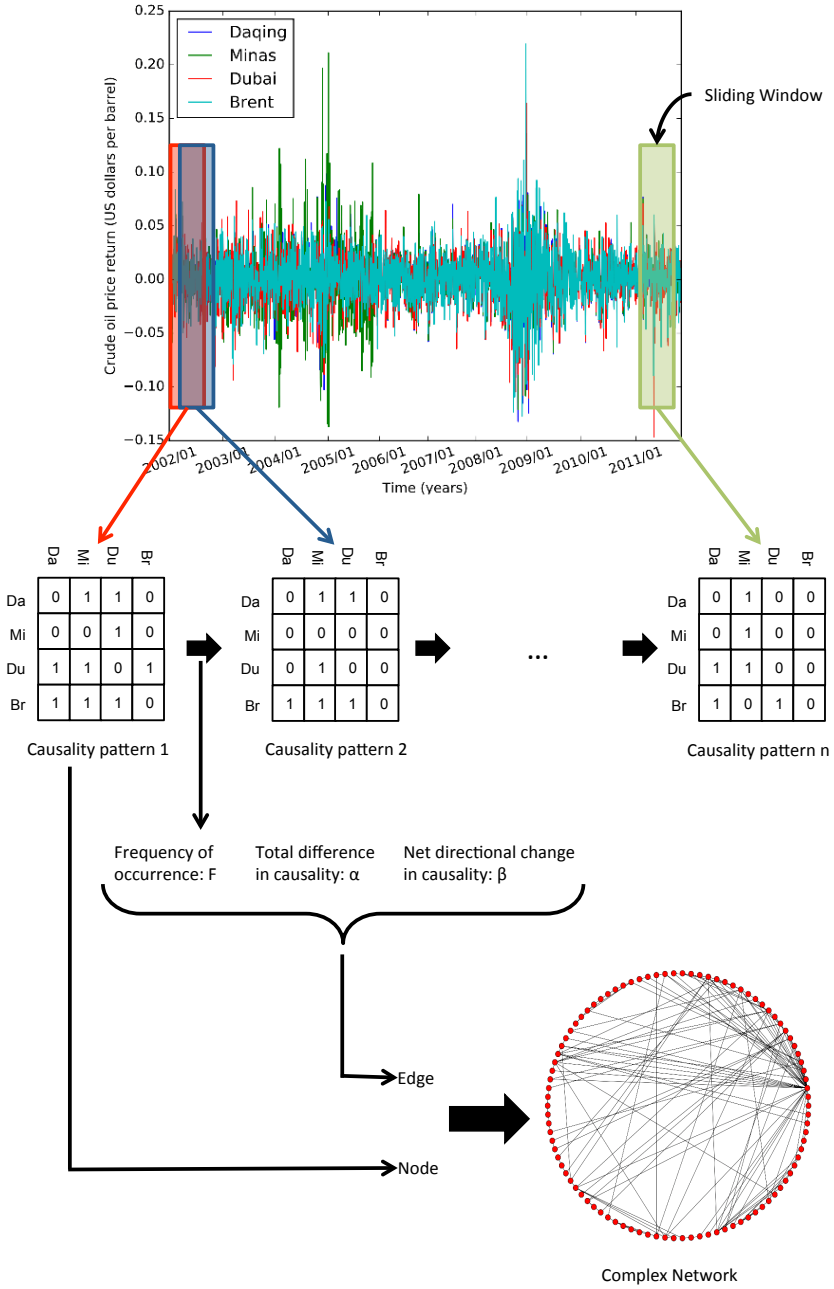

Figure 3.1: *Outline of the methodology of the USIC-Network model for a sub system of the international oil market comprising the returns of the spot price variables Daqng, Minas, Dubai, and Brent. Causality patterns and complex network displayed are for representation purposes only.*

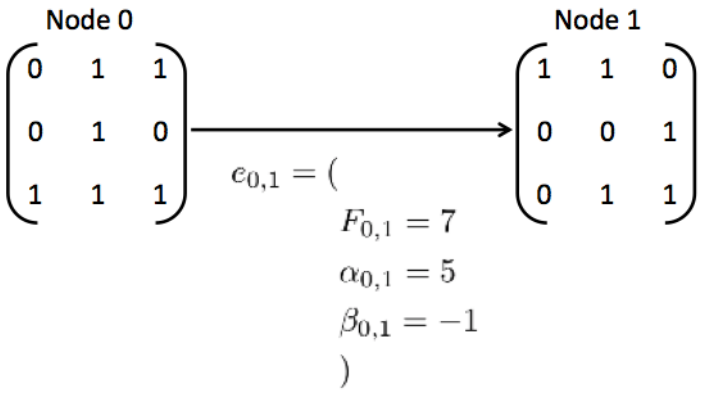

Figure 3.2: *Example of the edge weightings between two nodes (Node 0 and Node 1, where the transition $T_{0,1}$ has occurred seven times) that contain the causality patterns for a system composed of three variables in a USIC-Network model.*

Further three auxiliary functions can be defined *get-f()*, *get-α()*, and *get-β()* that return the three components of an edge weight:

$$\textit{get-f}((R_{l_1}, R_{l_2}, (f, \alpha, \beta))) = f$$
$$\textit{get-}\alpha((R_{l_1}, R_{l_2}, (f, \alpha, \beta))) = \alpha$$
$$\textit{get-}\beta((R_{l_1}, R_{l_2}, (f, \alpha, \beta))) = \beta$$

With this new definition of E and using the previous definition of N the USIC-Network can be defined as USIC-Network$= (N, E)$ for the set of time series variables V.

3.3 Node metrics and network properties using the USIC network

The previously mentioned studies employing complex networks for the evolution of multivariate systems by Jiang *et al*, Yu *et al*, Qi *et al*, Dong *et al*, and Yu *et al* [3, 2, 1, 56, 37] also subsequently analyse their constructed networks. This analysis takes a number of forms but primarily it determines properties and metrics of the network that can then be used to describe the evolution of the original system. This

analysis can be characterised as the following types: determination of the degree of the nodes, the distribution of this degree, and the comparison of this to standard network models, such as a scale free network [3, 2, 1, 56]; the distribution of edge weights (frequency), comparison to power law distribution, and total level of self-loops within the network (i.e. when a causality pattern does not change over a time step) [3, 56]. Common network methods are used, such as the closeness of the network [1], and the betweenness of nodes [2, 1, 56, 37]. Clustering is also often employed for more in-depth analysis of the behaviour of the system, for example modularity-based methods [3, 139]. These approaches do not explicitly allow for description of the evolution in terms of the changes in the structure of the causality pattern. The proposed USIC-Network encodes this information allowing it to become more accessible, and hence more amenable to analysis. Numerous approaches, including many classic examples, can be applied to the USIC-Network to make use of this data. However, it has been shown that unique network methods are often needed to uncover specific system related knowldge [143], for example generic network methods may look for high degree nodes, however for a specific system low degree nodes maybe just as important [144]. Therefore, in this section it is chosen to propose five properties/metrics of the USIC-Network, based on knowldge of what the network represents, that can be used to discover informative descriptions of the evolution of the original system, to allow for a richer overall understanding of the system. These are meant for any system that can be expressed via the USIC-Network model, however for context each approach is motivated with an example use case, demonstrating behaviour it is amenable to uncover.

3.3.1 Node metrics

The first set of properties/metrics are specific to each node, and do not require any grouping

3.3.1.1 Pattern stability

Interactions between financial instruments are often used when considering financial risk, and this is particularly true when portfolio construction is concerned.

Within a portfolio, assets that are correlated can be seen as increasing risk, as this limits diversification and makes the portfolio less resistant to shocks: conversely assets that are negatively correlated are some times used to mitigate risk by the increase in one counteracting the decrease in the other. Similarly the causality structure within a portfolio (where known) can be carefully selected to reduce overall risk, leading to the effectiveness of the portfolio being heavily connected with this causality structure. Due to this it is important for portfolio managers that this causality structure does not change as it could force them to have to reconstruct their portfolio, and this motivates a desire for a metric for how likely a system is to maintain its current causality structure. In the context of the USIC-Network this property can be observed using the metric of the likelihood of a transition from a causality pattern to have no Total Difference in Causality, i.e. $\alpha = 0$. This process is often referred to as a self-loop or self transition, where the causality pattern does not change over a time step (loops that taken multiple time steps to return to their original position are not classically considered self-loops and are not of interest here).

The concept of self-loops occurs in many branches of research, for example: (i) in modularity techniques, for the purpose of separating the internal links of a community from those connecting the community to others [145, 146, 147]; (ii) in Markov chains, where each state will have a probability of transitioning to itself, and is often a pre-defined value [148]; and (iii) in complex networks, where the overall percentage of self-loops for a network is analysed, and specific self-loop edges with a high frequncy are discussed [56]. For the above case this chapter is interested in a self-loop measure that can be discussed in terms of individual nodes, where each node has a probability of self-loop associated with it. Due to the variety of self-loop usage in the literature here a definition of this metric for use with the USIC-Network is given.

This problem can be seen as similar to that of determining the degree of a node, though unlike the degree that is often for a directed network separated out into a *in* and *out* degree it is desired to separate this further into a self-loop degree. Here a

definition of the chance of a self loop is presented, using the following functions that are applicable to any given representative node in network $\mathcal{N}$:

- The function $k^{loop}()$ gives the number of times a representative causality pattern R_l transitions over one time step to the same representative causality pattern:

$$k^{loop}(R_l) = get\text{-}f(e_{l,l})$$

- The function $k^{out}()$ gives the number of times a representative causality pattern R_l transitions over one time step to any next representative causality pattern in N (including itself):

$$k^{out}(R_l) = \sum_{R_x \in N} get\text{-}f(e_{l,x})$$

- The probability of a self-loop occurring for a representative causality pattern R_l is given by $\Omega^{loop}_{R_l}$, defined by:

$$\Omega^{loop}_{R_l} = \frac{k^{loop}(R_l)}{k^{out}(R_l)}$$

3.3.1.2 Directional change in causality

Causal interconnectivity within the financial markets has been shown to lead to a number of undesirable behaviours for market health when it becomes too high. This includes market behaviours such as crashes, bubbles, and other instabilities in price, leading to fallout that is both difficult and costly for regulators and governments to resolve [149]. It is therefore beneficial for market regulators to have a forecast for how the degree of causal interactions within the market are likely to change, to allow them to enact policy to limit adverse market behaviour before those changes occur. To give an indication of the change (either an increase, decrease, or no change) expected in the total causality at the next time step, it is desirable to have a one-

dimensional value representing the multi-dimensional data (edges) that describes the previous changes from the relevant node.

This problem regards the out links from the node, and their associated β weightings. For a classic network where weightings represent frequency of occurrences this problem can be seen as analogous to the out degree of the node, where the sum of the weights of all out links is calculated. Here a metric is presented that gives an indication of whether causality pattern (node) is likely to increase, decrease, or maintain the same total causality, based on its history. To do this first define the following functions that are applicable to any given representative node in network $\mathscr{N}$:

- An aggregation of the previous transitions of a specific representative causality pattern R_l can be found by summing the frequency of the out edges from that node, $k^{out}(R_l)$ (defined previously).

- To give knowledge of the direction of these transitions a new function $k^{weighted}()$ can be employed, that weights each edge as it is added to the sum by the sign of its β weight (giving information relating to whether the transition increase, decreases, or does not change the total causality). This summation hence will give a value representing the average directional change in causality, which can be taken as a prediction of future directional changes in causality. First define an auxiliary function $sign()$ as follows:

$$sign(x) = \begin{cases} -1, & if\, x < 0 \\ 0, & if\, x = 0 \\ 1, & if\, x > 0 \end{cases} \tag{3.1}$$

  And now define $k^{weighted}()$ as follows:

$$k^{weighted}(R_l) = \sum_{R_x \in N} (get\text{-}f(e_{l,x}) \times sign(get\text{-}\beta(e_{l,x}))$$

- The value of $k^{weighted}()$ can be heavily skewed by the number of edges, and

the frequency of those edges, making comparison of this value between different nodes difficult. To account for this a normalised measure of this value is proposed, $\Omega_{R_l}^{directed}$. This is normalised by dividing $k^{weighted}()$ by the total number of out transitions given by $k^{out}()$, and hence gives a value between -1 and 1.

$$\Omega_{R_l}^{directed} = \frac{k^{weighted}(R_l)}{k^{out}(R_l)}$$

3.3.2 Network properties

The second set of properties/metrics are network specific, and define groupings of nodes within the network.

3.3.2.1 Noise Clusters

In feature selection one potential aim is to determine from a set of variables those variables that Granger-cause a target variable. These variables can then be used to train a machine learning model for the purpose of forecasting the target variable. It is important for the correct subset of variables to be selected, with too large a subset increasing the cost and time of training, and too small a subset offering inferior forecasting results [150]. For feature selection it hence may be considered important to find the maximum number of potentially casual variables for a single underlying causality structure, while still minimising the total number of selected features, to reduce the number of retraining periods and present the best set of information for machine learning model to be trained on. However, many real systems, such as the financial markets, contain messy data and are susceptible to noise within their causality calculations. For example, the presence of noise can cause deviations in causality patterns and may cause a system with a singular underlying causality structure to be represented through analysis by a number of causality patterns. Therefore it can be considered beneficial to determine all potential causality patterns that may represent an underlying causal structure, allowing features to be selected from these patterns as a group, rather than just a single pattern.

Using the USIC-Network representation this problem can be seen as cluster-

ing nodes based on their Total Difference in Causality (α), with the objective of clustering together nodes with a small Total Difference in Causality. This problem can be seen as similar to those dealt with by density-based clustering methods, as this problem relates to a "small distance" measure and not a "large frequency" measure as used in alternative clustering approaches [151, 152, 153]. An initial starting point for density-based clustering is the single-linkage model, a hierarchical method that operates by grouping nodes within a given "distance" of each other and then increasing this distance till all nodes are clustered [154, 151]. A notable expansion to this model was introduced by Wishart to eliminate a "chaining" effect that could lead to the linkage of widely spaced nodes via a chain of more closely connected nodes. This expansion introduced the idea of a minimum number of nodes within a set "distance" from each node in the network, a node that does not meet this minimum is then removed from a cluster [155]. The single-linkage model and the Wishart expansion act as first step for the desired clustering, however in this use case not all nodes need be clustered and due to links not existing between all nodes the colorblue usage of minimum degree does not apply. These differences in scope therefore warrant further expansion of this model for this problem.

To define this clustering property, based on some notion of measurement-based noise or variation in the network, within a USIC-Network the following steps are taken:

- Here a parameter $Par^{noise}_{max\,\alpha}$ is introduced, which defines a maximum amount of deviation that can be expected between separate measures of the same causality pattern in a system (e.g. for a physical system a user may know that their measurement tools have an associated error and hence this metric embodies how that error translates to the measurement of a causality pattern). In the USIC-Network this measure takes the form of an α value, being the maximum expected α value between two causality patterns that could be considered the same within deviation. The exact choice of this parameter is complex and heavily system dependent, so this is left as a user defined value (a more formal definition of its exact value is considered to be out of the scope of this

thesis).

- To incorporate the notion of noise that $Par^{noise}_{max\,\alpha}$ provides into the complex network, an initial USIC-Network, $\mathscr{N} = (N,E)$, is updated by removing both edges with an α greater than $Par^{noise}_{max\,\alpha}$ and any nodes that are now unconnected. This defines a new network, $\mathscr{N}^{'} = (N^{'},E^{'})$, as follows:

$$E^{'} = \{e_{l_1,l_2}\} \forall e_{l_1,l_2} \in E \mid get\text{-}\alpha(e_{l_1,l_2}) \leq Par^{noise}_{max\,\alpha}$$
$$N^{'} = \{R_{l_1}\} \forall R_{l_1} \in N \; and \; \exists R_{l_2} \in N \mid e_{l_1,l_2} \in E^{'} \vee e_{l_2,l_1} \in E^{'}$$

- The network $\mathscr{N}^{'}$ now only contains transitions that are within this defined noise/deviation range. For this network this chapter hypothesise that causality patterns (nodes) that relate to each other and are just products of noise/error will exhibit some clustering behaviour. To discover this clustering a popular method known as modularity is employed, selecting the Clauset-Newman-Moore greedy modularity maximization algorithm [156, 127, 157]. For this the edge weightings and self-loops are not considered, and a set of non-overlapping clusters are produced. Although this algorithm may not be appropriate for cases where noise clusters overlap, it is assume that for most real world systems noise clusters will be adequately spaced, partly due to the binary values (quantisation) of causality, and hence overlapping clusters will not be a consideration in practice. It should also be noted that if $Par^{noise}_{max\,\alpha}$ is set such that all links are included, e.g. no consideration of α is taken, this approach reduces to the clustering approach employed by Jiang *et al* [3]. Each cluster produced is labelled as Λ^{noise}_{i}, where these sets, and the set Λ^{noise} comprising all these clusters, can be expressed as:

$$\Lambda^{noise} = \{(X,i)\} \forall X \subseteq N^{'}$$
$$\Lambda^{noise}_{i} = X \mid (X,i) \in \Lambda^{noise}$$

3.3.2.2 Regimes of Total Causality Level

The concept of regimes within the evolution of a system is a popular topic in many fields [149, 83, 158]. The type of grouping defining a regime can take a number of forms, in the context of the USIC-Network a potential construction for a regime is a grouping of system states (causality patterns) that have the same total causality (i.e. the level of interaction within the system is the same throughout the regime). A real world situation where this type of regime may be of interest is the behaviour between market makers during a flash crash. A particular example of this is the 2010 flash crash that has been connected to hot potato trading between market makers [159]: during a period of hot potato trading the amount of interaction between the market makers is likely to increase [160]. Therefore the regime of the total causality between market makers is likely to be different during periods of calm compared with periods of market instability such as hot potato trading, and detecting these changes in regime may be an indicator of coming instability [149, 83, 158].

This chapter is particularly interested in causal regimes whose internal edges have no net change in total causality. To this end a regime is defined as a set of nodes in the USIC-Network whose internal edges all have $\beta = 0$. This means that the total causality for each node in a regime is constant, but the causality structure may not be.

For this clustering, an approach inspired by single linkage can be used, where the minimum "distance" is set as $\beta = 0$ and does not increase iteratively [154, 151]. To tackle the potential problem of chaining as previously described, a parameter is introduced,$Par^{regime}_{minfreq}$ to represent the lowest allowed frequency level for a link. This allows for the removal of "pathways of low travel" between high travelled clustered regions. Though this can be set to some calculated value, such as the average frequency of edges, hence removing all edges below that, it is chosen to leave this as user selected as it is felt that an appropriate value will be system-dependent. The process to find these causal regimes can be described as follows:

- From an initial USIC-Network $\mathscr{N} = (N, E)$, a new network can be derived where all edges with a β value not equal to zero and a frequency value less

than $Par^{regime}_{minfreq}$ are removed. This network is labelled as $\mathscr{N}^{''} = (N^{''}, E^{''})$ and define it as follows:

$$E^{''} = \{e_{l_1,l_2}\} \forall e_{l_1,l_2} \in E \mid get\text{-}f(e_{l_1,l_2}) \geq Par^{regime}_{minfreq} \wedge get\text{-}\beta(e_{l_1,l_2}) = 0$$
$$N^{''} = \{R_{l_1}\} \forall R_{l_1} \in N \; and \; \exists R_{l_2} \in N \mid e_{l_1,l_2} \in E^{''} \vee e_{l_2,l_1} \in E^{''}$$

- The network $\mathscr{N}^{''}$ is constructed in such a way that if the desired causal regimes exist they will be components (groups of connected nodes that are not connected to the rest of the network). Therefore to find these causal regimes one extracts these components, labelling each component as a separate set Λ^{regime}_i. These sets, and the set of all these regimes can be expressed as:

$$\Lambda^{regime} = \{(X,i)\} \forall X \subseteq N^{''}$$
$$\Lambda^{regime}_i = X \mid (X,i) \in \Lambda^{regime}$$

3.3.2.3 Net causality change pathways

A broad type of structural feature that naturally arises when discussing complex networks is a "pathway", a series of sequential nodes connected by edges, that defines some route through the network. Specific instances of pathways can be defined in numerous ways, for example in the context of the USIC-Network these definitions could be based on the F, α, or β weightings. Pathways based on F or α naturally lead to the implementation of either minimum F constrains (i.e. highly travelled pathways) or a maximum threshold value for α (i.e. closely "spaced" pathways) However, pathways based on β can take more interesting formulations and to the authors knowledge have not been previously explored.

In the context of this thesis β is an interesting base for pathway construction for the previously discussed regimes (defined with edges of $\beta = 0$), with a system moving from one regime to another having to change its total causality and hence have a $\beta \neq 0$ between a node in one regime and a node in the other regime. The transition from one regime to another may occur over one time step or over many

time steps. In the latter case this results in a multi-step pathway existing between the two regimes. These pathways defining a change in total causality are a general structural feature of complex causality networks, whose start and end nodes need not be members of a regime.

Here a type of pathway is defined based on the β weighting of the network. This pathway is a structure that moves the causality pattern from one level of total causality to another, with no self-loops that might be deceptive during analysis. For this initial discussion of these types of pathways it is chosen to maintain the sign of the β weighting throughout the pathway (i.e. a pathway will either be composed only of edges weighted as $\beta > 0$ or only of edges weighted as $\beta < 0$). This decision is well motivated for pathways between regimes, as defined above, since as soon as a $\beta = 0$ edge is found a new regime may have been reached.

In defining these pathways the two following constraints are considered:

- As explained above, the initial interest is in pathways with a β constraint: constructed of either only edges containing $\beta > 0$ or only edges containing $\beta < 0$.

- For the USIC-Network (and most complex network representations) an edge is "representative" of one or many transitions at different times between observed causality patterns. As a result, these pathways are statistical in nature and don't necessarily represent a timed sequence of transitions: it is not necessarily true that two edges in a pathway occur in the same temporal order as their representative transitions were observed. Therefore typically only "common" pathways are considered by utilising a further constraint, which is to consider only edges with a frequency F that exceeds some threshold. The threshold will be system-dependent and experiment-dependent and therefore let it be expressed as the user-defined parameter $Par^{path}_{minfreq}$.

Based on these constrains a formal definition of the pathways described above is presented for the USIC-Network $\mathscr{N} = (N, E)$. It is elected to split this formal definition of these pathways into two components, the pathways themselves them

selves, and the first and last edges in each pathway. Label the set of all first edges (the head edges) as H and the set of all last edges (the tail edges) as T (where these are both subsets of E).

Employing these head and tail sets define each pathway as a 2-tuple of a set X (the pathway) and a label i such that X is a subset of E, and there exists exactly one first edge $e_{a,b}$ in H and exactly one last edge $e_{c,d}$ in T (where neither edges are self-loops, so $b \neq a$ and $d \neq c$) such that for all edges $e_{m,n}$ in X three conditions must hold:

1. the edge $e_{m,n}$ must either be the first edge $e_{a,b}$ or there must exist a unique edge $e_{p,q}$ in X such that $e_{p,q}$ directly precedes $e_{m,n}$ (i.e. $m = q$),

2. the edge $e_{m,n}$ must either be the last edge $e_{c,d}$ or there must exist a unique edge $e_{p,q}$ in X such that $e_{p,q}$ directly succeeds $e_{m,n}$ (i.e. $n = p$), and

3. all edges $e_{p,q}$ must have the same β sign, $sign(get\text{-}\beta(e_{p,q}))$ (where this sign will be defined via the head and tail edges).

Using the notation $\exists!$ to denote uniqueness quantification (i.e. $\exists! x$ means "there exists exactly one x"), define the set of all such labelled pathways $\Gamma^{H,T}$ in a given set of edges E with initial (head) edge in H and final (tail) edge in T as follows (with $\mathscr{N} = (N,E)$ assumed from here onwards):

$$\begin{aligned}\Gamma^{H,T} =& \{(X,i)\} \forall X \subseteq E \mid H \subseteq X \wedge T \subseteq X \wedge i \in \mathbb{N} \\ & \wedge (\exists! e_{a,b\neq a} \in H \; and \; \exists! e_{c,d\neq c} \in T \mid \forall e_{m,n} \in X \; ((m=a) \vee (\exists! e_{p,q} \in X | m = q)) \\ & \wedge ((n=d) \vee (\exists! e_{p,q} \in X | n = p)) \\ & \wedge (sign(get\text{-}\beta(e_{m,n})) = sign(get\text{-}\beta(e_{a,b}))) \\ & \wedge (get\text{-}f(e_{m,n}) \geq Par^{path}_{minfreq}))\end{aligned}$$

One can select a single pathway in $\Gamma^{H,T}$ by referencing its label as follows:

$$\Gamma^{H,T}_i = X \mid (X,i) \in \Gamma^{H,T}$$

Since the interest is in both $\beta > 0$ pathways and $\beta < 0$ pathways, four sets of edges are defined: head edges and tail edges for $\beta > 0$ pathways and head edges and tail edges for $\beta < 0$ pathways. Of particular interest are maximal pathways in a system: for example, a head node of a maximal $\beta > 0$ pathway will have no in-edges with $\beta > 0$ and a tail node of a maximal $\beta > 0$ pathway will have no out-edges with $\beta > 0$, and these constraints can be incorporated into the definitions of the sets of candidate head edges and tail edges for $\beta > 0$ pathways. Similarly, of interest are maximal $\beta < 0$ pathways whose head nodes have no in-edges with $\beta < 0$ and whose tail nodes have no out-edges with $\beta < 0$. For all these sets the frequency constraint discussed earlier also still applies.

The sets of head and tail edges for pathways are labelled $\beta > 0$ as $H+$ and $T+$ respectively, and for pathways of $\beta < 0$ as $H-$ and $T-$ respectively. For example, $H+$ is defined to be the set of all edges $e_{i,j}$ such that $e_{i,j}$ is in E, the edge has $\beta > 0$, and any edge $e_{k,i}$ whose end node is the same as the start node of $e_{i,j}$ has a $\beta \not> 0$. Therefore it can be written:

$$
\begin{aligned}
H+ = \{e_{i,j}\} \mid e_{i,j} \in E \;&\wedge\; \textit{get-}f(e_{i,j}) \geq Par^{path}_{minfreq} \\
&\wedge\; \textit{get-}\beta(e_{i,j}) > 0 \\
&\wedge\; \forall e_{k,i} \in E\ \textit{get-}\beta(e_{k,i}) \not> 0 \\
T+ = \{e_{i,j}\} \mid e_{i,j} \in E \;&\wedge\; \textit{get-}f(e_{i,j}) \geq Par^{path}_{minfreq} \\
&\wedge\; \textit{get-}\beta(e_{i,j}) > 0 \\
&\wedge\; \forall e_{k,i} \in E\ \textit{get-}\beta(e_{k,i}) \not> 0 \\
H- = \{e_{i,j}\} \mid e_{i,j} \in E \;&\wedge\; \textit{get-}f(e_{i,j}) \geq Par^{path}_{minfreq} \\
&\wedge\; \textit{get-}\beta(e_{i,j}) < 0 \\
&\wedge\; \forall e_{k,i} \in E\ \textit{get-}\beta(e_{k,i}) \not< 0 \\
T- = \{e_{i,j}\} \mid e_{i,j} \in E \;&\wedge\; \textit{get-}f(e_{i,j}) \geq Par^{path}_{minfreq} \\
&\wedge\; \textit{get-}\beta(e_{i,j}) < 0 \\
&\wedge\; \forall e_{k,i} \in E\ \textit{get-}\beta(e_{k,i}) \not< 0
\end{aligned}
$$

Therefore two sets of pathways can be written: $\Gamma^{H+,T+}$ is the set of maximal $\beta > 0$ pathways, and $\Gamma^{H-,T-}$ is the set of maximal $\beta < 0$ pathways. Individual pathways would for example be written as $\Gamma_i^{H+,T+}$ and $\Gamma_i^{H-,T-}$.

For the specific problem of the algorithmic detection of pathways a simple but effective approach is to start with a node in H and look along connected out edges to find a new node, and then repeat this process on the new node until no new nodes can be found (where each out edge must meet the requirements specified above). In the case where a node connects directly to multiple other nodes a pathway would be constructed for each possible choice (it should be noted that this may give rise to pathways with overlapping regions). To discover the set of all pathways this process would be applied to all possible start nodes.

3.3.3 Summary of properties and metrics

This section has presented a number of metrics/properties that can be used to characterise a system represented by the USIC-Network, and hence describe the behaviour of the underlying system. To review: two node level metrics were presented, pattern loop chance and directional change in causality, that both relate to the transitional likelihood of the node, with overlap in their presented information existing i.e. if a node is likely to self-loop then its directional change will be near zero. Two stationary node grouping properties were presented, noise grouping and regime grouping, that both group nodes together based on similarity related to the underlying causality patterns, both behavioural structures within the network. Finally one dynamic property was presented that categorised a type of temporal movement through the network, related to the a consecutive increase or decrease change in the total causality. These presented metrics/properties each can be used to characterise a system individually, though depending on their results they may be taken together to give a richer description of the systems dynamic behaviour.

3.4 USIC network validation on synthetic data

To demonstrate the characterisation of the properties presented above, and to validate the ability of the metrics to observe those properties, this section applies the

metrics to two types of synthetic networks; a random network model and edited random networks, designed to contain artefacts appropriate to the properties.

A random ErdősRényi model is used for generating the random network. This is done by predefining a number of nodes, whose causality pattern is randomly selected without repetition. A time series path of nodes (series of causality patterns) is produce by taking an initial node and then selecting the next node uniformly from all nodes in the system (where the probability of selecting node j while at node i is $T_{i,j}$), and then repeating this process till a path of desired length is achieved. This path can then be transformed into a network following the method described in Section 3.2 to produce a random network. Following this procedure the following parameter selections are made: Four variables are chosen to make up the underlying system, leading to each node representing a 4×4 causality pattern. A hundred nodes (causality patterns) are randomly produced (without repetition of their causality patterns), to create a hundred potential system states for the system to transition between. The network is then populated by by generating a path of ten-thousand steps with the transition probabilities $T_{i,j} = \frac{1}{100}\ \forall i, j$.

Characterisation of this random network is then achieved by investigating the proposed properties and metrics occurrence within it. For this the parameters of properties and metrics are set as follows: $Par^{noise}_{max\alpha} = 2$, $Par^{regime}_{minfreq} = 10$, $Par^{path}_{minfreq} = 10$, and $Par^{path}_{minlen} = 3$.

The results of these methods can be seen in Fig. 3.3, these results are as expected for a random network, not demonstrating any significant behaviour of interest. As may be expected for a random network there exists few noise clusters, that in total account for 10% of the networks nodes, with no single cluster accounting for more then 3% of the network. No Pathways or Regime Clusters were detected. The Causal Direction appears to be well distributed, not demonstrating any strong behaviour. The Loop Chance sits around zero, which is expected when each node only has a one in a hundred chance of transitioning to itself.

To confirm that the metrics and properties manifest as discussed above when present in a network, a Monte Carlo experiment is run for each approach separately.

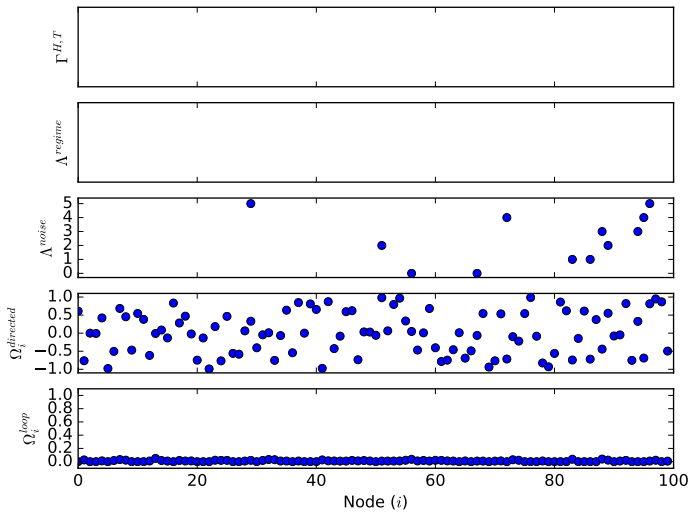

Figure 3.3: *Results of analysis methods for each node (0-99) on a random network.* Ω_i^{loop} *and* Ω_i^{sign} *display the associated value for each node,* Λ^{noise}*,* Λ^{equ}*, and* Γ *display the cluster/Pathways if any a node is a member of (node cluster/pathway label starts at 0).*

With the detection of these metrics and properties in these Monte Carlo experiments indicating the success of their definitions, i.e. that these definitions cover the existence of these behaviours when defined through the statistics of a networks formulation. These experiments are set up by defining an edited network, where this network is generated following the same procedure as above but during its creation an artefact pertaining to the properties or metrics of interest is implanted (described below). For each metric and property the experiment is run ten times, with the generation, probabilities, and networks being independent on each run. For each run the artefact is implanted on the same node numbers to aid in comparison between the runs and so that any unintended effects on the networks are not conflated with intentionally edited nodes (this does not impact the experiment in any way but visually since all the nodes are randomised for each run, and hence the node IDs are only for reference within a single experiment).

These artefacts are inserted by selecting a set of nodes and editing the causality patterns and the probability of transition between them, such that if a node was edited to have a 50% probability to transition to a certain other node, the remain-

ing 50% would be equally divided between all other nodes. The rest of the graph remains random, where all other nodes' causality patterns were randomly generated and the probability of transition between nodes was uniform for all non-edited probabilities.

Only one artefact is implanted in each network, with the node number being used for references in later visualisations and discussions but having no other meaning. The results of these Monte Carlo experiments are shown as the node (its ID) vs the metric or property being investigated (i.e. if it belongs to a regime, noise cluster, etc) to demonstrate the existence and detection of the implanted artefact. For each of the ten runs, and for each experiment, one run is selected and for the first hundred time steps (out of the ten thousand time steps run for) the node (its ID) and causality links of the causality pattern that the system is in are shown, to demonstrate the evolutionary behaviour of the system.

The five metrics assessed in this validation are (i) Ω^{loop}, to determine if it detects pattern stability, (ii) Ω^{sign}, to determine if it detects directional change in causality, (iii) Λ^{noise}, to determine if it detects noisy regimes, (iv) Λ^{equ}, to determine if it detects net causality equilibrium, and (v) $\Gamma^{H,T}$, to determine if it detects net causality pathways. For each of these the associated embedded artefact is defined below, together with its validation results.

Pattern stability. For pattern stability nodes 0-4 are edited, only the transition probability of these nodes was edited to $T_{0,0} = T_{1,1} = T_{2,2} = T_{3,3} = T_{4,4} = 0.7$, so that they each had a 70% chance of a self loop.

The detection for each run is shown in Fig. 3.4, where only the edited nodes showed any signs of being affected and are easily discernible from the non-edited nodes. The first one hundred time steps are shown in Fig. 3.5 appearing mostly random but showing a few instances where the pattern stability artefact occurs (e.g. from time steps 10-14 and 42-44).

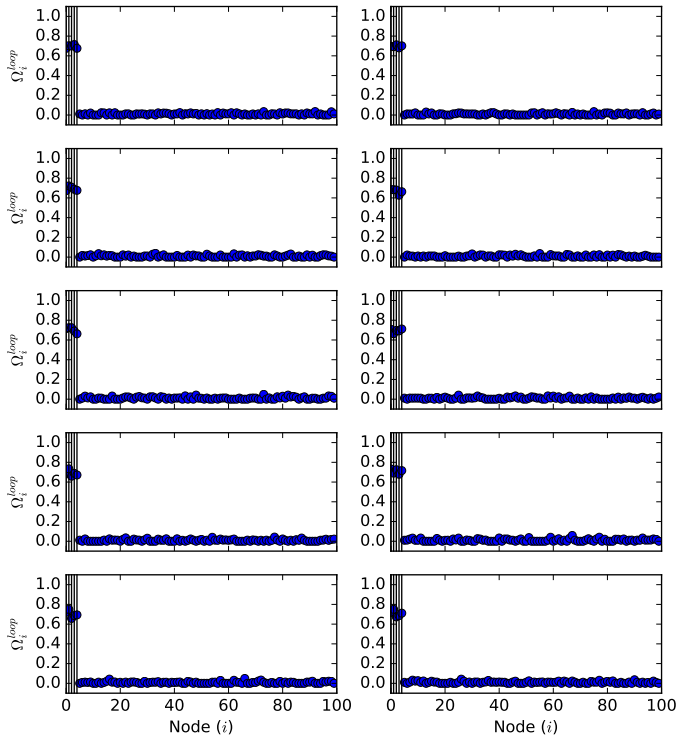

Figure 3.4: *Results of analysis methods for pattern stability, for each node (0-99) for the ten runs of the Monte Carlo experiment, the x-axis represents the node ID and the y-axis the loop probability Ω_i^{loop}. The edited nodes are marked with vertical lines representing where the pattern stability behaviour is expected to occur.*

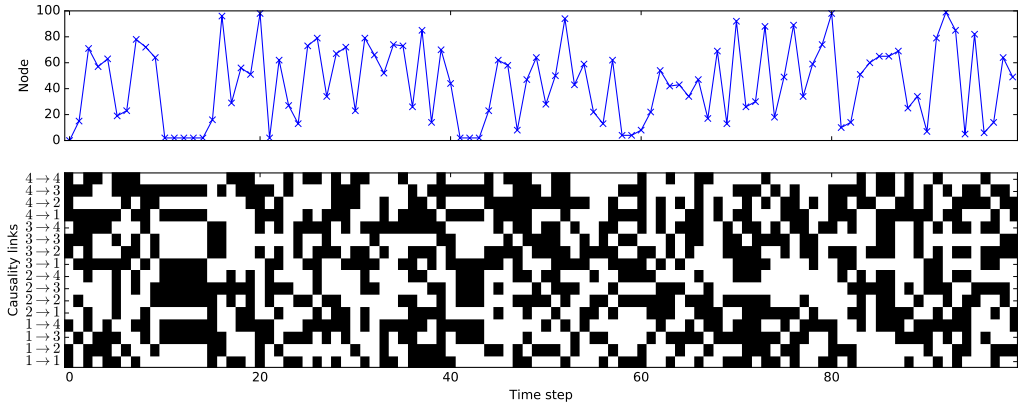

Figure 3.5: *Causality pattern of first hundred time steps of a single run shown in Fig 3.4, showing node ID (top, with continuous line for readability), and causal links (bottom), with a black square representing a link. Pattern stability is seen from time steps 10-14, 41-43, 58-59, and 85-86.*

Directional change in causality. For directional change in causality nodes 29-33 are edited so that they contain 8, 10, 12, 14, and 16 causality links respectively. Their transition probability was also edited such that $T_{29,32} = T_{29,33} = T_{30,32} = T_{30,33} = T_{32,29} = T_{32,30} = T_{33,29} = T_{33,30} = 0.35$ and $T_{31,29} = T_{31,30} = T_{31,32} = T_{31,33} = 0.2$, so that it is probable 29 and 30 will have a positive Causal Direction, 32 and 33 should have a negative Causal Direction, and 31 will have an approximately neutral Causal Direction (though for 31 the 20% random transition probability will likely shift it above or below a 0 Causal Direction, but it is still expected to sit between 29-30 and 32-33).

The detection for each run is shown in Fig. 3.6, this demonstrates random behaviour for the majority of the nodes (as expected) but shows the artefact appearing for the edited nodes in all runs (nodes 29 and 30 have a positive Causal Direction, 32 and 33 have a negative Causal Direction, and 31 has a Causal Direction between 29-30 and 32-33). The first one hundred time steps are shown in Fig. 3.7 appearing random.

Noisy regimes. For noisy regimes nodes 49-53 were edited to all be $\alpha = 1$ from a random causality pattern (that was not manually placed into the network, this selection of $\alpha = 1$ defines these nodes as "close" to the random node such that they could represent a noisy regime). Their transition probability was also changed to $T_{49,50} = T_{49,51} = T_{49,52} = T_{49,53} = T_{50,49} = T_{50,51} = T_{50,52} = T_{50,53} = T_{51,49} = T_{51,50} = T_{51,52} = T_{51,53} = T_{52,49} = T_{52,50} = T_{52,51} = T_{52,53} = T_{53,49} = T_{53,50} = T_{53,51} = T_{53,52} = 0.15$, so that the selected nodes are most likely to transition to each other, who are other close nodes, creating a noise regime.

The detection for each run is shown in Fig. 3.8, with the artefact (nodes 49-53 belonging to a single noise regime) being detected in all runs, there are also a number of other noisy regimes that appear, which is not unexpected for a random network and this type of behaviour. The first one hundred time steps are shown in Fig. 3.9 appearing mostly random but showing some repeated similar causality patterns when the system enters a noise regime e.g. at time steps 66-70 and 80-85.

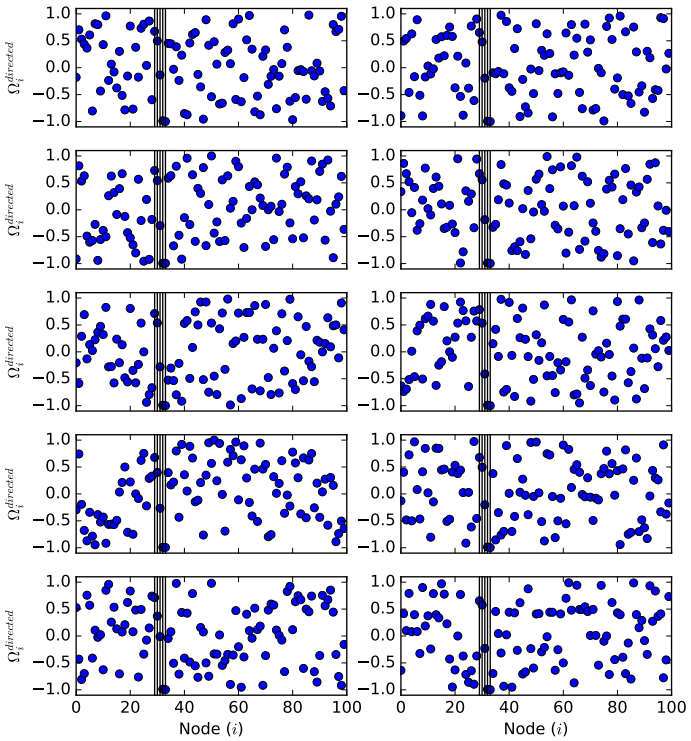

Figure 3.6: *Results of analysis methods for directional change in causality, for each node (0-99) for the ten runs of the Monte Carlo experiment, the x-axis represents the node ID and the y-axis the expected causality change Ω_i^{sign}. The edited nodes are marked with vertical lines representing where the directional change in causality behaviour is expected to occur.*

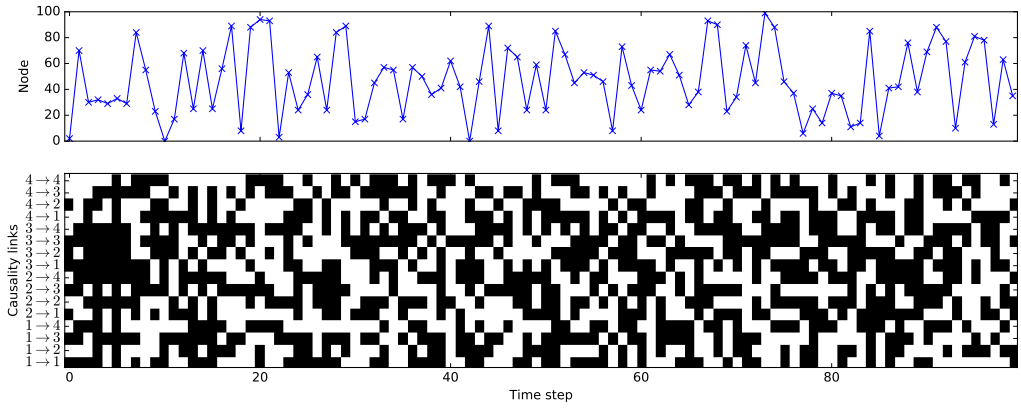

Figure 3.7: *Causality pattern of first hundred time steps of a single run shown in Fig 3.6, showing node ID (top, with continuous line for readability), and causal links (bottom), with a black square representing a link.*

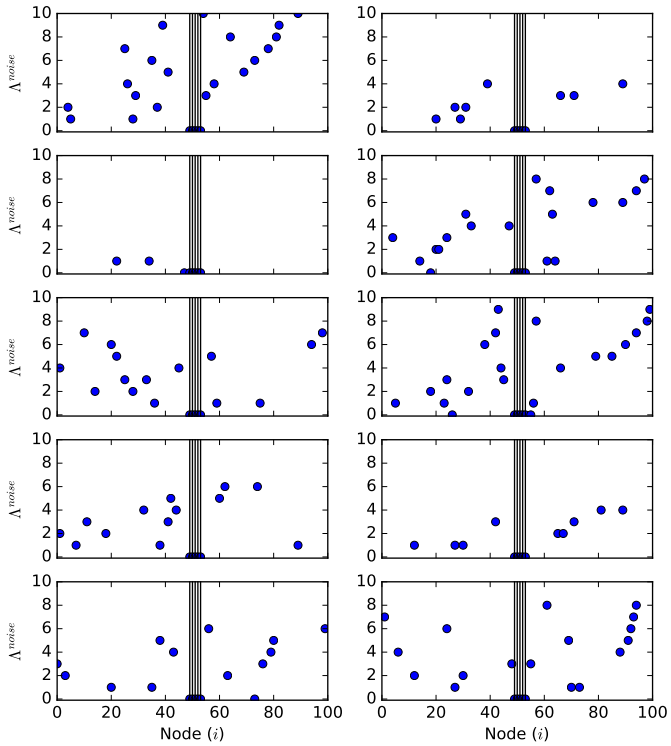

Figure 3.8: *Results of analysis methods for net causality equilibrium, for each node (0-99) for the ten runs of the Monte Carlo experiment, the x-axis represents the node ID and the y-axis the cluster the node is in (if it is in one) Λ^{noise}. The edited nodes are marked with vertical lines representing where the noisy regimes behaviour is expected to occur.*

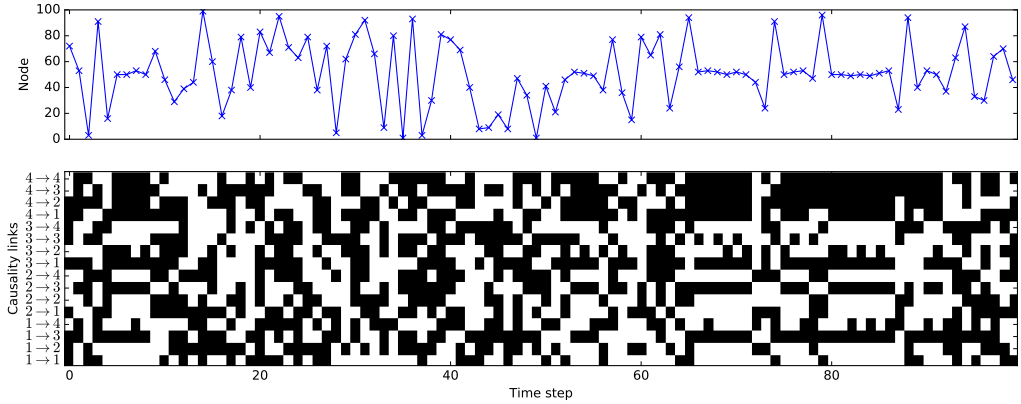

Figure 3.9: *Causality pattern of first hundred time steps of a single run shown in Fig 3.8, showing node ID (top, with continuous line for readability), and causal links (bottom), with a black square representing a link.*

Net causality equilibrium. For net causality equilibrium nodes 69-73 were edited so that they all contained 5 causal links, though their structure was left random. Their transition probability was also edited to $T_{69,70} = T_{69,71} = T_{69,72} = T_{69,73} = T_{70,69} = T_{70,71} = T_{70,72} = T_{70,73} = T_{71,69} = T_{71,70} = T_{71,72} = T_{71,73} = T_{72,69} = T_{72,70} = T_{72,71} = T_{72,73} = T_{73,69} = T_{73,70} = T_{73,71} = T_{73,72} = 0.15$, so that these nodes form a group that is more likely to transition internally than externally.

The detection for each run is shown in Fig. 3.10, with the artefact (nodes 69-73 belonging to the same cluster) being detected in all runs. The first one hundred time steps are shown in Fig. 3.11 appearing random.

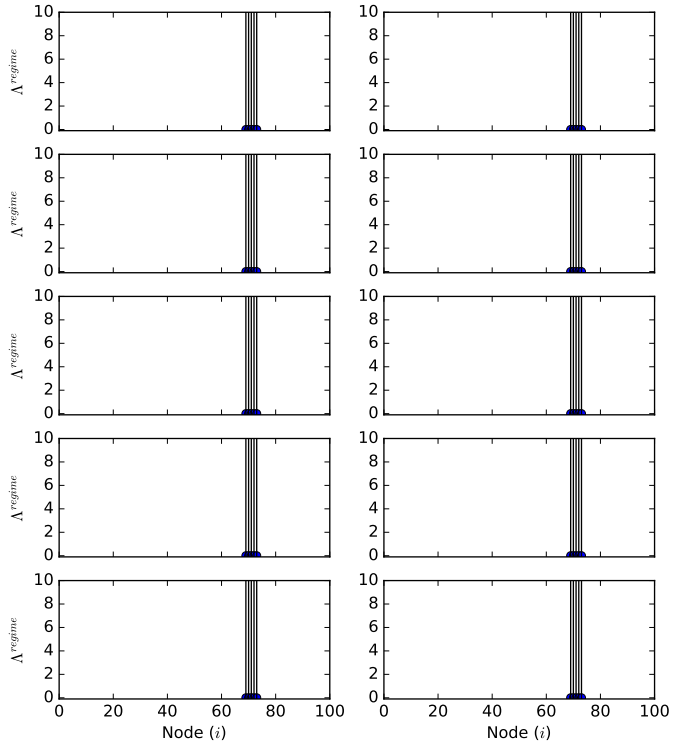

Figure 3.10: *Results of analysis methods for noisy regimes, for each node (0-99) for the ten runs of the Monte Carlo experiment, the x-axis represents the node ID and the y-axis the cluster the node is in (if it is in one) Λ^{equ}. The edited nodes are marked with vertical lines representing where the net causality equilibrium behaviour is expected to occur.*

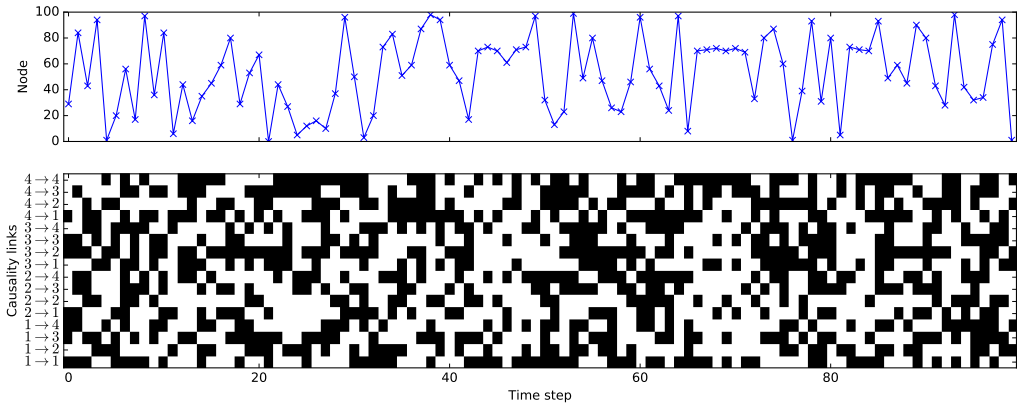

Figure 3.11: *Causality pattern of first hundred time steps of a single run shown in Fig 3.10, showing node ID (top, with continuous line for readability), and causal links (bottom), with a black square representing a link.*

Net causality pathways. For net causality pathways nodes 89-93 were edited so that they contained 2, 3, 5, 6, and 8 causality links respectively. The transition probabilities were also edited to $T_{89,90} = T_{90,91} = T_{91,92} = T_{92,93} = 0.8$, so that these nodes formed a pathway from node 89 to node 93.

The detection for each run is shown in Fig. 3.12, with the artefact (nodes 89-93 belonging to a pathway) being detected in all runs. The first one hundred time steps are shown in Fig. 3.13 appearing random.

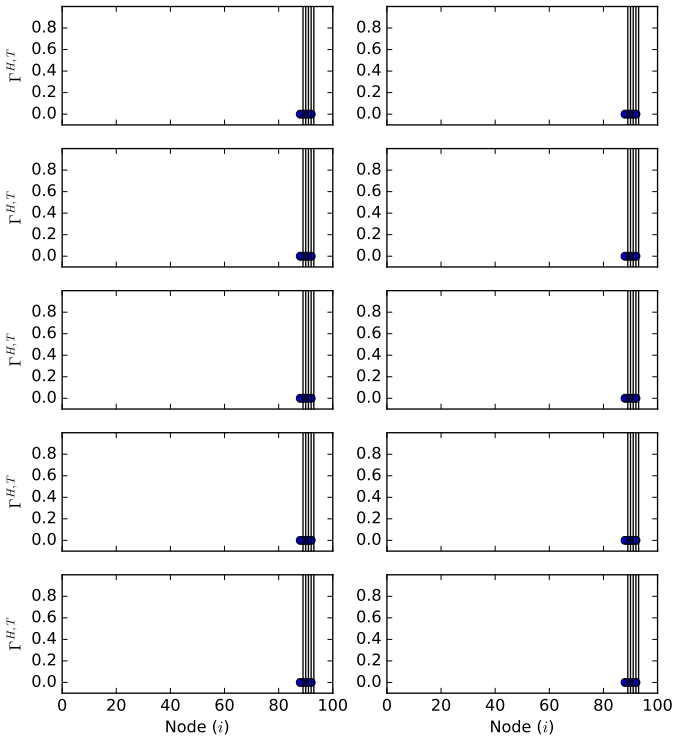

Figure 3.12: *Results of analysis methods for net causality pathways, for each node (0-99) for the ten runs of the Monte Carlo experiment, the x-axis represents the node ID and the y-axis is $\Gamma^{H,T}$, the pathway cluster the node is in (if it is in one). The edited nodes are marked with vertical lines representing where the net causality pathway behaviour is expected to occur.*

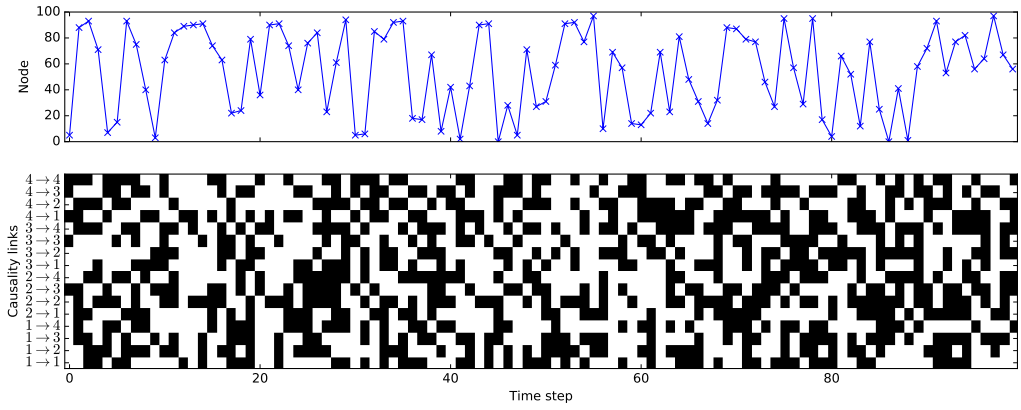

Figure 3.13: *Causality pattern of first hundred time steps of a single run shown in Fig 3.12, showing node ID (top, with continuous line for readability), and causal links (bottom), with a black square representing a link.*

3.5 USIC network case study: the international oil market

In this section the characteristics of the international oil market data presented in Chapter 2.4 are investigated using the methodology presented in this chapter. Using the series of causality patterns for this data defined in Chapter 2.4 the USIC-Network is constructed. Then the analysis introduced in this chapter is applied to this network, using parameters $Par^{noise}_{max\alpha} = 2$, $Par^{regime}_{minfreq} = 1$, $Par^{path}_{minfreq} = 2$, and $Par^{path}_{minlen} = 4$ (this configuration was selected through experimentation), with results for each node in the network shown in Fig. 3.14. From the results in Fig. 3.14

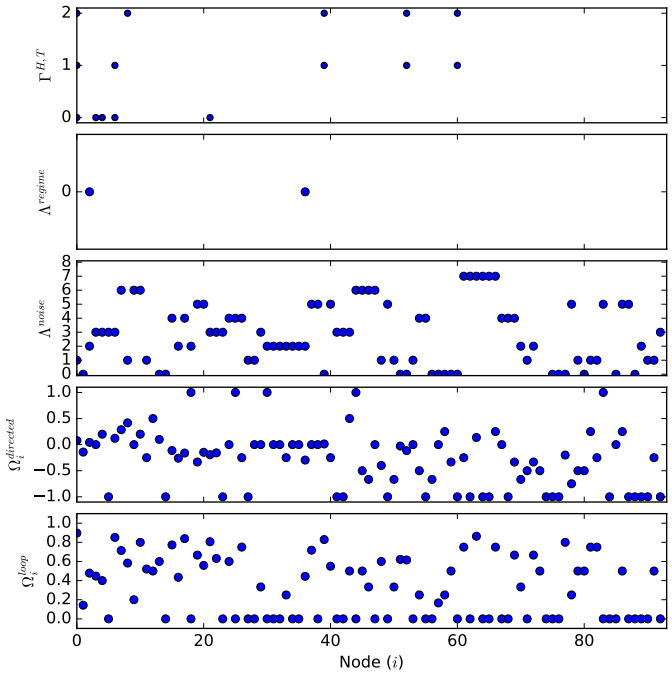

Figure 3.14: *Results of analysis methods for each node on a network of Oil spot prices.* Ω_i^{loop} *and* $\Omega_i^{directed}$ *display the associated value for each node.* Λ^{noise} *and* Λ^{regime} *display the cluster labels of the node if applicable.* $\Gamma^{H,T}$ *displays pathway labels of the node if applicable, these are for both* $\Gamma^{H+,T+}$ *and* $\Gamma^{H-,T-}$ *pathways. Note that the cluster and pathway labels starts at 0.*

looking at the individual analysis methods one can draw the following conclusions:

- On the whole nodes tend to favour having a chance of self-looping ($mean(\Omega_i^{loop}) = 0.32$), however a few nodes have a very high chance of

self-looping. This implies that there exists a few highly stable causality patterns within the evolution.

- The causal direction of nodes during transition is on average close to zero but favours negative transitions ($mean(\Omega_i^{directed}) = -0.27$), showing that on the whole transitions tend to decrease the total causality of the network.

- The majority of the network belongs to one of eight noise clusters (Λ^{noise}) with sizes of: 17, 14, 13, 12, 10, 10, 7, and 6. These clusters are defined such that no causality pattern within them may differ more then 12.5%. This demonstrates that the causality pattern of the international oil market transitions throughout time between groupings of very similar patterns. Implying that the specific causality patterns are non-random, and can be seen as deviations within these clusters.

- The network shows virtually no regime clustering (Λ^{regime}), implying that causality patterns do not maintain their overall causality during transitions in non-self-loop cases. This suggests that the evolution of the system is effected by individual causal links and not the overall causality of the system.

- There only exists a few pathways of Net Change in Causality $\Gamma^{H,T}$ that are longer than a few nodes within the network, implying that changes in net causality primarily take place over a short number of nodes. This indicates a lack of significance in Net Change in Causality in the systems evolution.

To further investigate the behaviour of noise clusters within the network (from the context of the methodology), the cluster the system is in at each time step during its evolution is plotted, results displayed in Fig. 3.15. From these results one can see that cluster 6 is very dominant within the evolution, with the system spending the majority of its time within this cluster. It can also be seen that on the whole when the system leaves this cluster it tends to stay in whichever other cluster it transitions to for a extended number of time steps. This demonstrates that the systems evolution is heavily dominated by these clusters, with causality patterns staying similar for extended time steps.

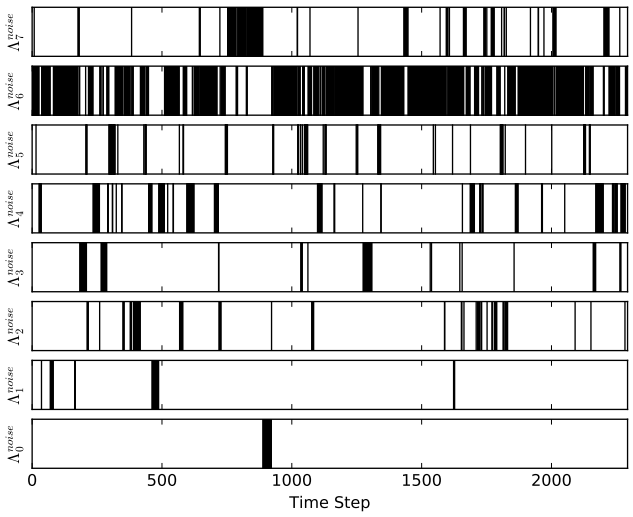

Figure 3.15: *Noise cluster occupancy for the current node against time, for the whole system evolution: a black line indicates the current node is in the indicated noise cluster at the indicated time step.*

The change in the self-loop metric during the evolution is also investigated, through smoothing the results with a rolling average of 50 time steps, shown in Fig. 3.16. The rolling average shows an approximate cyclic pattern to the evolution of this metric, where the system goes through periods of increasing self-loop chance before going through periods of decreasing self-loop chance. This implies that the system state may be moving between regions of stability, with unstable regions in-between.

These results taken together demonstrate that this evolution is highly dependent on the individual links and structure of the causality patterns and not on the overall causality. Furthermore the system favours transitions to causality patterns with a similar structure, illustrated by the evolution being able to be decomposed into a number of noise clusters. Nodes nodes with a high chance of self-looping may be considered more stable aspects of these noise clusters, with nodes with a low chance of self-looping being taken as noise/transition nodes around and between these clusters.

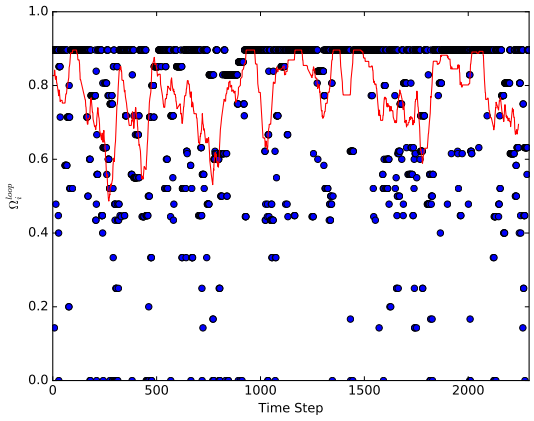

Figure 3.16: *Chance of self-loop at each time step for the current node (blue circles). The red line is the rolling average with a window of 50 time steps for the self-loop chance in the evolution.*

3.6 Discussion

The work presented in this chapter addresses the first research question presented in the Research Objectives, demonstrating that time-varying behaviour of a system can be more fully explained via the inclusion of information on the transition properties between causality patterns. This is achieved through the proposal of: a new methodology for the representation of a time-varying system using a complex network; five new properties/metrics for the characterisation of said complex network; synthetic validation of the methodology; and, investigation of the international oil market using said methodology. Below a more detailed summary of this work is given.

The work presented here aims to expand the field of research, presenting a new methodology for information extraction from evolving causality networks. Exploring the evolution characteristics of time-varying causality relationships holds the potential for a deeper understanding of the dynamics of many complex multivariate systems. This chapter a method to encode the evolution of the interaction dynamics within a multivariate system into a series of causality patterns. The work of authors, such as Jiang *et al* is expanded upon, by transferring these patterns to a multi-weighted directed network, the USIC-Network, capable of containing three key metrics of the evolution: frequency of transition, Total Difference in Causality,

and Net Change in Causality. The addition of the latter two metrics allows for information regarding the change in the underlying causality structure to be encoded into the network. This in turn supports further analysis methods to be performed on this network. Five novel approaches for the analysis of the evolution of interactions within a multivariate system are presented: these methods are based on the presented network model and take advantage of the information of the underlying causality pattern. Prior to the USIC-Network model presented in this chapter, this causality information was not available in a manner suitable for network analysis techniques.

This methodology is validated via application to two networks: a random network, and a random network edited to contain the above properties and metrics. This validation demonstrated that the above properties were discoverable using the metrics and differed from a similar analysis of a random network.

To demonstrate these proposed approaches they are applied to sample data from the international oil market, due to its popularity in research and its known underlying interaction dynamics. For this data a behavioural description not readily discoverable with current approaches to analysis employed in the research of complex network representations of time-varying interaction dynamics was constructed. The primary aspects of this description can be divided into four linked findings offer new understandings of this market data: (i) The transitions over a single time step primarily result in a small change in the overall causality of the system. (ii) The change in the causality pattern from a transition over a single time step changes the amount of causality in the system and not the structure of the causality (i.e. the causality pattern remains mostly similar). (iii) The evolution contains clustering, specifically eight clusters wherein the causality pattern of every member node is very similar (differing by no more than two causality links). (iv) The evolution goes through cycles of “high” and “low” stability (likelihood of self-loop), implying the existence of and movement between favoured causality patterns. These results as a whole can be taken to infer that the system favours a few causality (patterns plus some deviation around these) that it moves between through the addition or subtrac-

tion of a couple causality links. Therefore the structure/layout of the causality of the market is important, with the overall amount of causality in the system playing a less important role for the evolution.

This chapter offers new information on the dynamic behaviour of the international oil market, demonstrating that the underlying structure of the causality patterns has a role within its dynamic behaviour. This chapter also offers the methodology and validation that behaviour can be discovered and described from a low level perspective using a complex network.

Chapter 4

Paradigms of Temporal Dynamics in Time-Varying Systems

Chapter 3 identified that the inclusion of knowledge of the underlying structure leads to the discovery of behaviours within the time-varying dynamics of the system. Specifically, Chapter 3 discovered that the international oil market has a dynamic behaviour that appears to be primarily governed by changes in the causality of individual links rather than changes in the structure of the whole pattern. Based on this analysis, this chapter explores a hypothesis that knowledge of this low-level behaviour can be used to define two paradigms of temporal behaviour for systems of this type. Further, that these paradigms describe the dominate evolutionary aspect of this behaviour that can be used to reduce the complexity of the system during analysis. The applicability of this hypothesis is investigated for the international oil market through a methodology of dimensionality reduction based on the paradigms and a simplistic prediction algorithm for the reduced time series.

4.1 Background and related work

As discussed in Chapter 2, and further expanded upon in Chapter 3, much of the current literature concerned with time-varying multivariate systems takes a complex network approach [33]. Specifically, this work on complex networks primarily focuses on high-level behaviour within the system's dynamic behaviour. However, in Chapter 2 another approach to time-varying multivariate systems was discussed,

this being the representation of the time-varying system as a time-series of causality patterns [74] (or in some cases a few specific causality patterns, but with a temporal ordering considered [23]). This work focusing on the causality patterns (or equivalent interaction structure) often takes the form of comparison or tracking of a feature of the network (pattern) of the causality interactions (further discussed below). Methodologies of this form allow for a time series of the system's time-varying behaviour to be constructed. This time series is an alternative representation of the data than a complex network approach, allowing for different forms of analysis and investigation to be employed, and approaches of this manner should be of interest to those investigating the dynamic behaviour of a time-varying multivariate system (in particular, relating to the underlying causality structure).

Time-series analysis is a mature field consisting of numerous approaches and methods for the investigation of information in this form [6, 7]. In particular, time-series data can be seen as amenable to methodologies for prediction/forecasting [161], with an often employed approach being model fitting. This methodology fits a model to some training segment of the time series and then predicts future values of the series via this model (and its fitted parameters) [162]. The most basic usage of this is linear extrapolation, where a linear model is fitted to the data, with predictions following the resultant line [163].

Most time-series approaches assume that each value within the series is one-dimensional (a single value); this does not naturally fit with the time-series of causality patterns being constructed for time-varying interactions of a system. Therefore it is desirable to represent each causality pattern (a multi-dimensional value) as a one-dimensional value. However, this represents a problem as for a multi-dimensional value, as there does not exist a single one-dimensional value that can be said to fully represent it in all aspects, with Li *et al.* [40] stating that "during transformation process, information losses of multivariate time series become inevitable".

Due to this, several metrics (information aggregation/fusion strategies) have been used within the literature to act as this value [40, 164], such as a one-

dimensional value found via a Principle Component Analysis (PCA) [74], and the total causality of the pattern, found by summing the causality [98, 83]. However, due to this information loss, the analysis applied to these univariate time-series representations is only applicable to the information contained within them.

In Chapter 3 an investigation via a complex network method was undertaken, which implied that transitions in the time-varying behaviour of the international oil market are dominated by the increase or decrease of a few causal links and not a substantial change in the structure of the causality patterns. Based on this prior work, this chapter investigates a hypothesis: "can a meaningful one-dimensional time series representation of a time-varying system be constructed based upon that system's overall transition behaviour?". The use of a dimensionality reduction method based on a scientific understanding of the system opposed to a standard statistical method, such as PCA, is three fold. PCA has three important limitations in its application, which are that it assumes linearity and correlation of the causality results between the time series (of each of the oil returns), that it is not robust to outliers, and that the principle components (especially when reduced to one-dimension) have low interpretability [165]. All three of these problems can be addressed with an approach based on the scientific understanding of the system in question, with non-linear relations being able to be considered, outliers having a more limited effect, and the dimensionally reduced value having a pre-determined meaning.

4.2 Methodology: Paradigms of dynamic complex networks

Following the above background and the work in Chapter 3 this section defines two paradigms based on the transitions of the causality pattern that a system could be in (these are not intend to be exclusive to all temporal dynamics, merely to capture two dominate types of temporal dynamics based on the USIC-Network presented in Chapter 3). For each of these paradigms, a one-dimensional representation of the data is given that is deemed appropriate for the specific behaviour of the system. A methodology for determining which paradigm is most applicable for a system

is also presented. This paradigm approach splits the complexity of dimensionality reduction into two aspects, choice of paradigm and lower dimensionality representation in that paradigm. This allows the actual dimensionality reduction approach to be simpler and more meaningful to the specific system. These two paradigms are described as follows:

1. Change paradigm: this paradigm is best described by the β metric (introduced in Chapter 3), the transitions of a dynamic system under this paradigm will be limited to the addition or removal of causality links in a manner that maintains similarity between α and β. In other words, transitions under this paradigm will primarily change the overall causality of the causality pattern. For clarity examples are shown in Fig. 4.1.

2. Ordering paradigm: this paradigm is best described by the α metric (introduced in Chapter 3), the transitions of a dynamic system under this paradigm will primarily be rearranging the structure of the causality patterns, leading to a dissonance between α and β. In other words, transitions under this paradigm will primarily maintain the overall causality of the causality pattern. For clarity examples are shown in Fig. 4.1.

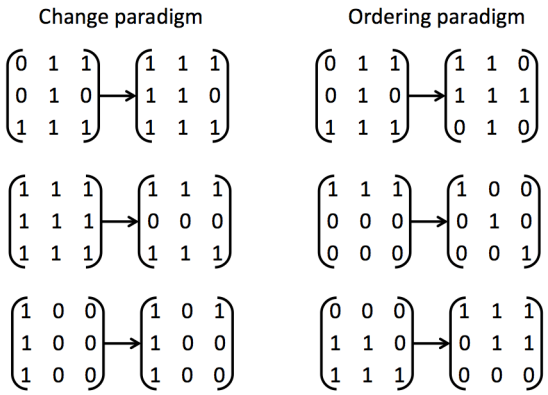

Figure 4.1: *Examples of causality patterns before and after a single transition, for a system containing three variables. Change paradigm examples shown on left exhibiting either addition or subtraction of causal links. Ordering paradigm examples shown on the right exhibiting combination of addition and subtraction of causal links.*

These paradigms encapsulate two behaviours of the time-varying dynamics of a system, these exist as a subset of a much large set of all possible behaviours, and are not intend to be an exhaustive set. These two paradigms were specifically chosen for where they sit in this larger set of all behaviours. For complex analysis (such as that proposed in Chapter 3) it is desirable and common to have a number of archetypal forms presented that represent the edges of behaviours (see small world and scale-free networks in network theory [166]). The two paradigms presented can be seen to represent two opposite types of evolution (that preserving total causality but changing link layout, and that preserving link layout but changing total causality) and sitting at either end of a spectrum of paradigms, with paradigms mixing these two evolutionary types occupying the space in the spectrum between them.

With these paradigms defined, an educated decision on what metric to employ for the one-dimensional representation of the causality pattern can be made. This choice of metric should be selected to give the best visibility to the behaviour of the underlying structure. These choices are as follows:

1. Change paradigm: in this paradigm, movement is primarily driven by the β metric, implying that transitions result in a change in the total causality of the system, i.e. the new causality pattern will have a different amount of causality than the previous causality pattern. Using this information it can be seen that the total causality of the system will change time step to time step (ignoring self-loops), allowing a time series of this measure to encode the information change of each transition. Therefore this metric is chosen as the one-dimensional representation of the causality pattern for this paradigm. For convenience, this metric will be referred to as Sum of Causality (SoC) and formally defined for a pattern i as $SoC = \sum C_i$ (where C_i is a causality pattern, defined in Chapter 3.2).

2. Ordering paradigm: for this paradigm movement is primarily driven by the α metric, implying that transitions result in a change in the structure of the causality pattern but an equivalent total causality. Due to equivalent total causality, SoC is not applicable as a metric for this paradigm. This reveals

that a transition should result in a new causality pattern with a different structure to the previous causality pattern (ignoring self-loops), i.e. resulting in an α metric value greater than zero between them. A reference causality pattern needs to be chosen to convert this form of transitional metric to a causality pattern metric. A reference causality pattern C_r can either be manually selected or taken as the mean causality pattern of some historical data, with the latter being recommended. Using this reference pattern, a metric for the causality patterns under this paradigm is proposed. For convenience, this metric will be referred to as Sum Difference from Mean (SDM) and is the value of the α metric between the causality pattern and the reference causality pattern. For a causality pattern i this can be formally defined as $SDM = \sum \sqrt{(C_i - C_r)^2}$

The determination of a systems current paradigm naturally follows on from the definitions and discussions given above. Given that a system is represented by the USIC-Network proposed in Chapter 3, a comparison of the α and β transition parameters for the network will allow for a determination of the paradigm. If for the majority of transitions (edges) $\alpha = \pm\beta$ the system can be said to be in a Change paradigm, and if for the majority of transitions $\alpha > |\beta|$ and $\beta \approx 0$ the system can be said to be in an Ordering paradigm. However, these determinations are qualitative, and hence the specific conditions and level to which they hold will depend upon the data and analysis being undertaken.

4.3 Paradigm results: International oil market

This chapter was motivated by the results in Chapter 3 which imply that the time-varying behaviour of the international oil market data, laid out in Chapter 2.4, is dominated by transitions changing a few links and changing the overall causality of the causality patter. Following the proposals of paradigms set out above, these findings suggest that the international oil market is dynamic behaviour is classifiable as the Change paradigm. To investigate this hypothesis qualitative determination approach set out above is followed below:

Taking the data presented in Chapter 2.4 and apply the method for complex

network construction proposed in Chapter 3, a USIC-Network representation of the data can be found. From this USIC-Network the α and β values for each edge can be extracted, these weightings are shown in Fig. 4.2. The comparison of α

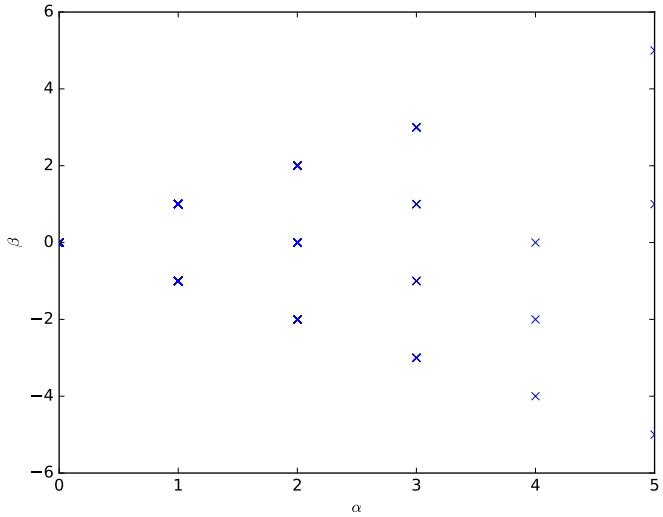

Figure 4.2: *α and β weightings for edges calculated in Chapter 3 for all edges within the network.*

against β shown in Fig. 4.2, do not appear to strongly support either paradigm, with data points occurring for $\alpha = \pm\beta$ and, $\alpha > |\beta|$ both for $\beta = 0$ and $\beta \neq 0$. Based on this a further investigation can be undertaken, noting that the interest is on the predominate behaviour, "outlier" and irrelevant edges (and their associated α and β) can be removed. To achieve this all edges with a frequency weighting of less then or equal to five ($F \leq 5$) are removed, and so are all self-loop edges (where the source and end node are the same). After this data cleaning the data can be re-plotted, with results shown in Fig. 4.3.

From the comparison of α against β shown in Fig. 4.3 it is clear that $\alpha = \pm\beta$ for all investigated edges. With this knowledge, a conclusion can be made that the international oil market is under the Change paradigm, concurring and furthering the results from Chapter 3.

Following the methodology set out above, this analysis implies that the time series data for the international oil market would be well represented by the Sum of Causality (SoC) metric. To achieve this the series of causality patterns discussed in

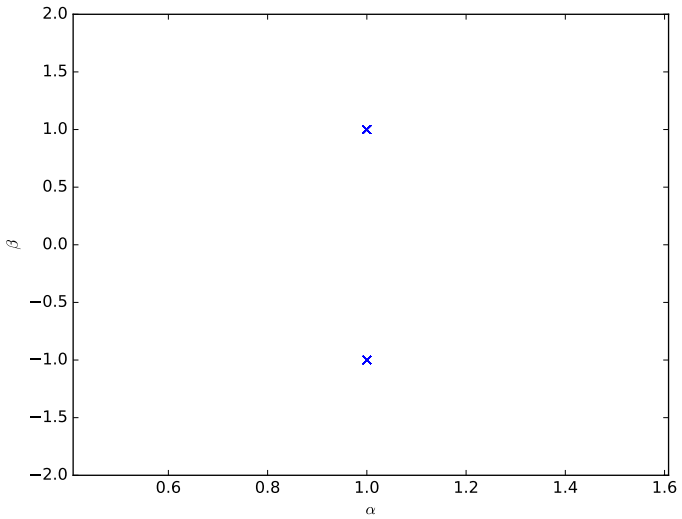

Figure 4.3: *α and β weightings for edges calculated in Chapter 3 for all edges within the network that are not self-loops and have $F > 5$.*

Chapter 2.4 can be transformed into a time series of one-dimensional values, with each value being the SoC for the causality pattern at that time step. This time series is plotted and shown in Fig. 4.4.

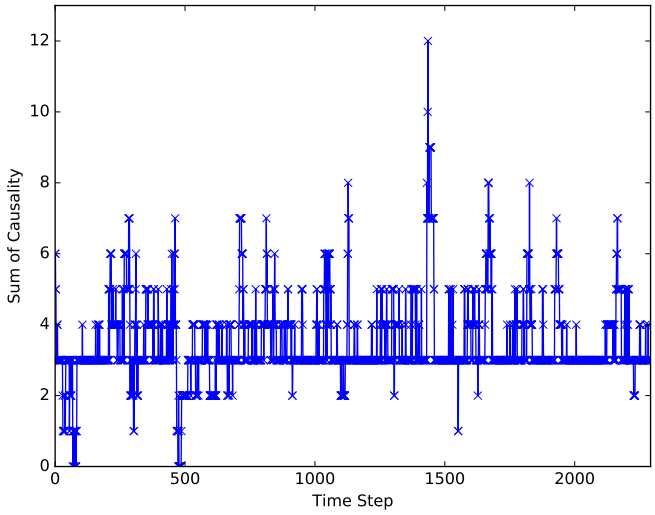

Figure 4.4: *Time series of Sum of Causality (SoC) for the series of causality patterns for the international oil market discussed in Chapter 2.4.*

4.4 Methodology: Predicting causality patterns under a paradigm

Following the results above for the international oil market, this section constructs a methodology for predicting this data, and hence validating the paradigms description. An investigation into this time series data is undertaken to achieve this, with methods being proposed where appropriate. As such, the results of this section can be seen as specific to the international oil market; however, the approach to the construction of these methods can be seen as a framework for other investigations. This section is split into three parts: a methodology for smoothing the time series of causality patterns, an investigation of this new time series for predictability, and a proposal of a simplistic prediction algorithm to demonstrate the strength of the information contained in the series.

4.4.1 Time series smoothing

From observing the behaviour shown in Fig. 4.4 it can be seen that there appears to be a dominant value to which the system reverts. The deviations from this value appear in some instances to have a wind up/down behaviour, taking a few steps to peak or come down. This implies that some behaviour may exist within this time series; however, due to the step-like nature of the data, this behaviour is challenging to observe. Therefore it may be informative for a smoother expression of this data to be investigated.

Due to the relatively short number of time steps that appear to be involved in the wind up/down behaviour, methods such as smoothing over multiple time steps cannot be employed [167]. It is known from the formation of the Granger causality test that a artificial quantisation is applied to the actual statistics, this is the significance level that is applied to the continuous *p-value*. Therefore, a measure called continuous Granger causality is proposed to investigate this system, based on the raw *p-value* result as a representation of the strength of the causality [168]. This measure is for use in intermediary analysis and should not be considered a definitive measure of causality its self, for that, a significance level should be applied as in

the traditional Granger causality test. For clarity the continuous Granger causality measure is defined as $1 - p\text{-}value$ (where the *p-value* is calculated as discussed in Chapter 2), such that a "strong" causality is near 1 and a "weak" causality is near 0. This usage of the *p-value* is inspired by the discussions on the topic of this quantity by R. Fisher [169], who describes the *p-value* as a rough numerical guide of the strength of evidence against the null hypothesis [170, 171]. This endorses the use of the continuous Granger causality metric as a rough guide to the system's dynamic behaviour. The aim of this metric being to give a smoother nature to the dynamics allowing for behaviours to be correctly characterised and discovered. It should also be noted that other interaction dynamics with a continuous measure of interaction may offer alternatives to this approach, such as transfer of entropy [37] and correlation [115], discussed in Chapter 2.

With this continuous Granger causality metric, the causality patterns can be reconstructed, and hence so can the time series, as described above. For clarity from here forward, the time series produced with the continuous Granger causality metric will be referred to as the "continuous time series", and the original series shown in Fig. 4.4 will be referred to as the "binary time series" (after the binary nature of classic Granger causality). The continuous time series can be seen in Fig. 4.5.

The data shown in Fig. 4.5 appears to exhibits a much smoother behaviour than that in Fig. 4.4, and to display oscillation around some value. Compared to the data in Fig. 4.4 the SoC values produced are larger; this is due to the summation of values that a significance filter would set to zero. In line with the earlier hypothesis, this data exhibits more information on the wind up/down behaviour in the time-varying behaviour, with more obvious movement both two and from the midpoint being observable.

The variation of this data set can be investigated further by plotting the histogram of its distribution; this is shown in Fig. 4.6. From this distribution, one can see that it appears approximately Gaussian, with a very often visited mean and decreasing visits to the extremities, creating tails. With this knowledge investigating

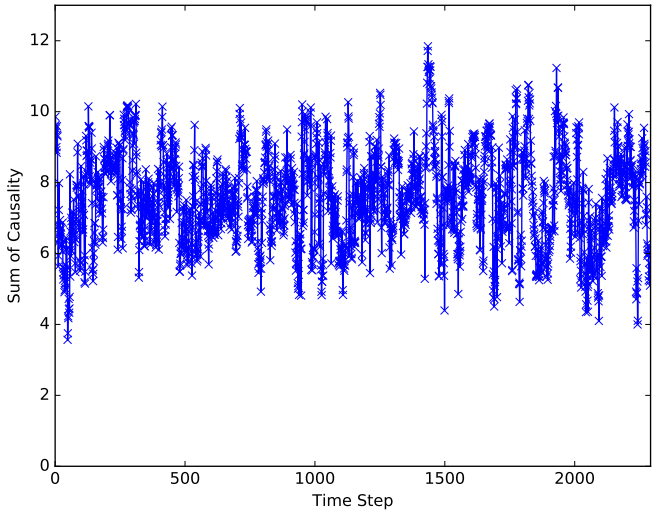

Figure 4.5: *Sum of Causality for each time step, where the measure of causality in each pattern is calculated as* $1 - p$*-value (the continuous time series).*

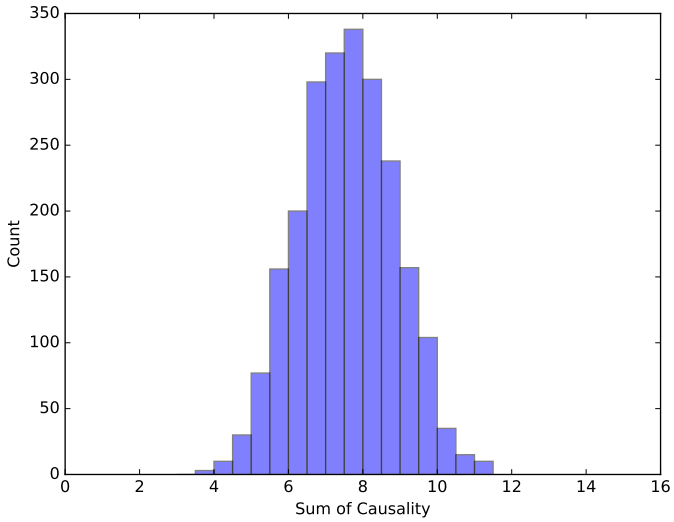

Figure 4.6: *Histogram of Sum of Causality for data shown in Fig. 4.5, bin size of* 0.5.

the data from Fig. 4.5 with a line of best fit should return how well this data is distributed around this central value. A line of best fit, calculated using least squares, is shown in Fig. 4.7 with a gradient of 2.354×10^{-5} and an intercept of 7.553. This line of best fit implies that the data is approximately centred on this intercept and oscillates around it. A further test can be done to test if this data is stationary, here the Augmented Dicker-Fuller (ADF) test was employed (results shown in Table 4.1),

that shows this data is stationary at a 1% significance level.

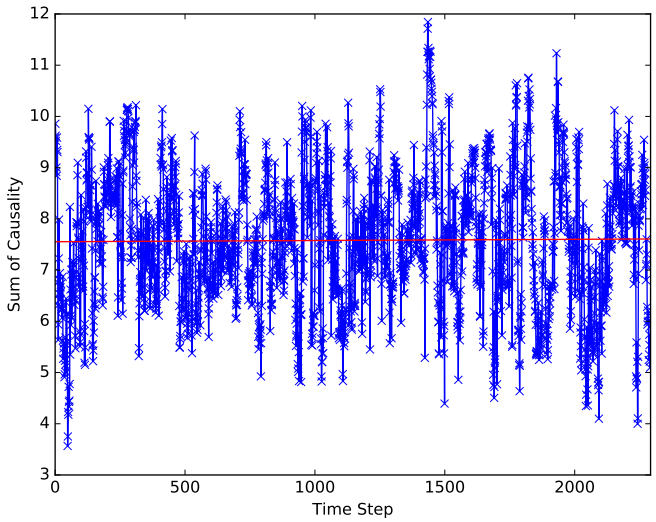

Figure 4.7: *Sum of Causality for each time step, where the measure of causality in each pattern is calculated as* $1 - p$*-value (the continuous time series). Line of best fit with gradient of* 2.354×10^{-5} *and a intersect of* 7.553.

Table 4.1: Results of stationary tests using a Augmented Dickey-Fuller (ADF) test. A p-value<0.01 indicates the rejection of the null hypothesis for the test at a 1% level.

	ADF-Statistic (3 sf)	p-value
Sum of Causality	-10.7	0.001

From this analysis, it appears that the usage of a continuous Granger causality measure allows for the expression of time-varying behaviour that the application of a significance level can obscure. Further, the continuous time series appears more amenable to an investigation of predictability than the binary time series due to its stable, smoother, and stronger behaviour.

4.4.2 Investigation of predictability

Taking the continuous time series produced above an investigation of its amenability to prediction is undertaken. This investigation can validate the usability and suitability of the paradigms for expressing key temporal properties of the system.

The above transfer to a continuous time series from a binary one was motivated by the presence of wind up/down periods in the series, where it increased to and

decreased from high SoC values (or the reverse for lower values). In the continuous time series, this behaviour is more apparent and can be investigated in more depth as a rate of change. Predictability within the rate of change can be expanded upon to determine predictability within the series itself. Here the rate of change from the line of best fit (shown in Fig. 4.7) is conducted by: (i) calculating the deviation of the value at each time step from the line of best fit, (ii) calculating the rate of change of each time step, as the difference in the value of the current time step from the last (thus these calculation create a list of values that ignores the first time step). The behaviour of these values can be explored by plotting the deviation from the line of best fit for each time step against its rate of change, these results are shown in Fig. 4.8. The results shown in Fig. 4.8 demonstrate an overall behaviour for the

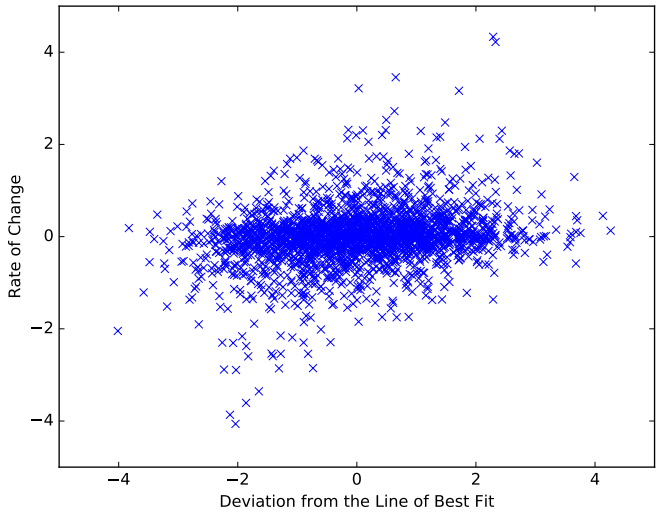

Figure 4.8: *Rate of change against deviation from the line of best fit for the data in Fig. 4.5.*

systems movement, with strong clustering around the 0 rate of change, spreading and increasing as the deviation approaches ± 2, representing the systems reversion to the line of best fit. This implies that the dynamic nature of this behaviour is approximately linear between peaks, where large rates of change occur.

The next investigation into the predictability of the data is through the lags. Taking lag time steps of 1, 5, 20, and 50, the comparison between values at these different intervals is taken. This investigation is aimed at determining how well a

current value can represent a future value. The results of this investigation are shown in Fig. 4.9. From this investigation it can be seen that for low lag values there is

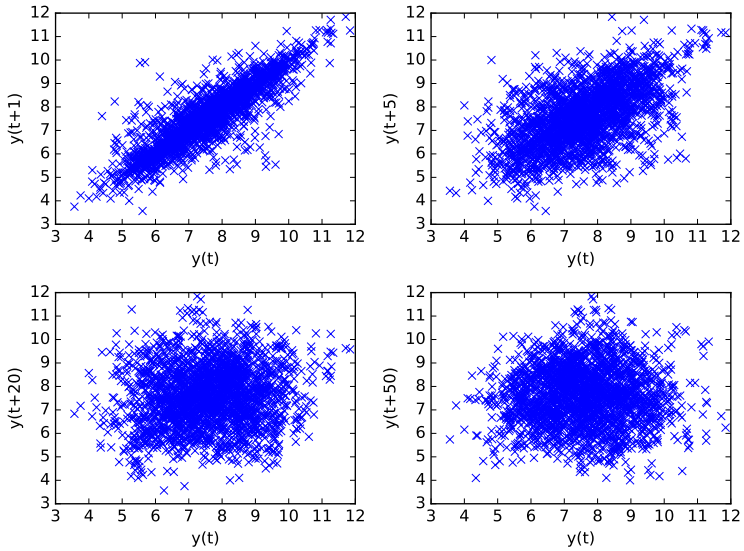

Figure 4.9: *Lag plot, with $y(t)$ being the SoC value at time step t from Fig. 4.5.*

some positive correlation, however this quickly drops off until there appears to be no correlation. Implying that the current value may operate as a potential predictor for a low number of time steps into the future.

This can be investigated further by looking at the autocorrelation of the continuous time series, with the results of this analysis shown in Fig. 4.10. From the results in Fig. 4.10, it can be seen that for low lag values there exists rapidly decreasing strong positive correlation, for high lag values there does not appear to be much significant correlation. This corroborates the results shown in Fig. 4.9, suggesting some predictability for a low lag order.

For clarity the results in Fig. 4.10 were re-plotted for low lag values, shown in Fig. 4.11. From these results, the predictive power quickly degrades as the lag increases, becoming non-significant at around a lag of 25 time steps, equivalent to a period of 55 days.

4.4.3 Prediction algorithm

Following the above analysis, a prediction algorithm is proposed for this data. This algorithm is designed to take advantage of the discovered auto predictability at low

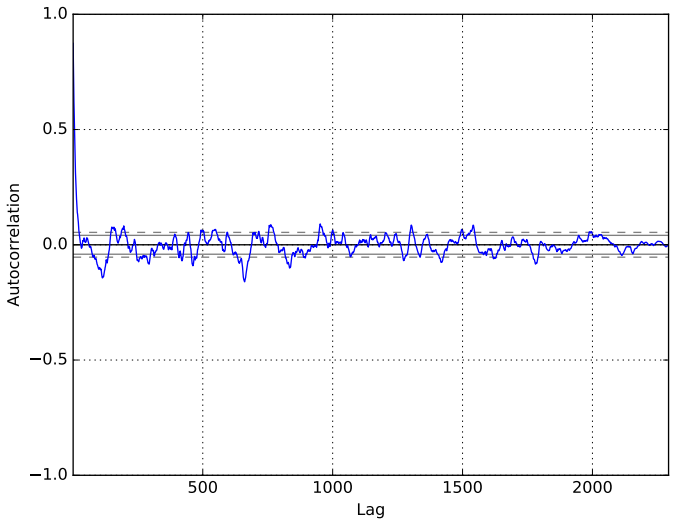

Figure 4.10: *Autocorrelation of data from Fig. 4.5, with the horizontal dotted lines displaying the 99% confidence band and the solid horizontal line displaying the 95% confidence band.*

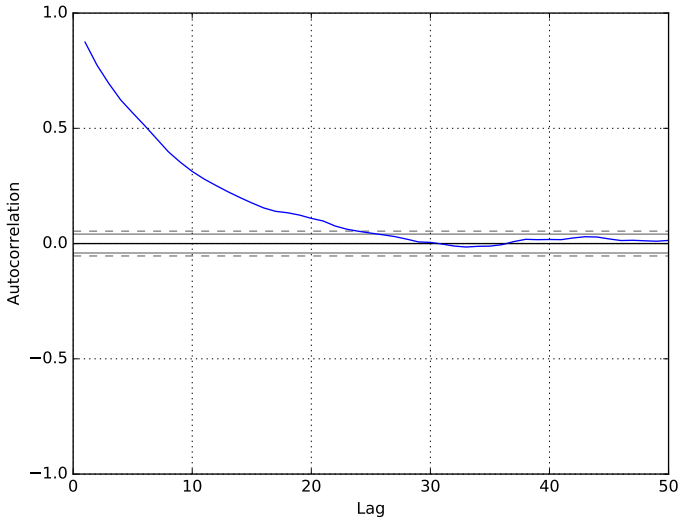

Figure 4.11: *Autocorrelation of data from Fig. 4.10, with the horizontal dotted lines displaying the 99% confidence band and the solid horizontal line displaying the 95% confidence band.*

lag values and the approximate linearity between peaks. This algorithm is intentionally simplistic to show the power of the paradigm at expressing the nature of the system, and not the power of a complex algorithm.

Based on this, a prediction algorithm is proposed called Rolling Linear Fit

(RLF) prediction. This algorithm can be broken into four steps, which are as follows:

1. Training: this method the previous x time steps as training data. This training data is then used to fit a linear model via least-squares [172, 173].

2. Prediction: this linear model is then used to predict the SoC values of the next y time steps by predicting that it falls along the constructed line.

3. Conversion to causality pattern: the predicted SoC values are compared to the SoC of all historic causality patterns, with the causality pattern that has the closest SoC value to the predicted SoC value being predicted as its causality pattern.

4. Retraining: after prediction is complete, the algorithm retrains on the new x previous time steps, producing a new linear model via least squares. This process is then repeated until no new data is being acquired.

As such, this algorithm can be seen to be parametrised on two user-defined quantities, the training window and the prediction length.

4.5 Prediction results: International oil market

Follow the methodology laid out above, an investigation of this prediction algorithm can now be undertaken on the international oil market, as described in Chapter 2.4.

Due to the above-stated training and prediction length parameters, the first step in the investigation is to determine these. To achieve this, several parameter combinations are investigated, based on the autocorrelation results shown in Fig. 4.11. Six training lengths are selected for testing: 5, 10, 15, 25, and 30 data points, and four prediction lengths: 5, 10, 15, and 20 time steps. To understand the accuracy of each combination of these parameters predictions were made over the full time series with each combination. To achieve this the time series was split into windows of training length + prediction length. Each predicted causality pattern (binary) was compared to the true causality pattern at that time step via the Hamming distance

between them [174]. This comparison results in a series of accuracy measurements; taking the mean, a single value of the accuracy can be found (mean error). The results of this experiment are shown in Fig. 4.12.

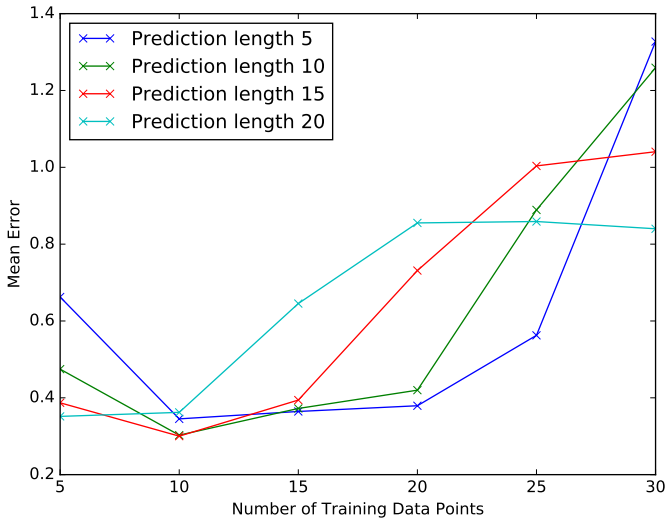

Figure 4.12: *The mean error of Rolling Linear Fit prediction for a number of training sizes and prediction lengths.*

From the results shown in Fig. 4.12 it can be seen that for the majority of the prediction lengths, a training range of 10-20 data points is optimal and that a prediction range of 5-10 maintains the best results. An explanation for this behaviour is, if the training + prediction lengths are too large the fluctuation in the data (shown in Fig. 4.5) is overshot, and a linear model is no longer appropriate. Based on these results, a train length of 10 data points and a prediction range of 10 time steps is chosen for the experiment going forward.

To understand this algorithm in context some benchmarking must be applied. To achieve this, three simple alternative approaches for prediction are introduced below:

- Line of Best Fit (LBF) prediction: This method assumes that the line of best fit for the entire data set can be calculated as is shown in Fig. 4.7, and predicts the associated SoC value for the relevant time step. This value is then compared to find the continuous causality pattern within the series whose sum is

closest to it. This pattern then has a significance level of 0.05 applied to produce a binary causality pattern that is compared via the Hamming distance to the actual binary causality pattern at that time step to calculate an error.

- Static Evolution (SE) prediction: This method assumes that the system will not change from the current causality pattern and will hence always predict the current causality pattern. To calculate the error, the current causality pattern is compared to the causality pattern at the predicted time step via the Hamming distance.

- Network Based Transition (NBT) prediction: This method uses a network prediction approach discussed by Jiang *et al.* [3]. For the current node in a network, the next node is predicted as the one most often transitioned to from the current. Therefore the current causality patterns node in complex network representation (discussed in Chapter 3) is discovered, and the edge from it with the highest frequency F is selected with the causality pattern relating to the node it connects to being chosen as the prediction. The error for this is calculated as the Hamming distance between this and the causality pattern at the predicted time step.

An experiment investigating the mean accuracy (as described above) for several different prediction lag lengths was conducted to compare the Rolling Linear Fit prediction to these alternatives. This experiment was designed to compare the result of the methods more accurately, considering the difference in their prediction ranges. This experiment for the Rolling Linear Fit prediction can be described as follows: the first 10 time steps were taken for training, with the next 10 being predicted. The accuracy of the prediction for each lag (1-10) was noted separately. This process was then rolled forward by one time step and repeated until the end of the data was reached. This resulted in a series of accuracy measurements for each lag value, the mean accuracy for each of these values was found. For the alternative three approaches, the prediction was compared to time steps 1-10 ahead of the current step. This process was then rolled forward by one time step and repeated,

the results for these predictions was then averaged for each lag. The comparison of these results can be seen in Fig. 4.13.

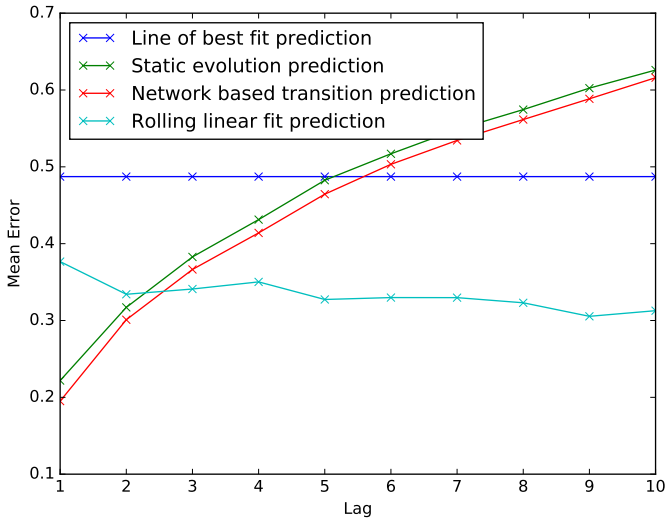

Figure 4.13: *Mean error for each of the presented methods of prediction for lags between 1-10.*

The significance of the results shown in Fig. 4.13 was compared via a t-test, with results shown in Table. 4.2. These results demonstrate that for a lag of 1 time step, RLF is statistically better than LBF, but is outperformed by both SE and NBT. This result is somewhat expected as, has been previously implied (Chapter 3) the system tends to remain in the same or a similar position over a single time step, and given the occurrence of changes in the rate of change, there are expected to be steps where SE and NBT outperform the RLF over a single time step. For a lag of 2 and 3 RLF outperforms LBF, but is not statistically different from SE and NBT. This demonstrates that at only a lag of 2 the performance of SE and NBT have degraded enough that the change in rate of change no longer offers enough of an impact on RLF for out-performance. For lags 4-10 RLF statically outperforms all presented alternatives. Notably, RLF can be seen as relatively stable in its prediction throughout the whole prediction range, this is in contrast to SE and NBT whose results continuously degrade.

Table 4.2: Results of t-tests between the presented methods and each lag level. A p-value$<$0.01 indicates the rejection of the null hypothesis for the test at a 1% level.

	Statistic/p-value (3 dp)		
Lag	RLF & LBF	RLF & SE	RLF & NBT
1	-5.772/0.001	9.587/0.001	11.496/0.001
2	-8.140/0.001	1.018/0.309	2.000/0.046
3	-7.747/0.001	-2.400/0.0164	-1.472/0.141
4	-7.214/0.001	-4.509/0.001	-3.571/0.001
5	-8.530/0.001	-8.560/0.001	-7.609/0.001
6	-8.374/0.001	-10.155/0.001	-9.446/0.001
7	-8.353/0.001	-11.725/0.001	-10.981/0.001
8	-8.761/0.001	-13.417/0.001	-12.739/0.001
9	-9.739/0.001	-15.804/0.001	-15.106/0.001
10	-9.311/0.001	-16.494/0.001	-15.999/0.001

4.6 Discussion

In this chapter several methodologies, investigations, and experiments, were conducted. Demonstrating that the usage of the proposed paradigms and continuous Granger causality can lead to better predictions of the causality patterns then the previously proposed Network Based Transition (NBT) approach from complex networks. This investigation can is summarised as follows:

A methodology was proposed to encapsulate two forms of dynamic behaviour based on the underlying causality pattern of a system: Change paradigm and Ordering paradigm. For both of these paradigms, a metric was proposed that would be appropriate for the one-dimensional representation of a causality pattern for a system under the paradigm: Sum of Causality (SoC) and Sum Difference from Mean (SDM). Following this, an investigation of the international oil market concluded that this data follows the Change paradigm, corroborating the results from Chapter 4; this allowed for the expression of the time-varying data as a time series.

Based on this, a methodology was presented for the prediction of data in this series. First, a hypothesis was constructed, and an experiment was conducted to demonstrate that the dynamic behaviour could be more fully explored by taking a continuous measure of causality. Following this a measure called continuous Granger causality ($1-p\text{-}value$) was proposed and the time series was transformed

into a "continuous time series". An investigation of predictability was conducted on this time series, demonstrating that for a low lag number, the series exhibits some autocorrelation characteristics. Based upon these findings, a prediction algorithm called Rolling Linear Fit (RLF) prediction was proposed. This method constructs a linear prediction for short lag values and then retrains this model, rolling it along the data.

Using this algorithm, an experiment was run on the international oil market. This experiment concluded that the optimal training and prediction range were both 10 time steps. Furthermore, three alternative prediction methods were also proposed and it was discovered that for lag ranges of 4 to 10 the RLF statistically outperformed the presented alternative methods for the prediction of future causality patterns. These results imply that RLF can predict the causality pattern of the international oil market 10 days ahead with, on average a deviation of less than one causal link (with a mean error of 0.316).

This chapter demonstrates the usefulness of a univariate time series representation of a multivariate system for the analysis of dynamic behaviour, in a manner not capture by, but complementary to, the complex network approach (shown in Chapter 3). This chapter also showcases methodology for the informative selection of dimension reduction to capture known aspects of the dynamic behaviour.

Chapter 5

Configuration Space Construction Methodology for Behavioural Investigations

The analysis of the international oil market presented in Chapter 3 and Chapter 4 found results that imply the presences of a overall behaviour in the evolution of the system causality structure. From this analysis a hypothesis of a potential behaviour described at a low level can be formed; clustering of similar causality patterns, connected through repeated transitions (discussed in detail in this chapter). Though this hypothesis is formed from validated investigations, its validity as a higher order description of the system needs also to be validated.

This analysis has so far been built on construction of complex networks, these offer a powerful tool for the investigation of a system but may be limited in the full description of the behaviour of that system. Therefore, to explore the hypothesised behaviour this chapter explores the usage of a configuration space as a framework for constructing a description of the systems dynamic behaviour for investigation, allowing the hypothesis to be tested. This is achieved via the proposal of a methodology to encode the evolution of the system in a manner that maintains the knowledge of the underlying structure, with this methodology a formal description of the hypothesis can be constructed. This is validated and the international oil market is analysed via the creation of a algorithmic analysis and prediction method.

This chapter hence presents an approach to expressing and validating high order descriptions of a systems behaviour that have been drawn from the previously shown analysis methodologies.

5.1 Background and related work

This chapter is specifically interested in representative methods that have a high level of information encoded relating to the original time series. The limitations of standard network transformation approaches and time series analysis for the capture of the full information regarding the time series evolution has been noted by a number of authors [21, 175, 66, 38]. Transforming to these representations often causes some level of obscuring of information contained within the original series. A response to this, that has achieved success, has been through the usage of state space representations for the behaviour of a dynamic time series [33]. In particular phase space representations have been employed, often as an intermediary stage to the construction of a complex network. In this thesis a phase space will not be employed, however within the literature it has been, therefore below an overview of its usage is given for completeness.

A phase space operates as a multi-dimensional space in which the time series can be encoded (discussed further on), the benefit of this approach is that the space allows for distance between points to be measured (this also applies where points are represented as vectors from the origin). This phase space additionally supports the construction of a complex network whose nodes correspond to selected points within the phase space and whose edges selected pairs of points in the phase space. An example selection of edges might be all those whose points in phase space have a distance less then some defined threshold. In prior work a transformation of a time series into a phase space has been achieved via time delay embedding, with Taken's embedding theorem often cited as the motivation [176]. A phase space representation such as this is based upon the concept of subdividing a single time series into a number of disjoint time series, such that each of these disjointed time series can be expressed by a single vector in the phase space. For this purpose a

phase space for this is constructed such that each dimension represents a "delay" i.e. a number of time steps from the current time step. Therefore a point in an m dimensional phase space such as this represents a time series of m steps, that can be seen as some subdividing of an original time series. Following the work of Gao and Jin [21] a vector in this phase space, for a time series z_t, where $t = 1,2,...,M$, can be expressed as follows:

$$\begin{aligned}\vec{X_k} &= \{x_k(1), x_k(2), ..., x_k(m)\} \\ &= \{z_k, z_{k+\tau}, ..., z_{k+(m-1)\tau}\}\end{aligned}$$

Where $k = 1,2,...,N$, $N = M-(m-1)\times\tau/t$, τ is a selected time delay, and m is the dimension of the space. This and similar approaches have been used to great effect by other a number of authors [66, 21, 50, 175, 92], and can be seen as defining a space that is able to specify the state of the system. However, similar to much of the literature in the field, the concern of this approach is on the dynamics of time series, and not on the structure of the underlying causality generating the series (in the case where the series is one of causality patterns).

For the methodology and investigations presented in this thesis a different representation is required. A state space approach that is more suited to the representation of dynamic causality in a multivariate system, is a configuration space. Where each point within this space encodes a causality pattern, and any causality pattern can be represented. This state space is selected for usage in this chapter, and is discussed further in the next section.

This chapter also explores groups of points in the above mentioned state space, and their detection. Detection of groups of points in a state space is achieved through clustering approaches, such as k-nearest neighbours, shared nearest neighbours, single linkage, and other density or linkage based methods. However, these classic approaches aim to define clusters, instead of finding pre-defined clustering behaviour. In many systems clusters will often be the same, using a reduction in the density of points in the space to determine the edge of a cluster, but for the specific behaviour being investigated in this chapter this separation approach is not

immediately applicable. Therefore, in this chapter an advancement to the clustering literature is presented, to separate out groupings of points that are connected via dense pathways. This work and its background is discussed further in Section 5.4.

5.2 Methodology: Configuration space

Following the above background a configuration space representation for time varying causal interactions in a multivariate system of time series is proposed. A configuration space consists of a space, which contains all possible configurations (states) of the system, allowing for any state to be expressed within it [177]. In the context of a causality pattern, this means a configuration space contains every possible causal value for every possible causal link, hence any causality pattern can be expressed as a single point within the configuration space. This will allow for a space in which the trajectory of a time series can be encoded, maintaining full information regarding the underlying causality and the geometric relationship between them.

A configuration space (an N-dimensional Euclidean space) is created in which all possible states (causality patterns) of a system are represented, with each possible state corresponding to a unique point in the configuration space. In this configuration space it is possible to plot the trajectory of a series of causality patterns. This is done by projecting each causality pattern onto the configuration space (i.e. plotting them as points in the space), and tagging each point with its time stamp (the time stamp is not a dimension of the space).

This configuration space is defined with a separate axis to represent each possible causal link, with these axes having a range from 0 to 1, in line with the continuous Granger causality measure defined in Chapter 4. Thus, for example, any point in a configuration space with four dimensions could be represented by a point of four coordinate elements, where each element can have values ranging from 0 to 1. The dimensionality of the configuration space will be directly linked to the number of variables. For a bivariate system with variables X and Y, the axes would be: $X \to Y$, $Y \to X$, $X \to X$, and $Y \to Y$. This gives a configuration space with four dimensions, which can be generalised to N^2 dimensions for N variables. This order

can be reduced by ignoring self causality (e.g. $X \rightarrow X$), which is not of interest in this investigation, and with this restriction the order of the configuration space can be collapsed to $N^2 - N$.

Therefore the systems state of a multivariate time series can be expressed in this representation as follows. For an arbitrary multivariate system of N time series, each of length T, the system state at time t (the causality pattern of the multivariate system at time t) can be expressed as:

$$s_{t,n} = \{c_{i,j}\}_t \forall i \neq j | i, j \in \{1,2,...,N\} \quad t \in \{1,2,...,T\}$$

Where $s_{t,n}$ is the system state (point in the configuration space) at time t for a space of dimension n, $n = N^2 - N$, and $c_{i,j}$ is the continuous Granger causality from variable (time series) i to j (the measure introduced in Chapter 4).

5.3 Constructing behaviour description

The configuration space representation discussed above, allows for a framework within which descriptions of dynamic behaviours can be constructed. Using this framework a description of the behaviour for the sample data from the international oil market can be constructed and formally defined, where this behaviour hypothesis is constructed based upon the results found in Chapter 3 and 4.

Based on the results found in the previous chapters two aspects of behaviour are hypothesised: (i) Chapter 3 demonstrates the existence of noisy clusters that the system state moves between, often staying within a cluster for a number of time steps, it also implies a cyclic behaviour where the system state moves between more stable states (high chance of self-looping) via less stable ones. Chapter 4 along with Chapter 3 demonstrate that the transitions between system states is dominated by small changes in the causality patterns with groupings designated by causality patterns close in α. This behaviour underlies a hypothesis that there exists regimes within the time-varying behaviour, where causality patterns that have a small difference in α are clustered and that the systems time-varying behaviour is dominated by these. (ii) Chapter 4 shows that the transitions throughout the time-series of

the causality patterns has a degree of smoothness to it, with a sharp but incremental build up and down between different levels of sum of causality, and Chapter 3 shows the incremental cyclic behaviour between the stability of causality patterns. This behaviour underlies a hypothesis that there exists pathways of defining the transitions between the above mentioned regimes, with these pathways being composed of a number of system states over consecutive time steps. Taken together this hypothesised behaviour can be described as causal regime shifting, as the system moves between regimes of causality, both of these hypothesis are more formally explored below.

Described in terms of binary causality, a causal regime shifting system (also called a "regime changing" system) is a system that repeatedly transitions between a fixed number of "stable" binary causality patterns, where a "stable" binary causality pattern is described as one that the system state remains in for a defined minimum number of time steps. An example of a similar network behaviour to this can be seen in the work of Jiang *et al* [3] where it is shown that for four international crude oil market benchmark prices there exists a number of well visited causality patterns between which the system state predominately transitions.

The previous mentioned approach of Jiang et al [3] uses binary causality patterns as nodes in a complex network (with edges representing the transitions through time), as discussed in Chapter 2. To make use of richer information, in this chapter the continuous measure of Granger causality is used, as introduced in Chapter 4, to allow for discussion of small variations and gradual changes within the space.

It is assumed that observed system states (causality patterns) are affected by noise and errors in detection and measurement, and therefore the observed states are approximations of the "true" state of the system. For simplicity it is assumed that such deviations have an identical and independent Gaussian probability density function across all dimensions, therefore repeated visits to the same "true" system state will gradually produce a symmetrical cluster within the configuration space.

The dynamic nature of the system will is a time ordered path through the configuration space. This is defined as an ordered set S_n of observations $s_{t,n}$ in the n-

dimensional configuration space, where each observed point occurs at a consecutive time step t (one observation per time step) and is a tuple of n real-valued continuous Granger causality values $c_{i,j}$ (from variable i to variable j, with this measure introduced in Chapter 4):

$$s_{t,n} = \{c_{i,j}\}_t \forall i \neq j \mid c_{i,j} \in \mathbb{R},\ t \in \mathbb{N}^*$$

$$S_n = (\{s_{1,n}, s_{2,n}, \ldots s_{T,n}\}, \prec) \mid \forall s_{a,n}, s_{b,n} \in S_n \quad s_{a,n} \prec s_{b,n} \iff a < b$$
$$\wedge\ s_{i,n} \in S_n \iff s_{i,n} \in \{s_{1,n}, s_{2,n}, \ldots s_{T,n}\}$$

It is assumed that for regime shifting behaviour the "true" system state will take one of two positions: (i) being in a "stable" state, analogous to a node in a binary network representation, where the state will remain for a number of time steps, or (ii) being in a "moving" state, analogous to an edge in a binary network representation (but where the edges consist of sequences of observed intermediate states, which are not normally captured in a binary network representation), where the system state will transition between "stable" states. This thesis refers to the noise distribution centred on a "stable" state as a Point Cluster, and the sequence of noise distributions centred on intermediate states in a "moving" state as a Trajectory Cluster.

These Point and Trajectory clusters can be thought of using the analogy of K-nearest neighbours and Markov-Chains. A Point cluster is made up of a group of nearest neighbours, and Point Clusters are connected via Trajectory clusters each of which is a path of points, each a nearest neighbour to the last. The temporal pathway describing how the system moves from one neighbour to another can be though of as a Markov-Chain, not jumping further then a nearest neighbour point in a single time step, allowing the system to move within a Point Cluster and then via a connected Trajectory Cluster to other Point Clusters throughout time. This is visually analogous to two probability density functions representing Point Clusters connected by a probability density representing a Trajectory cluster, as shown in Fig. 5.1. The actual description of this behaviour is more formally described in the following.

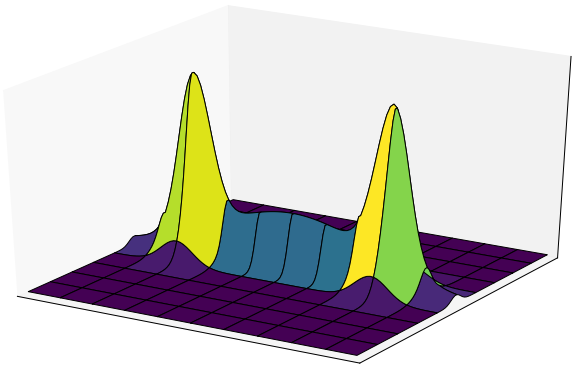

Figure 5.1: *A visual representation of what the probability density function of two described Point Clusters connected by one Trajectory Cluster may look like if reduced in dimensionality to a two dimensional plane.*

For this work the Euclidean distance D^P between any two points (system states) is defined as:

$$D^P((a_1,a_2,\ldots,a_n)_{t_1},(b_1,b_2,\ldots,b_n)_{t_2}) = \sqrt{\sum_{i=1}^{n}\{|b_i-a_i|\}^2}$$

Given the assumption that noise has a Gaussian distribution, it is expected that there will be more observed states close to the "true" system state in the centre of a Point Cluster, and fewer observed states further away from that "true" system state, and in practice it is expected that there be a maximum distance between observed states in a Point Cluster — this maximum distance is referred to as δ. Following this, the probability density function of these distances should peak for low values and decline for larger values of δ. Further a Point Cluster is defined to be "stable" if at least one visit to that Point Cluster lasts at least ζ time steps (where for $\zeta = 1$ neither the previous or the next state are in the same Point Cluster as the current

state). This allows us to define a Point Cluster PC_a as follows:

$$PC_a = \{s_{i,n} \in S_n\} \mid \forall j,k \; D^P(s_{j,n}, s_{k,n}) \leq \delta$$
$$\wedge \; (\forall b)_{b \neq a} \; s_{i,n} \in PC_a \iff s_{i,n} \notin PC_b$$
$$\wedge \; \exists z \;\; s_{z,n}, s_{z+1,n}, \ldots s_{z+\zeta,n} \in PC_a$$

These Point Clusters are connected together by trajectories, a consecutive progression in time of system states from one Point Cluster to another. A single trajectory $T_{s,e,m}$ from PC_s to PC_e is defined as:

$$T_{s,e,m} = \{s_{i,n}, s_{i+1,n}, \ldots, s_{i+\nu,n} \in S_n\} \mid s_{i-1,n} \in PC_s$$
$$\wedge \;\; s_{i+\nu+1,n} \in PC_e$$
$$\wedge \;\; \forall o \;\; s_{i,n}, s_{i+1,n}, \ldots, s_{i+\nu,n} \;\notin PC_o$$

It is assumed that for two Point Clusters connected by more then one trajectory those trajectories can be grouped into Trajectory Clusters, categorising different routes through the configuration space (though there may exist only one Trajectory Cluster between two Point Clusters). Therefore a Trajectory Cluster is defined as an unordered set of trajectories that connect the same Point Clusters and are within a certain Euclidean distance κ of each other, given by the function D^T (defined on Page 133).

$$TC_k = \{T_{s,e,m} \, \forall m\} \mid D^T(T_{s,e,i}, T_{s,e,j}) \leq \kappa$$

A Point Cluster may have zero or more Trajectory Clusters entering it and zero or more Trajectory Clusters leaving it.

The nodes and edges of Jiang *et al's*[3] network representation map to the Point Clusters and Trajectory Clusters of the presented configuration space representation. The Trajectory Clusters hence capture the information (including error deviation) of the transition from one node to another, with the point cluster capturing the deviation that can exist within a node. The two new contributions of the

configuration space representation are (i) the use of continuous Granger causality patterns, and (ii) the explicit representation of deviations due to noise and measurement error.

5.4 Constructing analysis and predictive algorithm

To validate the hypothesised behaviour described above an approach is required, here an algorithm is proposed to discover Point Clusters and Trajectory Clusters present in a data set (as described above). Furthermore this algorithm can be employed to predict the next Point Cluster, based on current state of the system, using the Trajectory Cluster and Point Clusters that have been discovered. This functionality allows for validation of the hypothesised behaviour against the actual behaviour (discussed in more detail later in this chapter).

To do this first a configuration space must be populated with historic data of causality patterns (each with a time stamp to provide temporal ordering). For the following discussion it is assumed that a new series of at least two consecutive causality patterns is then introduced, where two points are required to establish direction of movement in the configuration space. This new series is referred to as a series of "test values" and it is desired to predict a future point in this series. When introduced to the configuration space each of these test values will fall into one of three states: part of a Point Cluster, part of a Trajectory Cluster, or neither a part of a Point Cluster nor of a Trajectory Cluster.

The prediction aims to find the next Point Cluster (analogous to the next regime) that the system will be in, by reference to where the test data is currently located in the configuration space, in what direction it is moving in the configuration space, and whether it overlaps with an historical Trajectory Cluster or Point Cluster. It is of course possible the test values might establish a new route that does not exist in the historic data, but for the purpose of prediction it is assumed that all test values will follow historical data.

The discussion of the novel prediction method is split into three parts, first the method of detection of Point Clusters is given, then the method for Trajectory

Clusters, and finally the prediction method which utilises the previous two methods.

Point Clusters

To ground a Point Cluster in its relation to the underlying time series of the system, it can be thought of as the collection of times (throughout time) that the Granger causality between the time series of the system was very similar. This does not mean that the actual time series themselves were similar during these periods, merely that the interactions between them were. ζ then governs the length of time the time series need to maintain these similar causal interactions to be included in this set. Point Clusters can hence be seen as capturing meaningful behaviour across time where a system returns to a similar state. Changes between Point Clusters can hence be seen has changing from one repeated (throughout time) causal structure to a sufficient different repeated causal structure, signifying a change in the interaction between the time series. Compared to previous clusters (such as that done by Jiang et al. [3]) which groups causality patterns that frequently transition between each other, but may be distinctively different, this approach groups causality patterns that are distinctively similar and hence may be treated as representing a single system state.

A Point Cluster has two notable characteristics that can be used for detection: (i) it is a region of high density within the configuration space, and (ii) it is a region where at least one path will remain for at least ζ time steps. To detect Point Clusters in the historical data first high density regions are detected within the configuration space, and then these regions are filtered based on the second criterion to discover those that are Point Clusters.

The detection of regions of high density is a problem often approached using clustering algorithms. There is a large breadth of literature on clustering algorithms; however, since a key characteristic in the problem is the density of the points, a natural fit is the density-based clustering algorithms [151]. Density clustering is based on the principle that regions of greater density exist within the space, which can be separated into clusters[178]. In particular influence is taken from Shared Nearest Neighbours (SNN)[179] and single linkage[154, 155], with neither of these

methods being fully sufficient due to their assumption of greatly reduced density between clusters. For example, with causal regime-shifting behaviour a Point Cluster may occur where a Trajectory Path slows down temporarily and yet sufficiently to be identified as a Point Cluster, and furthermore several Trajectory Paths may pass through a Point Cluster; these are both problematic for standard algorithms because the density boundary for a Point Cluster may not be sufficiently well defined (it merges with the density distributions of the Trajectory Clusters entering and leaving the Point Cluster) leading to possible unwanted linkage effects[76].

To address this problem, a network is defined $G(N,E)$ with the set of nodes N and the multiset of edges E where each edge gives the Euclidean distance between two nodes. N is initialised as the set of all observations in the configuration space, and E as all edges between those nodes, defining a fully connected network.

$$N = \{n_i\} \mid n_i \in S_n$$
$$E = \{e_{i,j}\} \mid (\forall i,j)_{i \neq j} \;\; n_i, n_j \in N \;\; \wedge \;\; e_{i,j} = D^P(n_i, n_j)$$

Stage 1: From E an edge distance is estimated that within which the majority of points in a Point Cluster are connected. It is expected that the elements in E will have a multimodal probability distribution with modes relating to frequently occurring distances such as the distance between Point Clusters, and between steps in Trajectory Clusters. Following the given definition of perturbations due to noise and errors one can expect the smallest edge lengths to be caused by noise and error perturbations, and on this assumption the mode with the smallest value in the probability distribution of E (if such a mode exists, which it is assumed to) gives the edge lengths generated by these perturbations.

To determine this initial mode an estimation of the probability distribution of E is made as follows: after removing values above the median to remove outliers, a histogram is generated of the remaining edge lengths using k fixed-width bins across the range (k is chosen empirically — 500 is typical). The popular "find_peaks" library function[180] is used to determine the first mode in the (possibly noisy) probability distribution of E; this is labelled as *firstmode*. The minimum element of

E is also determined, which is labelled *min*.

Looking ahead to Stage 2 it is required that retained nodes have a sufficiently high number of neighbours to facilitate a connectivity metric to distinguish between the nodes within a Point Cluster and the nodes on the periphery of a Point Cluster that connect to Trajectory Clusters; therefore a threshold edge length is calculated that is somewhat larger than the first mode. Specifically, the threshold edge length is defined as $\hat{S} = \mathit{firstmode} + (\mathit{firstmode} - \mathit{min})$.

Therefore, E and N can be updated as follows:

$$E' = \{e_{i,j} \in E\} \mid e_{i,j} \leq \hat{S}$$
$$N' = \{n_i \in N\} \mid \forall n_j \in N \;\; e_{i,j} \in E' \vee e_{j,i} \in E'$$

Stage 2: It is assumed that the majority of e_i will now connect points within the same Point Cluster, though some e_i may connect points in a Point Cluster to those in Trajectory Clusters (for example, where a Trajectory Cluster enters a Point Cluster). It is assumed that the relevant n_i (those with edges only to points within the same Point Cluster) will have a different number of connections than those with edges to points in a Trajectory Cluster. To remove edges between Point Cluster and Trajectory Clusters, a filter is employed based on the number of connections (the "degree") of a node, given by $deg(n_i)$:

$$E'' = \{e_{i,j} \in E'\} \mid \forall n_a, n_b \in N' \;\; deg(n_a) = deg(n_b)$$
$$\wedge deg(n_i) = |\{e_{c,d} \in E' \mid c = i \vee d = i\}|$$
$$N'' = \{n_i \in N'\} \mid \forall n_j \in N' \;\; e_{i,j} \in E'' \vee e_{j,i} \in E''$$

Stage 3: Finally the time condition ζ is used to construct a set of Point Clusters $PC = \{PC_i\}$ where each Point Cluster PC_i is a non-empty set of nodes that only have edges to other nodes in the same Point Cluster and where a path through the

Point Cluster takes at least ζ time steps:

$$PC_i = \{n_k \in N''\} \mid (\forall q)_{e_{p,q} \in E''} \; n_p \in PC_i \implies n_q \in PC_i$$
$$\wedge \; \exists z, \; \exists n_z, n_{z+1}, \ldots n_{z+\zeta} \in PC_i$$
$$where \; (\forall v)_{0 \leq v \leq \zeta - 1} \; timediff(n_{z+v}, n_{z+v+1}) = 1$$
$$timediff(s_{t_1,n}, s_{t_2,n}) = t_2 - t_1$$

Trajectory Clusters

The detection of Trajectory Clusters is dependent on the discovered Point Clusters, PC_i. The detection of Trajectory Clusters is split into two stages: (i) the grouping of paths with the same beginning and terminating Point Clusters, PC_s and PC_e, and (ii) analysing these groups to produce Trajectory Clusters.

Stage 1: Here a trajectory, $T_{s,e,k}$, is defined as an ordered set of time-consecutive nodes between two Point Clusters, PC_s and PC_e:

$$T_{s,e,k} = (\{n_i, n_{i+1}, \ldots, n_{i+\delta} \in S_n\}, \prec) \mid (\exists a \in PC_s \wedge \exists b \in PC_e)_{s \neq e}$$
$$\wedge \; timediff(a, n_i) = 1 \wedge timediff(n_{i+\delta}, b) = 1$$
$$\wedge \; (\forall v)_{0 \leq v \leq \delta - 1} \; timediff(n_{i+v}, n_{i+v+1}) = 1$$

Hence a group of trajectories is defined as having the same start and end components as $T_{s,e} = \{T_{s,e,k} \; \forall k\}$.

Stage 2: Differences in trajectories between two Point Clusters may be due to two causes; the inherent noise from the "true" state or the existence of fundamentally different routes. It is desired to broadly cluster together those trajectories that follow approximately the same route, referring to these clusters as Trajectory Clusters. To transform these groups into Trajectory Clusters requires some concept of the distance between trajectories. Two popular distance metrics are Longest Common Sub-Sequence [181, 182] and Hausdorff distance [183, 184, 185], however

for simplicity in this thesis the Euclidean distance between two trajectories is used, assuming that their observations are identically spaced in time and have an equal number of observations, n: [183]

$$D^T(T_{s,e,i}, T_{s,e,j}) = \frac{1}{n}\sum_{k=1}^{n} D^P(select^T(T_{s,e,i}, k), select^T(T_{s,e,j}, k)) \quad | \quad n = |T_{s,e,i}| = |T_{s,e,j}|$$

$$select^T((\{a_1, a_2, \ldots, a_n\}, \prec), k) = a_k$$

With this distance measure a single linkage approach is used, clustering together trajectories which are closer than a set distance to another trajectory already in the cluster. For simplicity this distance is set as the mean Euclidean distance of the group of trajectories, assuming that separate routes will have a separation of more than the mean. Therefore a Trajectory Cluster can be expressed as follows:

$$TC_{s,e,c} = \{T_{s,e,k} \forall k\} \mid \forall T_{s,e,i}, T_{s,e,j} \in TC_{s,e,c} \;\; D^T(T_{s,e,i}, T_{s,e,j}) \leq mean(T_{s,e}),$$

$$mean(T_{s,e}) = \frac{\sum_{i,j} D^T(T_{s,e,i}, T_{s,e,j})}{|T_{s,e}|} \;\; (\forall \; T_{s,e,i}, T_{s,e,j} \in T_{s,e})_{i \neq j}$$

Prediction

The presented novel prediction approach is referred to as Assisted Prediction of Causal States (APOCS), and describe its procedure as follows. When applied to data APOCS can be separated into three stages: (i) create a causality configuration space from historical time series data, (ii) identify Point Clusters and Trajectory Clusters in the configuration space, and (iii) on the basis of newly observed values, predict future values. Predictions can be made after stages (i) and (ii) are completed, however in practice it is likely that a user will periodically wish to rerun stages (i) and (ii) to include the newer data points into the historical data.

The method presented focusses on predicting the next Point Cluster, and for usability transforming that prediction into a binary causality pattern. It is important to note that this approach does not aim to predict a constant number of time steps ahead, and hence the time range of the prediction will vary (but will always be

given). It should also be noted that time steps are measured in causality patterns, and hence one may wish to relate this to a number of data points by referring to the windowing method. For a data point within a Point Cluster the prediction will be for a single time step, but for a point in a Trajectory Cluster the prediction will be for the number of time steps left in the trajectory plus one, to bring it to the next Point Cluster.

Taking a set of newly observed data that gives two consecutive causality patterns, $s_{i,n}$ and $s_{i+1,n}$ respectively, a prediction can be made as follows:

1. Identify the Point Cluster PC_j that is closest to $s_{i+1,n}$, measured to the nearest point within PC_j, with this distance to the closest Point Cluster being d_{PC}.

2. Identify the closest Trajectory Cluster TC to the Trajectory comprising the two points $s_{i,n}$ and $s_{i+1,n}$. Select two historically observed points PT_1 and PT_2 forming a two-point trajectory T' as follows:

$$\begin{aligned} T' = \{PT_1, PT_2\} \mid\ & PT_2 \in TC \\ & \wedge\ \mathit{timediff}(PT_1, PT_2) = 1 \\ & \wedge\ d_O = D^P(s_{i,n}, PT_1) \\ & \wedge\ d_N = D^P(s_{i+1,n}, PT_2) \\ & \wedge\ d_{TC} = (d_N + d_O)/2 \\ & \wedge \min_{PT_1, PT_2} d_{TC} \end{aligned}$$

3. - If $d_{PC} < d_{TC}$, predict the Point Cluster PC_j, and give a prediction time of one time step.
   - If $d_{TC} \leq d_{PC}$ predict the Point Cluster at which TC terminates. Give a prediction time of the average remaining time steps in TC plus one.

4. Let the predicted Point Cluster be PC_x. Then predict the average causality pattern $\overline{PC_x}$ of PC_x. This done by calculating the mean values at each dimension (where the mean value for each dimension is $x_{a,b}$ for the dimension

characterised as the causality between variable a and b) for all points in PC_x:

$$\overline{PC_x} = \{x_{a,b}\}\forall a,b \mid x_{a,b} = \frac{1}{|PC_x|}\sum_{k=1}^{|PC_x|} index(select^{PC}(PC_x,k),(a,b))$$

$$index(\{c_{i,j}\},(a,b)) = c_{a,b}$$

$$select^{PC}(\{a_1,a_2,\ldots a_n\},k) = a_k$$

5. Apply an alpha level filter (0.05) to the causality pattern to produce the binary causality pattern.

6. Predict the binary causality pattern and its associated time step from the current time step.

5.5 Results: Synthetic data

To validate the APOCS algorithm it is applied to synthetic data generated via a Monte Carlo approach. For benchmarking two alternative approaches are also applied to the same data set. These methods are (i) A Stationary approach, that takes the current continuous Granger causality pattern in the test data and predicts that the pattern will remain unchanged; and (ii) A Maximum Connected approach, that turns all causality patterns into binary causality patterns and then predicts a change in causality pattern will be the most frequent transition from the current pattern. As mentioned previously the APOCS algorithm has a dynamic prediction range, since the trajectory governing the transition may take a number of time steps to be realised. For the purpose of comparison and not to overweight the accuracy of the APOCS algorithm its predictions are compared to those of one time step for the alternative approaches, as shown in Chapter 4, this range gives the highest prediction accuracy for these methods. Due to this the Rolling Linear Fit (RLF) method, from Chapter 4, will not be used as a bench mark here, since it was demonstrated that at a prediction range of one time step this methods accuracy is lower then the alternatives.

APOCS is stress tested by investigating its average performance over a range of

system parameters for a system that follows the regime-shifting behaviour outlined on Page 123. This uses a Monte Carlo simulation where each run is a new instance of the configuration space and the historical data within it, with the parameters defining this being reselected for each run (but remaining constant within a given run). This simulation is not designed to directly simulate a real world system, rather to test APOCS with numerical variations of regime shifting behaviour.

The following describes a single run of the above Monte Carlo simulation, and an analysis of the results for the full Monte Carlo simulation.

Monte Carlo Run

For a single Monte Carlo run, initially generate a sequence of causality patterns whose evolution follows the regime-shifting behaviour introduced on Page 123. Then split this sequence so that the first two thirds are used as historical data to populate the configuration space and the last third represents "newly observed" test data whose behaviour will be predicted. Each run produces a sequence of prediction results, one for each subsequent point in the test data.

For each run a number of system parameters are set to generate the sequence; these parameters are probabilistically selected at the beginning of the run but stay constant throughout a single run.

First the sequence of causality patterns is produced without noise or errors. To achieve this the behaviour that this sequence will obey is define as follows:

- First define the points in the configuration space that represent the regimes between which the system will shift. Initially only the centre points are defined — later, noise and errors will be added to create Point Clusters.

  This is done by first selecting the number of points from a uniform distribution (limit this distribution to be in the range $[3,6]$ for simulation performance), and setting the number of variables that the causality pattern will have (set this for each run, with runs being done at $3,6,9$ and 12). The actual *p-values* of the causality patterns for these points are then produced, this is done through a two stage process; (i) each link in the causality pattern is randomly selected to either be causal or not causal, with a $50/50$ probability. (ii) The p-value of

each link is then selected from a uniform distribution, where for a causal link the range is $[0, \alpha]$ and for a non-casual link the range is $[\alpha, 1]$ (here $\alpha = 0.05$ is used).

- Second, randomly choose one of these points to be the starting point for the simulated historical sequence.
- The length of the sequence is parameterised on the number of transitions to occur, here this is set as 100.
- The generation of the sequence proceeds via the following logic:
  - If the most recently generated point is one of the previously defined regime points, and if fewer than ζ immediately preceding points in the sequence were also at the same regime point, then the next generated point will be at the same regime point (typically set $\zeta = 3$). However, if this test fails then the sequence must consult a transition matrix (described below) to move to the next regime (it is permitted that the target might be the same as the current regime).
  - To move away from one regime point to another, the target regime is chosen from the transition matrix and four new points in the sequence are generated, evenly spaced between the current regime point and the target regime point (so that a transition between the same two points will always follow the same path). If the current and target regime points are different, the path will be a straight line, but if they are the same then the path will be an arc.
  - The transition matrix defining these transitions is given by $\mathbf{M}$ where $\mathbf{M}_{i,j}$ represents the probability of point i transitioning to point j where $\sum_j \mathbf{M}_{i,j} = 1$. This matrix is considered a parameter and the values within it are selected from uniform distribution at the beginning of the run such that $\sum_j M_{i,j} = 1 \; \forall i$, and held constant throughout a the whole run.
- Once the sequence is produced then add an error in all dimensions to every

causality pattern within the sequence. This error is drawn from a Gaussian distribution centred on 0 with a standard deviation of 0.005.

The Monte Carlo simulation described above was run for a total of a hundred runs evenly split between four systems each with a different number of variables: 3, 6, 9, and 12. For each run two different metrics are produced and employed as follows:

1. Metric 1; *p-value* comparison: For the APOCS and the Stationary method (the Maximum Connected method was not included as it does not produce *p-value* results). The Euclidean distance between the predicted value and the true value (based on the methods prediction range) in *p-value* space divided by the maximum possible distance for the dimensionality of the space (e.g. $\sqrt{n}$ in n-dimensional space).

2. Metric 2; binary comparison: For the APOCS, the Stationary, and the Maximum Connected methods (an alpha level filter of 0.05 was applied to APOCS and Stationary to produce binary causality patterns — the Maximum Connected method results are already in the form of a binary causality pattern). The Euclidean distance between the predicted binary causality pattern and the true binary causality pattern divided by the maximum possible distance for the dimensionality of the space.

The results for metric 1 are shown in Fig. 5.2 and the results for metric 2 are shown in Fig. 5.3, for convenience the t-tests of the combined series across variable numbers of these results are shown in Table. 5.1.

From the analysis one can see that in this synthetic stress test APOCS outperforms both presented alternatives to a statistically significant degree, of 1% for the continuous case and for $n = 9, 12$ for the binary case, and at a 5% level for $n = 3, 6$ for the binary case. It should also be noted that while APOCS outperforms the alternatives, it also provides predictions over a larger time scale than the alternatives. When comparing the alternative methods to each other it can be seen from Figure 5.3 that they produce very similar results, this is driven by the fact that for a

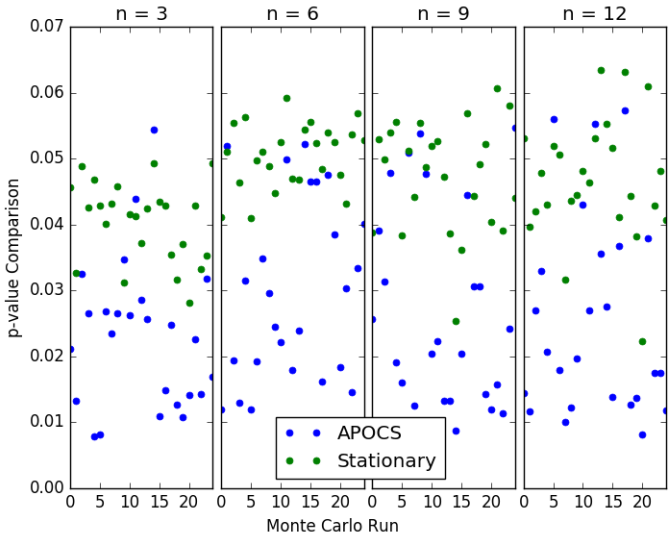

Figure 5.2: *Results following metric 1 in p-value space, comparing the prediction of results of APOCS (on dynamic time scale) and a Stationary (one time step) approach. Results are for four separate Monte Carlo simulations run with 3, 6, 9, and 12 variables.*

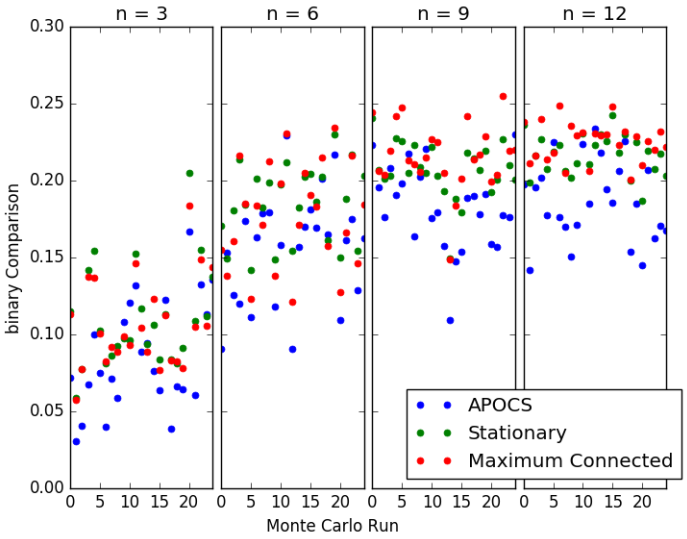

Figure 5.3: *Results following metric 2 in binary space, comparing the prediction of results of APOCS (on dynamic time scale), a Stationary (one time step), and a Maximum Probability (one time step) approach. Results are for four separate Monte Carlo simulations run with 3, 6, 9, and 12 variables. The Monte Carlo runs are the same as those shown in Fig. 5.2*

binary causality pattern often its most frequent transition is to itself, making the prediction consistent with a Stationary approach. It can also be seen that the addition

of a binary mask to the *p-value* results for APOCS and Stationary lead to a decrease in accuracy of the predictions, this is particularly apparent with higher numbers of variables. This is due to the hard cut in the result space implemented by a binary approach potentially causing points in a *p-value* space that are close by to appear further apart, hence these results show how a *p-value* approach can hence lead to a better understanding of results, by not obscuring this information.

Table 5.1: Results of significance tests between APOCS method and the alternative methods for data presented in Fig. 5.2 and Fig. 5.3. A $p-value < 0.01$ represents rejection of the null hypothesis at the 1% level, and a $p-value < 0.05$ represents rejection of the null hypothesis at the 5% level.

Methods	t-Statistic/p-value (3 sf)			
num. var. (n)	3	6	9	12
Metric 1: p-value comparison				
APOCS/Stat.	-6.86/0.001	-7.13/0.001	-5.89/0.001	-5.98/0.001
Metric 1: binary comparison				
APOCS/Stat.	-2.50/0.0159	-3.32/0.00175	-3.66/0.001	-5.01/0.001
APOCS/Maxi. Con.	-2.25/0.0291	-2.18/0.0341	-4.70/0.001	-6.47/0.001

5.6 Results: International oil market

Now that the regime shifting behaviour has been defined and a method for investigating it proposed, an experiment can be run on the international oil market, using the data presented in Chapter 2.4, to analysis the hypothesised behaviour for this system.

The data set minus the last twenty values are taken for investigation (these last twenty values are removed for validation, which will be explained later), this data has the APOCS algorithm applied to determine the presence of Point Clusters and Trajectory Clusters for a varying ζ parameter; the results of this analysis are shown in Fig. 5.4. From this experiment it can be seen that as the ζ parameter is increased the number of Point Clusters and Trajectory Clusters decreases rapidly. This demonstrates that only a few Point Clusters in the system have "stability" of more then a few time steps, however clustering with lesser stability is very common. There is substantial difference in the number of Trajectory Clusters (maximum 6) compared with the number of Point Clusters (maximum 382). This implies that

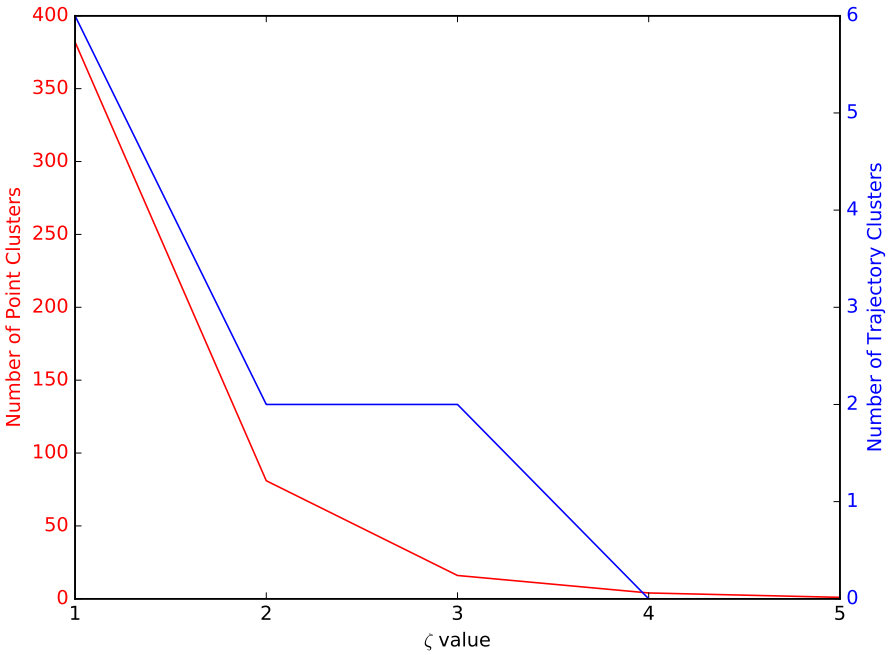

Figure 5.4: *Number of Point Clusters and Trajectory Clusters for data presented in Chapter 2.4 with the last 20 values removed, for ζ values 1, 2, 3, 4, and 5, with (Point Clusters, Trajectory Clusters), (382, 6), (81, 2), (16, 2), (4, 0), and (1, 0).*

for this international oil market data most transitions between Point Clusters do not form Trajectory Clusters according to the definition in Sec. 5.3; instead, most transitions appear to follow non-repetitive paths in the configuration space (or change Point Cluster over a single step). The decrease in Trajectory Clusters can be seen as linked to the decrease in Point Clusters, and is not directly determined by ζ but by the number of Point Clusters that trajectories can go between. Therefore, this decrease and lower number of Trajectory Clusters compared to Point Clusters implies that only certain Point Clusters are linked by trajectories and that as these are removed by increasing ζ the Trajectory Clusters are also removed. Further, this implies that the Point Clusters that are linked by Trajectory Clusters are those with a lower ζ and hence are less stable throughout time. From this it can be concluded that some transitions between less stable Point Clusters may be governed by Trajectory Clusters, but for more stable Point Clusters they are not.

This analysis determines the existence of clustering within the configuration

space of the system (Point Clusters). Based on the prior investigations of this data (Chapter 3 and 4) that demonstrates the system maintains similar causality patterns throughout its dynamic behaviour, it can be inferred that if these Point Clusters accurately represent the dynamic behaviour there should capture these prominent causality patterns. If this is the case, it can be hypothesised that taking the nearest Point Cluster as a prediction for the causality pattern at the next time step (following the average process presented earlier) will produce similar results to that of taking the current causality pattern (or the next node in a complex network, as demonstrated in Chapter 4).

Here the hypothesis in the above paragraph is investigated, by taking a prediction for APOCS of the closest Point Cluster for the next time step, and comparing to the two alternative approaches. To run this experiment a parameter selection of $\zeta = 3$ is made, this is done to isolate the more stable Point Clusters within the system while maintaining a level of granularity that is lost with a ζ above this level. As stated at the start of this section the last 20 time steps of the data have been kept for validation, hence here these will be used as the test data, with prior data used for training. The continuous Granger measure results are shown in Fig. 5.5, the binary Granger results are shown in Fig. 5.6, the statistics for these results are shown in Table. 5.2, and significance tests between these methods are shown in Table. 5.3. This experiment shows that there is no statistical difference (at a 1% significance level) in prediction quality between the APOCS method and the two presented alternatives in for this sample data, demonstrating that the hypothesis in the previous paragraph is correct at this level, and hence that the Point Cluster found via APOCS put presented Section 5.3 accurately represent the system.

5.7 Discussion

This chapter explored a hypothesis for the dynamic nature of the regimes/cluster and transition behaviour discovered in Chapters 3 and 4 for the data discussed in Chapter 2.4. The hypothesis investigated was, is the dynamic evolution of the sample data from the international oil market governed by a series of of noisy perturbations

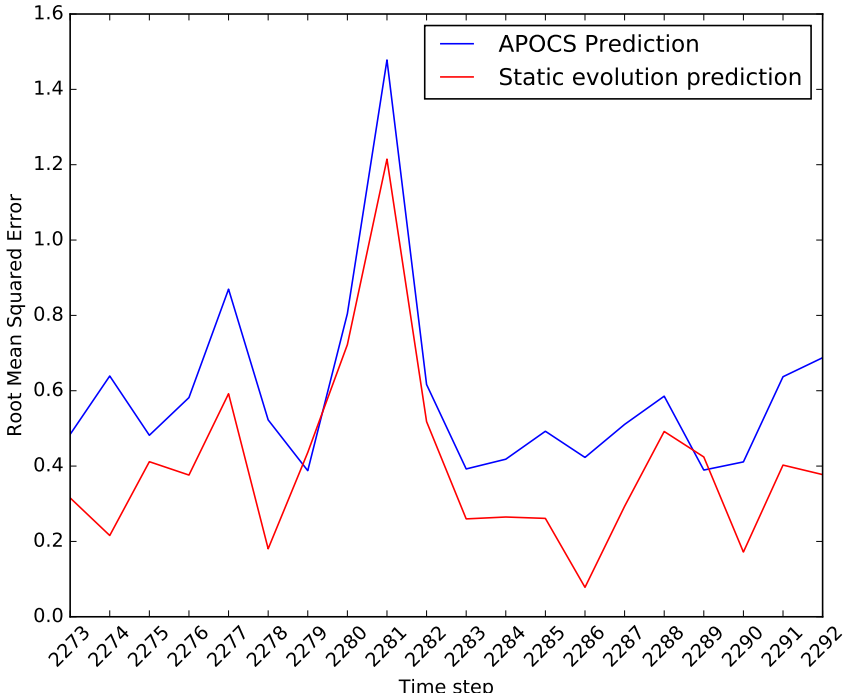

Figure 5.5: *Results following metric with a continuous Granger measure, comparing the prediction results of the three presented methods.*

Table 5.2: Mean and variance values over all test data for: Assisted Prediction of Causal States (APOCS), Static Evolution (SE), and Network Based Transition (NBT).

Method	Mean (3 dp)	Variance (3 dp)
Metric 1: continuous comparison		
APOCS	0.591	0.059
SE	0.400	0.057
Metric 2: binary comparison		
APOCS	0.524	0.425
SE	0.624	0.410
NBT	0.745	0.395

around causality patterns that are connected in a repeated manner throughout time.

To investigate this first a configuration space is defined in which the dynamic series of causality patterns can be encoded in a manner that preserves their structure and time ordering. Next an archetypal definition of a hypothesised behaviour based on the results from Chapter 3 and 4 is proposed within the configuration space. Following this an algorithm is proposed for the investigation and prediction of systems following the presented archetypal behaviour, this algorithm is named Assisted Pre-

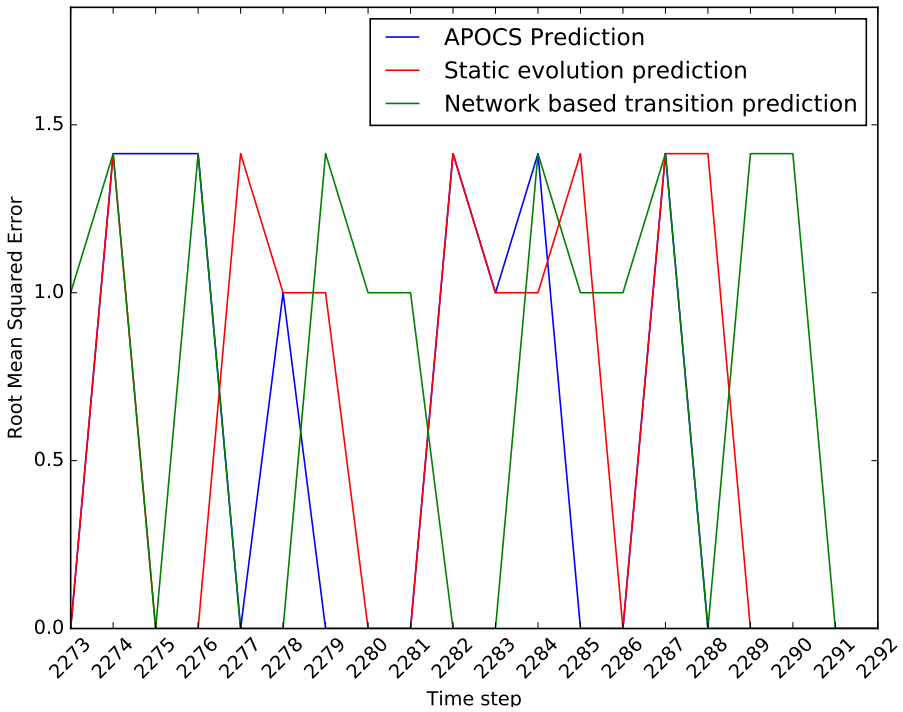

Figure 5.6: *Results following metric with a binary Granger measure, comparing the prediction results of the three presented methods.*

Table 5.3: Results of t-tests between prediction methods on test data for: Assisted Prediction of Causal States (APOCS), Static Evolution (SE), and Network Based Transition (NBT). Note that a $p-value \geq 0.001$ indicates no statistical difference at a 1% significance level.

Methods	Statistic (3 dp)	P-value (3 dp)
Metric 1: continuous comparison		
APOCS/SE	2.433	0.020
Metric 2: binary comparison		
APOCS/SE	-0.477	0.636
APOCS/NBT	-1.062	0.295

diction of Causal States (APOCS).

The presented methodology undergoes a synthetic test by constructing data following the defined behaviour and applying a Monte Carlo approach to benchmark APOCS against two alternatives. For this experiment APOCS statistically outperformed the alternatives in every set-up.

Finally the hypothesised behaviour for the data of the international oil market was investigated using the presented methodology. The results for a number of

ζ values demonstrate that the majority of the clustering in this data either has no stability or very little stability, with the number of detected Point Clusters rapidly decreasing as the ζ value increased. The behaviour of the Trajectory Clusters corresponded with this, decreasing with an increase of ζ, this being driven by the removal of Point Clusters to which the Trajectory Cluster is connected. There was also a notable difference in the number of Trajectory Clusters compared to Point Clusters at all ζ values, with the number of Trajectory Clusters being significantly lower. From this a conclusion on this data of the international oil markets dynamic behaviour can be drawn, being that the transitions between Point Clusters does not primarily follow a repeated path through the configuration space. Implying the existence of non-repetitive transition behaviour or transitions that move between Point Clusters in a single time step. To validate that the Point Cluster detected (specifically those at a $\zeta = 3$ were chosen) accurately represent the systems dynamics a comparison with a complex network (demonstrated as informative in Chapter 3) representation and a static representation (demonstrated as informative over at least one times step in Chapter 4) was made, with validation being considered if these give similar high level information. This was achieved via a prediction comparison of the three, with APOCS being used to predict the closest Point Cluster, a complex network predicting the most common transition, and a static approach predicting the same causality pattern. This comparison was tested and the results shown to be statically the same at a 1% significance level, confirming that the Point Clusters represent an important aspect of the systems behaviour. This demonstrates that through the use of a configuration space a richer low level description of the dynamic behaviour of the multivariate data of the international oil market can be found. With the Point Clusters in the configuration space inherently giving a more informative description of the system then the nodes in a complex network (with the nodes representing causality patterns). Thus this chapter presents a methodology that reveals new information from the data on the international oil market.

Chapter 6

Conclusions and Future Work

In this final chapter, the main contributions of the research are summarised prior to proposals being given as to how the work can be extended in the context of dynamic time-varying multivariate systems.

6.1 Discussion and summary of contributions

The main objective of this research was to investigate whether the behaviour of the underlying causality links within a system could be harnessed to gain information relevant to the analysis and prediction of the evolution of a systems causality pattern, specifically in the case of the international oil market. Therefore the contribution to science put forth in this thesis is to construct methodologies, algorithms, and experiments, in order to address this scientific challenge. As well as, the expansion of knowledge on the dynamics of the international oil market.

The beginning of this thesis discusses how the analyses of time-varying complex systems is becoming a major area of interest, following on from the mature field focussed on time-varying univariate systems. In particular work relating to the international oil market is referenced, and the importance of this archetypal system is highlighted. An overview of the methods employed in current investigations of dynamic systems is given, discovering that much of the current approaches focus on "high level" behaviour of the evolution and ignore "low level" behaviour within the specific changes of the causality pattern (where the "low level" behaviour is the complex network of the interactions between the variables of the system, the

"high level" behaviour is hence the movement between these structures). From this it is concluded that there is a need to investigate useful aspects of this low level behaviour, and design methodologies to incorporate these into analysis approaches.

To achieve this, this thesis has investigated the nature of this low level behaviour within sample data of the international oil market, approaching this investigation from a number of directions to give a more holistic understanding of the systems dynamics. The work here has provided a number of different methodologies, that have been demonstrate to be successful in uncovering information on the dynamic behaviour of time-varying complex systems. Such methods and findings would be of interest to practitioners in areas where an understanding of these complex time-varying dynamics is required, such as traders in the financial markets.

The work presented in this thesis achieved its objectives, which in overview were: 1) can time-varying behaviour of a system be more fully explored via the inclusion of information on the transition properties between causality patterns; 2) can a meaningful one-dimensional time series representation of a time-varying system be constructed based upon that system's overall transition behaviour; 3) Can a configuration space be utilised to test a hypothesis of specific system's dynamic behaviour, described at both low- and high-level. The work undertaken in these areas is discussed more thoroughly below.

Chapter 3 focused on incorporating knowledge on the change in the causality pattern over a transition, through directly expanding on the current state of the art. Further, based on this proposing a number of relevant properties and metrics for the systems evolution. To capture information regarding the transition between causality patterns two metrics were proposed: α the total difference in causality, and β the net directional change in causality. To harness the information stored within these metrics a new complex network methodology for capturing the dynamic nature of a time-varying complex system is constructed, which is referred to as the Underlying Structural Information Consideration Network (USIC-Network). This network encodes the evolutionary transitions between the causality patterns, as well as their; frequency, total difference in causality, and net directional change in causal-

ity. With this network construction a number of properties and metrics associated with a system represented by USIC-Network are defined and motivated. These are as follows: pattern stability, measuring the likelihood of a pattern not transition on the next time step; directional change in causality, measuring if the system is likely to gain or lose causality in the next time step; noise clusters, a clustering property that defines groupings of similar patterns by structure; regimes of total causality level, a clustering property that defines groupings of patterns whose total causality is equivalent; net causality change pathways, a property defining pathway structures where the total causality is either consistently decreasing or increasing. Employing this methodology, and comparing to synthetic results, the following behaviour of the sample data of the international oil market is discovered: the evolution appears to be dependent on small net changes in the causality pattern leading to eight noise clusters and nearly no regimes of total causality level or net causality change pathways. Implying that the system moves between a number of regimes in which it maintains a similar causality pattern.

Building on the work of Chapter 3, Chapter 4 uses knowledge of the low level behaviour to investigated two specific types of movement within the causality pattern structure, allowing for a more in-depth level of analysis. These movements are defined as paradigms under which a dynamic systems behaviour may evolve. These two paradigms cover the following behaviour: a evolution that is driven by changes a net change in causality; and a evolution that is driven by a change in the causality structure, but not a net change in the causality. For each of these paradigms a one-dimensional metric is proposed; that may be used as a proxy to represent the current state of the system, and allows for a informative separation between states, as based on the paradigm behaviour. These metrics are: Sum of Causality (SoC) for the first metric, this is the total causality of the causality patter; and Sum Difference from Mean (SDM) for the second metric, this is the difference (α) between the current causality pattern and a mean/reference causality pattern. Using these metrics a system evolving under one of the paradigms can have its evolution reconstructed in terms of a metric, allowing for a time series that is amenable to information dis-

covery. A methodology for analysis of a system to determine which paradigm it is best described by is then presented. Using this methodology a series of experiments were conducted on the sample data of the international oil market. These determined that the this data belongs to the first paradigm and hence can be represented the using SoC metric. Further investigation of the data under this representation, found that it exhibited a a time-varying nature best explored when expressed without the binary nature of classic Granger causality. To allow for this investigation a continuous measure of the Granger causality test based on the *p-value* is presented. The evolution under this measure exhibited a smooth behaviour, with investigation determining that this behaviour contained autocorrelation under a short lag (under 25 time steps). Based on these findings a prediction method is presented for this data, designed to construct linear predictions for a short lag, this is referred to as Rolling Linear Fit (RLF) prediction. Using this an experiment is run on the data of the international oil market, demonstrating that RLF prediction statistically outperforms three apparent alternatives for predicting the causality pattern of the series. Specifically this demonstrates that RLF is able to predict up to ten time steps into the future with out lose of accuracy and that this increases the predictive power over what is achievable with a complex network representation.

Chapter 5 presents work expanding on the results from the previous two chapters, to further investigate the regime behaviour discovered in Chapter 3 and hypothesising based on Chapter 4 that these regimes can be more precisely explored under a continuous causal measure and that a repeated transition between them may exist. To explore this hypothesis this chapter first presents a configuration space, within which the trajectory of the evolution can be mapped, in a information dense manner. With the introduction of this configuration space a formal definition of the hypothesis is able to be presented, in the form of a archetypal regime transition behaviour. Based on this an algorithmic approach to analysing data in the configuration space for the presented behaviour is presented, with a discovery of the behaviour allowing for predictions to be made. This methodology is presented on synthetic data, demonstrating that for a system under going this behaviour predictions can be made

against two apparent alternatives that statistically out perform them, while also having a larger prediction range. This investigation also demonstrates that an analysis in continuous Granger measure produces better results then that of a binary measure. Using this methodology an investigation of the sample data of the international oil market was undertaken, determining the following results: this data contains clustering around specific causality patterns that can be seen as noise/deviation, that is demonstrated to accurately encapsulates the data's evolution, the evolution between these clusters does not demonstrate a repeated path through the configuration space implying a more complex/random behaviour.

Throughout this thesis a number of methodologies are presented, algorithms, properties, and metrics, these can be taken together as a body of work. This body of work can be seen as a set of tools to aid in the investigation of dynamic behaviour of time-varying systems of interacting variables. Specifically allowing for a deeper understanding of the systems evolution through a understanding of its behaviour in terms of the causality patterns underlying it. Allowing for a low level analysis to supplement and expand on the high level analysis that already exist.

Using the presented methodologies a number of experiments and analyses on the sample data of the international oil market have been conducted, furthering the research on this market. As a whole the findings can be summarised as follows: the data does not exhibit; movement maintaining its total causality, repeated pathways changing the net causality, or repeated pathways through the configuration space. The data's evolution appears to be restricted to clusters of similar causality structures, seen as noise/deviations around a central causality pattern in the configuration space, with transitions predominately changing a single causality link. The evolution displays autoregressive behaviour over short lag times, allowing for predictions, however, over long lag times the systems evolution is does not contain correlation. Therefore the sample data of the international oil market can be said to move in a unrepeated manner between a number of causality patterns that it deviates around.

6.2 Future work

This research pursued its objectives within the pre-defined scope successfully. The work presented here also inspires future research directions. A number subsequent topics that may be worth investigation are as follows: First, future work may expand upon the properties and metrics constructed for the USIC-Network model, with the ones presented in this thesis being only a selection of possible approaches. Second, the construction of new paradigms to describe dynamic behaviour of the underlying structure may be undertaken by future work. The paradigms presented in this thesis are appropriate to the data and cover two conceptual important behaviours, but leave room for a broad number of other behaviours. Third, a descriptions of transition behaviours between regimes (clusters of causality patterns that are "similar") may be undertaken by future work. This thesis investigates one transition behaviour, that being the repetitive smooth movement through configuration space, however due to the lack of evidence for the existence of this type of transition there exists space for other (potentially more complex) transition behaviours to be defined and tested for in future work. Finally, future work can be inspired more broadly by this thesis to create new methodologies and approaches for the analysis of time-varying multivariate systems where the underlying structurally changes are considered when constructing the description of the overall dynamic behaviour.

6.3 Concluding remarks

The analysis of complex dynamic behaviour within time-varying multivariate systems is, and will remain, an important aspect of research in many disciplines, partly due to the ever increasing interconnectedness between systems of interest. Though much work is done with univariate data, no-one would doubt the importance of considering confounding variables and the impact they may have on a systems evolution. This field has seen much interest in recent years, however this work has primarily been focused on high level behaviour. Thus the consideration of the structural change of the systems causality pattern opens up new approaches to the analysis of the dynamic behaviour of systems. The hope of this work is that it will

prompt further use of this underlying structure for the understanding of systems dynamic behaviour.

Bibliography

[1] Y. Qi, H. Li, S. Guo, and S. Feng. Dynamic transmission of correlation between investor attention and stock price: Evidence from china's energy industry typical stocks. *Complexity*, 2019.

[2] X. Yu, S. Shi, L. Xu, J. Yu, and Y. Liu. Analyzing dynamic association of multivariate time series based on method of directed limited penetrable visibility graph. *Physica A*, 545, 2020.

[3] M. Jiang, X. Gao, H. An, H. Li, and B. Sun. Reconstructing complex network for characterizing the time-varying causality evolution behavior of multivariate time series. *Scientific Reports*, 7, 2017.

[4] X. Jia, H. An, W. Fang, X. Sun, and X. Huang. How do correlations of crude oil prices co-move? a grey correlation-based wavelet perspective. *Energy Economics*, 49, 2015.

[5] E. Keogh and S. Kasetty. On the need for time series data mining benchmarks: A survew and empirical demonstration. *Data Min. Knowl. Discov.*, 7(4), 2003.

[6] T-C. Fu. A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24, 2011.

[7] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah. Time-series clustering – a decade review. *Information Systems*, 53, 2015.

[8] R. H. Shumway and D. S. Stoffer. Arima models. *Time Series Analysis and Its Applications*, 2017.

[9] Y. Zou, R. V. Doner, N. Marwan, and J. F. Donges. Complex network approaches to nonlinear time series analysis. *Physics Reports*, 787, 2019.

[10] L. Lacasa, V. Nicosia, and V. Latora. Network structure of multivariate time series. *Science Reports*, 2015.

[11] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 1969.

[12] H. Ye, E. R. Deyle, L. J. Gilarranz, and G. Sugihara. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci Rep*, 5, 2015.

[13] E. Ghysels, J. B. Hill, and K. Motegi. Testing for granger causality with mixed frequency data. *Journal of Econometrics*, 192, 2015.

[14] B. Candelon and S. Tokpavi. A nonparametric test for granger causality in distribution with application to financial contagion. *Journal of Business & Economic Statistics*, 34, 2016.

[15] A. B. Barrett, L. Barnett, and A. K. Seth. Multivariate granger causality and generalized variance. *Phys. Rev. E*, 81, 2010.

[16] M. Asai, M. McAleer, and J. Yu. Multivariate stochastic volatility: A review. *Econometric Reviews*, 25, 2006.

[17] M. J. Schervish. A review of multivariate analysis. *Statistical Science*, 2(4), 1987.

[18] Marco Bardoscia, Stefano Battiston, Fabio Caccioli, and Guido Caldarelli. Pathways towards instability in financial networks. *Nature communications*, 8(14416), 2017.

[19] Robert M. May. Will a large complex system be stable? *Nature*, (238):413–414, 1972.

[20] R. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 2002.

[21] Z. Gao and N. Jin. Complex network from time series based on phase space reconstruction. *Chaos*, 19, 2009.

[22] R. Jammazi, R. Ferrer, F. Jareno, and S. J. H. Shahzad. Time-varying causality between crude oil and stock markets: What can we learn from a multiscale perspective? *International Review of Economics & Finance*, 49, 2017.

[23] S. An, X. Gao, M. Jiang, and X. Sun. Multivariate financial time series in the light of complex network analysis. *Physica A*, 503, 2018.

[24] Y.-J. Zhang and L. Zhang. Interpreting the crude oil price movements: Evidence from the markov regime switching model. *Applied Energy*, 143, 2015.

[25] Y.-J. Zhang and T. Yao. Interpreting the movement of oil prices: driven by fundamentals or bubbles? *Economic Modelling*, 2016.

[26] X. Huang and S. Huang. Identifying the comovement of price between china's and international crude oil futures: A time-frequency perspective. *International Review of Financial Analysis*, 2020.

[27] X. Gao, H. An, W. Fang, X. Huang, H. Li, W. Zhong, and Y. Ding. Transmission of linear regression patterns between time series: From relationship in time series to complex networks. *Physical Review E*, 90, 2014.

[28] X. Huang, H. An, X. Gao, X. Hao, and P. Liu. Multiresolution transmission of the correlation modes between bivariate time series based on complex network theory. *Physica a-Statistical Mechanics and Its Applications*, 2015.

[29] W. Long, L. Guan, J. Shen, and L. Cui. A complex network for studying the transmission mechanisms in stock market. *Physica A: Statistical Mechancis*, 2017.

[30] A.-L. Barabasi. *Network Science*. Cambridge University Press, 2016.

[31] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1), 2007.

[32] A. Vespignani. Twenty years of network science. *Nature*, 558, 2018.

[33] V. Silva, M. Silva, P. Ribeiro, and F. Silva. Time series analysis via network science: Concepts and algorithms. *Advanced Review: Data Mining and Kwowledge Discovery*, 2020.

[34] E. Bradley and H. Kantz. Nonlinear time-series analysis revisited. *Chaos*, 9, 2015.

[35] E. Ott. *Chaos in Dynamical Systems, second ed.* Cambridge University Press, 2002.

[36] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis, second ed.* Cambridge University Press, 2004.

[37] X. Yu, S. Shi, L. Xu, J. Yu, Y. Liu, and L. Wang. A directed limited penetrable visibility graph (dlpvg)-based method of analysing sea surface temperature. *Remote Sensing Letters*, 10(7), 2019.

[38] M. Wang and L. Tian. From time series to complex networks: The phase space coarse graining. *Physica A*, 461, 2016.

[39] N. Takeishi and T. Yairi. Anomaly detection from multivariate time-series with sparse representation. *in: Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on, San Diego, USA, IEEE*, 2014.

[40] J. Li, W. Pedrycz, and I. Jamal. Multivariate time series anomaly detection a framework of hidden markov models. *Applied Soft Computing*, 60, 2017.

[41] Marco Bardoscia, Paolo Barucca, Stefano Battiston, Fabio Caccioli, Giulio Cimini, Diego Garlaschelli, Fabio Saracco, Tiziano Squartini, and Guido

Caldarelli. The physics of financial networks. *Nat Rev Phys*, 5(3):490–507, 2021.

[42] Duc Thi Luu, Mauro Napoletano, Paolo Barucca, and Stefano Battiston. Collateral unchained: Rehypothecation networks, concentration and systemic effects. *Journal of Financial Stability*, 52:100811, 2021. Network models and stress testing for financial stability: the conference.

[43] X. Sun, W. Fang, X. Gao, S. An, S. Liu, and T. Wu. Time-varying causality inference of different nickel markets based on the convergent cross mapping method. *Resources Policy*, 74, 2021.

[44] A. Bovet, C. Campajola, F. Mottes, Valerio Restocchi, Nicolo Vallarano, Tiziano Squartini, and Claudio J Tessone. The evolving liaisons between the transaction networks of bitcoin and its price dynamics. *arXiv*, arXiv:1907.03577, 2019.

[45] X. Zhao, M. Ji, N. Zhang, and P. Shang. Permutation transition entropy measuring the dynamical complexity of financial time series. *Chaos, Solitons and Fractals*, 139, 2020.

[46] T. Kreuz, E. Satuvuori, M. Pofahl, and M. Mulansky. Leaders and followers: quantifying consistency in spatio-temporal propagation patterns. *New J. Phys.*, 19, 2017.

[47] A. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286, 1999.

[48] M. E. J. Newman and D. J. Watts. Renormalization group analysis of small-world network model. *Phys. Lettt. A*, 263, 1999.

[49] Z. k. Gao, M. Small, and J. Kurths. Complex network analysis of time series. *EPL*, 115, 2016.

[50] X. Xu, J. Zhang, and M. Small. Superfamily phenomena and motifs of networks induced from time series. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 2008.

[51] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. Carlos Nuno. From time series to complex networks: the visibility graph. *PNAS*, 2008.

[52] Z.-K. Gao, Q. Cai, Y.-X. Yang, and W.-D. Dang. Time-dependent limited penetrable visibility graph analysis of nonstationary time series. *Physica A*, 476, 2017.

[53] X. Li, M. Sun, C. Gao, D. Han, and M. Wang. The parametric modified limited penetrable visibility graph for constructing complex networks from time series. *Physica A*, 492, 2018.

[54] Z.-K. Gao, Q. Cai, Y.-X. Yang, W.-D. Dang, and S.-S. Zhang. Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series. *Science Reports*, 2016.

[55] L. Lacasa, B. Luque, and J. C. Nuno. The visibility graph: A new method for estimating the hurst exponent of fractional brownian motion. *Europhys. Lett.*, 2009.

[56] X. Dong, X. Gao, Z. Dong, H. An, and S. Liu. Network evolution analysis of nickel futures and the spot price linkage effect based on a distributed lag model. *International Journal of Modern Physics B*, 33(19), 2019.

[57] X. Gao, H. An, W. Fang, X. Huang, and H. Li. Characteristics of the transmission of autoregressive sub-patterns in financial time series. *Science Reports*, 2014.

[58] W. Wang, Y. Chen, L. Tian, S. Jiang, Z. Tian, and R. Du. Fluctuation behavior analysis of international crude oil and gasoline price based on complex network perspective. *Applied Energy*, 2016.

[59] X.X. Wang, L.-Y. Xu, J. Yu, H.-Y. Xu, and X. Yu. Detection of correlation characteristics between financial time series based on multi-resolution analysis. *Advanced Engineering Informatics*, 42, 2019.

[60] J. Zhang and M. Small. Complex network from pseudoperiodic time series: Topology versus dynamics. *Physical Review Letters*, 96, 2006.

[61] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths. Recurrence networks—a novel paradigm for nonlinear time series analysis. *New J. Phys.*, 12, 2010.

[62] J. Tang, F. Liu, W. Zhang, S. Zhang, and Y. Wang. Exploring dynamic property of traffic flow time series in multi-states based on complex networks: Phase space reconstruction versus visibility graph. *Physica A*, 450, 2016.

[63] M.C. Qian, Z.Q. Jiang, and W.X. Zhou. Universal and nonuniversal allometric scaling behaviors in the visibility graphs of world stock market indices. *J. Phys. A Math. Theoret.*, 43, 2009.

[64] L. Lacasa. On the degree distribution of horizontal visibility graphs associated to markov processes and dynamical systems: diagrammatic and variational approaches. *Nonlinearity*, 27, 2014.

[65] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque. Horizontal visibility graphs: exact results for random time series. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.*, 80, 2009.

[66] Z. Gao, N. Jin, W. Wang, and Y. Lai. Motif distributions in phase-space networks for characterizing experimental two-phase flow patterns with chaotic features. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.*, 82, 2010.

[67] Y. Yan, S. Zhang, J. Tang, and X. Wang. Understanding characteristics in multivariate traffic flow time series from complex network structure. *Physica A*, 477, 2017.

[68] M. Xu and M. Han. Adaptive elastic echo state network for multivariate time series prediction. *IEEE Transactions on Cybernetics*, 46, 2016.

[69] Z-K. Gao, Y-X. Yang, P-C. Fang, N-D. Jin, C-Y. Xia, and L-D. Hu. Multi-frequency complex network from time series for uncovering oil-water flow structure. *Scientific Reports*, 5, 2015.

[70] S. Lu, H. Zhang, X. Li, Y. Li, C. Niu, X. Yang, and D. Liu. Modeling the global ionospheric variations based on complex network. *Journal of Atmospheric and Solar-Terrestrial Physics*, 192, 2019.

[71] T. Tanizawa, T. Nakamura, and F. Taya. Directed networks with underlying time structures from multivariate time series. *arXiv preprint:1406.5247*, 2014.

[72] H. Li, H. Ren, H. An, N. Ma, and L. Yan. Multiplex cross-shareholding relations in the global oil & gas industry chain based on multilayer network modeling. *Energy Economics*, 95, 2021.

[73] T. Robinson. Have european gas prices converged? *Energy Policy*, 35, 2007.

[74] G. Zachmann. Electricity wholesale market pricesin europe: Convergence? *Energy Economics*, 30, 2008.

[75] Z. Banko and J. Abonyi. Correlation based dynamic time warping of multivariate time series. *Expert Systems with Applications*, 39, 2012.

[76] Y. Liu, X. Gao, and J. Guo. Network features of the eu carbon trade system: An evolutionary perspective. *energies*, 2018.

[77] I. Seabrook, P. Baricca, and F. Caccioli. Evaluating structural edge importance in temporal networks. *Preprint - UCL*, 2020.

[78] F. Wu, W.-L. Zhao, Q. Ji, and D. Zhang. Dependency, centrality and dynamic networks for internationalcommodity futures prices. *International Review of Economics & Finance*, 67, 2020.

[79] Z. Wang, X. Gao, R. Tang, X. Liu, Q. Sun, and Z. Chen. Identifying influential nodes based on fluctuation conduction network model. *Physica A*, 514, 2019.

[80] M. Starnini, M. Boguna, and M. A. Serrano. The interconnected wealth of nations shock propagation on global trade-investment multiplex networks. *Scientific Reports*, 2019.

[81] K. S. Tuncel and M. G. Baydogan. Autoregressive forests for multivariate time series modeling. *Pattern Recognition*, 73, 2018.

[82] Q. Huang, C. Zhao, X. Wang, X. Zhang, and D. Yi. Predicting the structural evolution of networks by applying multivariate time series. *Physica A: Statistical Mechancis*, 2015.

[83] E. Saggioro, J. Wiljes, M. Kretschmer, and J. Runge. Reconstructing regime-dependent causal relationships from observational time series. *Chaos*, 2020.

[84] V. Mwaffo, J. Keshavan, T. Hedrick, and S. Humbert. A data-driven method to dissect the dynamics of the causal influence in complex dynamical systems. *2018 IEEE Workshop on Complexity in Engineering (COMPENG)*, 2018.

[85] B. Wu and T. Duan. Nonlinear dynamics characteristic of risk contagion in financial market based on agent modeling and complex network. *Complexity*, 2019.

[86] F. Bashir, W. Qu, and A. Khokhar. Hmm-based motion recognition system using segmented pca. *Conference on Image*, 2005.

[87] J. Wang, C. Yang, R. Wang, H. Yu, Y. Cao, and J. Liu. Functional brain networks in alzheimer's disease: Eeg analysis based on limited penetrable visibility graph and phase space method. *Physica A*, 460, 2016.

[88] D. Tsiotas and A. Charakopoulos. Visibility in the topology of complex networks. *Physica A*, 505, 2018.

[89] Z.-K. Gao, X.-W. Zhang, N.-D. Jin, R. V. Donner, N. Marwan, and J. Kurths. Recurrence networks from multivariate signals for uncovering dynamic transitions of horizontal oil-water stratified flows. *EPL*, 103, 2013.

[90] T. Weng, J. Zhang, M. Small, R. Zheng, and P. Hui. Memory and betweenness preference in temporal networks induced from time series. *Sci. Rep.*, 2017.

[91] Q. Ji, E. Bouri, and D. Roubaud. Dynamic network of implied volatility transmission among us equities,strategic commodities, and brics equities. *International Review of Financial Analysis*, 57, 2018.

[92] Z.-K. Gao, Y.-X. Yang, P.-C. Fang, Y. Zou, C.-Y. Xia, and M. Du. Multiscale complex network for analyzing experimental multivariate time series. *EPL*, 109, 2015.

[93] X.-X. Wang, L.-Y. Xu, J. Yu, H.-Y. Xu, and X. Yu. Detection of correlation characteristics between nancial time series based on multi-resolution analysis. *Advanced Engineering Informatics*, 42, 2019.

[94] X. Liu and C. Jiang. The dynamic volatility transmission in the multiscale spillover network of the international stock market. *Physica A*, 560, 2020.

[95] G. Sui, H. Li, S. Feng, X. Liu, and M. Jiang. Correlations of stock price fluctuations under multi-scale and multi-threshold scenarios. *Physica A*, 490, 2018.

[96] Q. Zheng and L. Song. Dynamic contagion of systemic risks on global main equity markets based on granger causality networks. *Discrete Dynamics in Nature and Society*, 2018.

[97] S. Kumar and N. Deo. Correlation and network analysis of global financial indices. *Physical Review E*, 86, 2012.

[98] G. Castagneto-Gissey, M. Chavez, and F. De Vico Fallani. Dynamic granger-causal networks of electricity spot prices: A novel approach to market integration. *Energy Economics*, 44, 2014.

[99] M. Oskarsdottir, T. V. Calster, B. Baesens, W. Lemahieu, and J. Vanthienen. Time series for early churn detection: Using similarity based classification for dynamic networks. *Expert Systems with Applications*, 106, 2018.

[100] X. Yang, S. Chen, H. Liu, X. Yang, and C. Huang. Jump volatility spillover network based measurement of systemic importance of chinese financial institutions. *International Review of Finance & Economics*, 2021.

[101] X. Yang, X. Zhao, X. Gong, X. Yang, and C. Huang. Systemic importance of china's financial institutions: A jump volatility spillover network review. *Entropy*, 2020.

[102] Z. Wang, X. Gao, H. An, R. Tang, and Q. Sun. Identifying influential energy stocks based on spill over network. *International Review of Financial Analysis*, 68, 2020.

[103] P. Li and Z. Dong. Time-varying network analysis of fluctuations between crude oil and chinese and u.s. gold prices in different periods. *Resources Policy*, 68, 2020.

[104] Z. K. Gao, P. C. Fang, M. S. Ding, and N. D. Jin. Multivariate weighted complex network analysis for characterizing nonlinear dynamic behavior in two-phase flow. *Experimental Thermal and Fluid Science*, 60, 2015.

[105] Y. An, M. Sun, C. Gao, D. Han, and X. Li. Analysis of the impact of crude oil price fluctuations on china's stock market in different periods—based on time series network model. *Physica A*, 492, 2018.

[106] A. G. Asuero, A. Sayago, and A. G. Gonzalez. The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry*, 36, 2006.

[107] K. Friston, R. Moran, and A. K. Seth. Analysing connectivity with granger causality and dynamic causal modelling analysing connectivity with granger causality and dynamic causal modelling. *Current Opinion in Neurobiology*, 2013.

[108] P.-O. Amblard and O. J. J. Michel. The relation between granger causality and directed information theory: A review. *entropy*, 2012.

[109] R. Taylor. Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, 1990.

[110] S. Shaw. Perturbation techniques for nonlinear systems. *Encyclopedia of Vibration*, 2001.

[111] Takashi Isogai. Dynamic correlation network analysis of financial asset returns with network clustering. *Applied Network Science*, 2(8), 2017.

[112] B Podobnik and H. Stanley. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. *Phys Rev Lett.*, 100(8), 2008.

[113] Caterina Schiavoni, Siem Jan Koopman, Franz C. Palm, Stephan Smeekes, and Jan van den Brakel. Time-varying state correlations in state space models and their estimation via indirect inference. *Tinbergen Institute Discussion Paper*, 2021.

[114] J. Geweke and G. Amisano. Hierarchical markov normal mixture models with applications to financial asset returns. *J. Appl. Econ.*, 26:1–29, 2011.

[115] J. Benesty, J. Chen, Y. Yiteng, and I. Cohen. *Pearson Correlation Coefficient*. Springer Topics in Signal Processing, 2009.

[116] C.-Z. Yao and H.-Y. Li. Effective transfer entropy approach to information flow among epu, investor sentiment and stock market. *Front. Phys.*, 2020.

[117] C. Shannon. A mathematical theory of communication. *Bell Syst Tech J.*, 27, 1948.

[118] E. Siggiridou and D. Kugiumtzis. Granger causality in multivariate time series using a time-ordered restricted vector autoregressive model. *IEEE Transactions on Signal Processing*, 64, 2016.

[119] S. L. Bressler and A. K. Seth. Wiener-granger causality: A well established methodology. *NeuroImage*, 58, 2011.

[120] W. H. Greene. *Econometric Analysis*. Prentice-Hall Upper Saddle River, fifth edition, 2002.

[121] L. Barnett, A. D. Barrett, and A. K. Seth. Granger causality and transfer entropu are equivalent for gaussian variables. *Physical REview Letters2009*, 103.

[122] E. Pereda, R. Q. Quiroga, and J. Bhattacharya. Nonlinear multivariate analysis of neurophysological signals. *Progress in Neurobiology*, 77, 2005.

[123] Axel. Wismüller, Adora. M. DSouza, and Anas. Z. Abidin. Large-scale nonlinear granger causality: A data-driven, multivariate approach to recovering directed networks from short time-series data. *arXiv:2009.04681*, 2020.

[124] L. Wang, J. Zhang, Y. Zhang, R. Yan, H. Liu, and M. Qiu. Conditional granger causality analysis of effective connectivity during motor imagery and motor execution in stroke patients. *BioMed Research International*, 2016.

[125] Tomaso Aste and T. Di Matteo. Sparse causality network retrieval from short time series. *Complexity*, 2017, 2017.

[126] Reka. Albert and Albert-Laszlo. Barabasi. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(47), 2002.

[127] M. E. J. Newman. Networks: An introduction. *Oxford University Press*, 2011.

[128] A. Barabasi and E. Bonabeau. Scale-free networks. *Scientific American*, 288, 2003.

[129] S. H. Strogatz and D. J. Watts. Collective dynamics of 'small-world' networks. *Nature*, 393, 1998.

[130] E. M. Syczewska and Z. R. Struzik. Granger causality and transfer entropy for financial returns. *Proceedings of the 7th Symposium FENS*, 2015.

[131] P. K. Narayan, S. Sharma, W. C. Poon, and J. Westerlund. Do oil prices predict economic growth? new global evidence. *Energy Economics*, 41, 2014.

[132] B. N. Huang, M. J. Hwang, and H. P. Peng. The asymmetry of the impact of oil price shocks on economic activities: An application of the multivariate threshold model. *Energy Economics*, 27, 2005.

[133] M. Gisser and T. H. Goodwin. Crude oil and the macroeconomy: Tests of some popular notions: Note. *J. Money Credit Bank*, 18, 1986.

[134] Q. Zhang. The impact of international oil price fluctuation on china's economy. *Energy Proc.*, 5, 2011.

[135] Y. Wei and X. Guo. An empirical analysis of the relationship between oil prices and the chinese macro-economy. *Energy Economics*, 56, 2016.

[136] W. Tang, L. Wu, and Z. X. Zhang. Oil price shocks and their short- and long-term effects on the chinese economy. *Ssrn. Eletrcon. J.*, 32, 2009.

[137] L. Zhao, X. Zhang, S. Wang, and S. Xu. The effects of oil price shocks on output and inflation in china. *Energy Economics*, 53, 2014.

[138] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwanga. Complex networks: Structure and dynamics. *Science Reports*, 424, 2006.

[139] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *J. Stat. Mech.*, 2008.

[140] Edoardo Fadda, Junda He, Claudio J. Tessone, and Paolo Barucca. Consensus formation on heterogeneous networks. *EPJ Data Science*, 11(1):34, 2022.

[141] C. Gray, A.G. Hildrew, X. Lu, A. Ma, D. McElroy, D. Monteith, E. O'Gorman, E. Shilland, and G. Woodward. Chapter ten - recovery and nonrecovery of freshwater food webs from the effects of acidification. In Alex J. Dumbrell, Rebecca L. Kordas, and Guy Woodward, editors, *Large-Scale Ecology: Model Systems to Global Perspectives*, volume 55 of *Advances in Ecological Research*, pages 475–534. Academic Press, 2016.

[142] Xueke Lu, Clare Gray, Lee Brown, Mark Ledger, Alexander Milner, Raúl Mondragón, Guy Woodward, and Athen Ma. Drought rewires the cores of food webs. *Nature Climate Change*, 6, 05 2016.

[143] Isobel E. Seabrook, Paolo Barucca, and Fabio Caccioli. Evaluating structural edge importance in temporal networks. *EPJ Data Science*, 10(1):23, 2021.

[144] A. Ma and R.J. Mondragón. Evaluation of network robustness using a node tearing algorithm. *Physica A: Statistical Mechanics and its Applications*, 391(24):6674–6681, 2012.

[145] G. Bastille-Rousseau, I. Douglas-Hamilton, S. Blake, J. M. Northrup, and G. Wittemyer. Applying network theory to animal movements to identify properties of landscape space use. *Ecological Applications*, 28(3), 2018.

[146] J. Xiang, Y.-N. Tang, Y.-Y. Gao, Y. Zhang, K. Deng, X.K. Xu, and K. Hu. Multi-resolution community detection based on generalized self-loop rescaling strategy. *Physica A*, 2015.

[147] J. Xiang, T. Hu, Y. Zhangyan, and K. Hu. Local modularity for community detection in complex networks. *Physica A*, 2016.

[148] P. A. Gagniuc. *Markov Chains: From Theory to Implementation and Experimentation*. NJ: John Wiley & Sons, 2017.

[149] H. Moon and T.-C. Lu. Network catastrophe: Self-organized patterns reveal both the instability and the structure of complex networks. *Scientific Reports*, 2015.

[150] I. Guyon. Practical feature selection: from correlation to causality. *Mining massive data sets for security*, 2008.

[151] H.-P. Kriegel, P. Kroger, J. Sander, and A. Zimek. Density-based clustering. *Advanced Review*, 2011.

[152] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. Prakash. Patel, A. Tiwari, M. Joo, D. Weiping, and L. Chin-Teng. A review of clustering techniques and developments. *Neurocomputing*, 2017.

[153] M. Ester, H-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd ACM International Conference on Knowledge Discovery and Data Mining (KDD), Portland, OR,*, 1996.

[154] P.H.A. Sneath. The application of a data matrix. *J Am Stat Assoc*, 67, 1957.

[155] D. Wishart. Mode analysis: a generalization of nearest neighbor which reduces chaining effects. *In: Cole AJ, eds. Numerical Taxonomy, London and New York: Academic Press*, 1969.

[156] NetworkX Developers. networkx.algorithms.community.modularity_max.greedy_modularity. https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.modularity_max.greedy_modularity_communities.html, 08 2020.

[157] A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Phys. Rev. E*, 70, 2004.

[158] J. Park and P. S. C. Rao. Regime shifts under forcing of non-stationary attractors: Conceptual model and case studies in hydrologic systems. *Journal of Contaminant Hydrology*, 2014.

[159] T. A. Vuorenmaa and L. Wang. An agent-based model of the flash crash of may 6, 2010, with policy implications. *Available at SSRN: https://ssrn.com/abstract=2336772*, 2014.

[160] A. Golub, J. Keane, and S.-H. Poon. High frequency trading and mini flash crashes. *Available at SSRN: https://ssrn.com/abstract=2182097*, 2017.

[161] J. G. De Gooijer and R. J. Hyndman. 25 years of time series forecasting. *International Journal of Forecasting*, 22, 2006.

[162] R. Lewis and G. C. Reinsel. Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis*, 16, 1985.

[163] W. K. Li. Time series models based on generalized linear models: Some further results. *Biometrics*, 50, 1994.

[164] Z. Xu and X. Gou. An overview of interval-valued intuitionistic fuzzy information aggregations and applications. *Gran. Comput.*, 1, 2016.

[165] Jonathon. Shlens. A tutorial on principal component analysis. https://arxiv.org/abs/1404.1100, 2014.

[166] Supun Perera, Michael G. H. Bell, and Michiel C. J. Bliemer. Network science approach to modelling the topology and robustness of supply chain networks: a review and perspective. *Applied Network Science*, 2(1):33, 2017.

[167] Robert Goodell Brown. *Smoothing, Forecasting and Prediction of Discrete Time Series*. Dover Publications Inc., Robert Goodell Brown.

[168] Tao Wu, Xiangyun Gao, Sufang An, and Siyao Liu. Time-varying pattern causality inference in global stock markets. *International Review of Financial Analysis*, 77:101806, 2021.

[169] R. A. Fisher. Statistical methods for research workers. *Breakthroughs in Statistics*, 1992.

[170] S. Goodman. P-values, hypothesis tests and likelihood: Implications for epidemiology of a neglected historical debate. *American Journal of Epidemiology*, 137, 1993.

[171] S. Goodman. A dirty dozen: Twelve p-value misconceptions. *Seminars in Hematology*, 45, 2008.

[172] John Fox. *Applied regression analysis, linear models, and related methods*. Sage Publications, 1997.

[173] S. R. Searle. *Linear Models*. Wiley, 1997.

[174] P. Wegner. A technique for counting ones in a binary computer. *Communications of the ACM*, 3(5), 1960.

[175] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths. Ambiguities in recurrence-based complex network representations of time series. *Phys. Rev. E*, 81, 2010.

[176] F. Taken. Detecting strange attractors in turbulence. *Lecture Notes in Math*, 898, 1981.

[177] Gregor Klančar, Andrej Zdešar, Sašo Blažič, and Igor Škrjanc. Chapter 4 - path planning. In Gregor Klančar, Andrej Zdešar, Sašo Blažič, and Igor Škrjanc, editors, *Wheeled Mobile Robotics*, pages 161–206. Butterworth-Heinemann, 2017.

[178] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. Optics: Ordering points to identify the clustering structure. *ACM Sigmod record*, 28(2), 1999.

[179] L Ertoz, M Steinbach, and V Kumar. Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data. *Proceedings of the 3rd SIAM International Conference on Data Mining (SDM)*, 2003.

[180] SciPy. scipy.signal.find_peaks. https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal 17/09/2020.

[181] C. Rick. Efficient computation of all longest common subsequences. *Lecture notes in computer science*, 1851, 2002.

[182] V Michail, H Marios, and G Dimitrios. Indexing multidimensional time-series. *Int J Very Large Data Bases*, 15, 2006.

[183] G. Yuan, P. Sun, J. Zhao, D. Li, and C. Wang. A review of moving object trajectory clustering algorithms. *Artificial Intelligence Review*, 47, 2017.

[184] J.Y. Chen, R.D. Wang, L.X. Liu, and J.T. Song. Clustering of trajectories based on hausdorff distance. *In: Proceedings of the 2011 international conference on electronics, communications and control*, 2011.

[185] L.X. Wei, X.H. He, Q.Z. Teng, and M.L. Gao. Trajectory classification based on hausdorff distance and longest common subsequence. *J Electron Inf Technol*, 2013.