Theoretische Physik

Understanding Complexity in Socio-Economic Systems at Various Time Scales: Data-Driven and Modelling Approaches

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Physik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Münster

vorgelegt von

Tobias Wand

aus Datteln

2024

Dekan	Prof. Dr. Rudolf Bratschitsch
Erster Gutachter	Prof. Dr. Uwe Thiele
Zweite Gutachterin	Prof. Dr. Svetlana Gurevich
Tag der mündlichen Prüfung	_____
Tag der Promotion	_____

Abstract

Complexity, collective behaviour and self-organisation are found in various systems in real life. Socio-economic systems are no exception from this and show the emergence of macroscopic phenomena that can not be trivially derived from the behaviour of human beings as their microscopic constituents. Socio- and econophysics have, for example, helped to understand how crashes appear on financial markets because of the individual decisions of traders and how opinions can drastically shift in a network of humans. Additionally, modern access to data and statistical methods provide numerous opportunities to analyse socio-economic phenomena quantitatively.

This thesis analyses complex socio-economic systems on various time scales from high frequency trading on financial markets to the long-term growth of socio-economic complexity in pre-modern societies: Algebraic extensions of the geometric Brownian motion for asset prices are shown to give rise to stable equilibria in Langevin equations for high frequency and daily price dynamics. The daily market correlation is found to be affected by memory effects and unequal contributions of the various business sectors to the collective dynamics. The Granger causality network between daily sector returns exhibits a strong hierarchical structure as identified via the Helmholtz-Hodge-Kodaira decomposition. An agent-based model of non-ergodic wealth dynamics shows the slow emergence of cooperation clusters from simple growth rate optimisation of each agent. These clusters resemble the kin selection mechanism from network attachment rules, but emerge endogenously as large-scale structures over long periods of time. Analysis of the Seshat databank on socio-economic complexity in pre-modern societies shows that a survivorship bias needs to be taken into account for resilience analyses of historical data. Finally, the development of socio-economic complexity in pre-modern states is shown to follow a universal trajectory with a characteristic time scale of two and a half millennia.

Kurzfassung

Komplexität, kollektives Verhalten und Selbstorganisation treten in verschiedenen realen Systemen auf. Sozioökonomische Systeme sind hiervon keine Ausnahme und zeigen die Emergenz von makroskopischen Phänomenen, die sich nicht trivial aus dem Verhalten der einzelnen Menschen herleiten lassen. Die Sozio- und Ökonophysik hat beispielsweise dazu beigetragen zu verstehen, wie Krisen im Finanzmarkt wegen des individuellen Verhaltens der Händler plötzlich entstehen und wie Meinungen sich in einem sozialen Netzwerk in kurzer Zeit ändern können. Zudem können sozioökonomische Phänomene dank moderner Datenquellen und statistischer Methoden nun auch quantitativ besser erforscht werden.

Diese Dissertation untersucht sozio-ökonomische dynamische Prozesse auf verschiedenen Zeitskalen vom Hochfrequenzhandel der Finanzmärkte zu der langsamen Herausbildung von komplexen Gesellschaften in der Menschheitsgeschichte: Algebraische Erweiterungen der geometrischen Brownschen Bewegung als Langevin-Gleichung für Hochfrequenz- und Tagesschlusskurse von Aktienpreisen weisen stabile Fixpunkte auf. Die tägliche Marktkorrelation zeigt Gedächtniseffekte im Kollektivverhalten des Markts und wird von verschiedenen Geschäftssektoren in ungleichem Maße beeinflusst. Für das Netzwerk von Granger-Kausalität zwischen den Renditezeitreihen von Industriesektoren identifiziert die Helmholtz-Hodge-Kodaira-Deokomposition eine stark hierarchisch geprägte Struktur. Ein agentenbasiertes Modell von nichtergodischer Wohlstandsdynamik zeigt die langsame Ausbildung von Kooperationsregionen, indem jeder Agent seine Wachstumsrate optimiert. Diese Regionen erinnern an *kin selection* als Mechanismus der Netzwerkentstehung, bilden sich aber endogen als Strukturen über lange Zeitperioden heraus. Anhand der Seshat-Datenbank, einem interdisziplinären Datenprojekt zur sozioökonomischen Komplexität in vormodernen Gesellschaften, wird gezeigt, dass bei der Resilienzanalyse von historischen Daten ein möglicher *survivorship bias* berücksichtigt werden muss. Zuletzt wird in der Entwicklung von sozioökonimischer Komplexität in vormodernen Staaten eine universelle Zeitskala von zweieinhalb Jahrtausenden identifiziert.

Contents

Part I

Introduction

1 Complexity Science, Econophysics and Sociophysics

"Social physics is that science which occupies itself with social phenomena, considered in the same light as astronomical, physical, chemical, and physiological phenomena, that is to say as being subject to natural and invariable laws, the discovery of which is the special object of its researches."

— Auguste Comte, Physique sociale - Cours de philosophie positive

Perhaps the fundamental feature of a complex system is that the whole is more than the sum of its parts [1, 2]. The behaviour of the whole system is not replicated by enlarging the isolated behaviour of its microscopic constituents but emerges from the interactions between them. As an example, consider the microscopic cells in a human body and the complexity of human biology which cannot be trivially derived from the fundamental cells. Complex systems can experience macroscopic phase transitions as qualitatively different behaviour after a small change in a control parameter. Perturbations of a complex system may show the resilience of the system as it returns to its initial equilibrium state or lead to feedback loops with increasing strength which requires the study of far-from-equilibrium dynamics. Some systems behave chaotically as small differences in initial conditions grow exponentially, making the long-term behaviour impossible to predict despite the deterministic nature of the system. Macroscopic spatio-temporal patterns can emerge whose length scales far exceed the size of the microscopic constituents and whose self-similarity can encompass various orders of magnitude. And most of the systems in nature which show self-organised behaviour are open systems and coupled to an environment with which they interact with and may adapt to external stimuli. Examples of these various phenomena include the Ising model where microscopic spin-spin interactions lead to macroscopic magnetisation and responses to external magnetic fields [3, 4], synchronised dynamics in the Kuramoto oscillator model [5], the superconductivity phase transition [6] or the synchronised flashes of fireflies [1]. Multiple length scales are

found in the development of turbulence [7] or the combination of many microcracks to a large fracture and material failure [8]. And systems like the Brownian particle influenced by the fast microscopic thermal fluctuations, the cooling of metals via rapid quenching or slow annealing or the slow climatic changes coupled to the much faster change in weather can be analysed by taking the different scales into account at which dynamics take place [9, 10, 11]. One can either regard the slowest dynamics as approximately constant from the point of view of the fastest degrees of freedom or, vice versa, the fastest dynamics approximated as noise. These approaches can be extended to multiscale phenomena in solid state physics and fluid dynamics [12]. In all of these cases, the superposition principle does not hold and the combination of two possible states of the system is not necessarily a valid state. Mathematically, this is described by the notion of nonlinearity that a function f does not fulfil $f(\alpha x + y) = \alpha f(x) + f(y)$. The tools and methods of complexity science have found application in many other disciplines as it became evident that many fields of research are dealing with complex systems [1].

One such example is the set of socio-economic relationships which characterise human cooperation, competition and coexistence. Using methods and concepts from quantitative natural sciences has become more widespread due to the increasing availability of data in the social and economic sciences, but is not a new idea. In fact, the founder of modern sociology August Comte explicitly envisioned this discipline as "social physics" and argued that social phenomena are subject to mathematical laws which can be uncovered by researchers [13]. In particular, the Austrian school of economics has focused on the complexity of the economy and society while emphasising the emergent socio-economic structure that arose from free decision-making. As argued by the economist Hayek, knowledge is dispersed among the many individual agents who participate in the economy and no single entity or planner can organise the whole breadth of knowledge on their own [14]. Instead, if the agents are free to participate in the price discovery process, the price can condense the dispersed knowledge about a product's scarcity, offer and demand into a single number [15, 16]. This school of thought naturally connects economic theory to the emergent properties studied in complexity science.

This thesis analyses complex human socio-economic systems at different time scales with methods from physics. The broader context of this thesis will be further explained in the remainder of this chapter with a focus on the complexity of financial markets. Additionally, a brief introduction to the quantitative analysis of social complexity in pre-modern societies in the field of cliodynamics will be given in this chapter and the statistical methods used throughout this thesis will be explained in chapter 2. The majority of the research in this thesis is focused on the financial markets whose price dynamics are

analysed on daily and high frequency intraday time scales alongside the analysis of daily correlation and causality between financial time series in part II. The agent-based simulation in part III explains how the microscopic optimisation of individual agents can give rise to large-scale cooperation clusters. The dynamics of states and cultures as the largest forms of social complex systems is analysed in part IV and it is shown that universal statistical laws of their development can be found in the data. Finally, the conclusion in part V gives a brief summary of the main results of each chapter and discusses the broader implications of this work.

1.1 The Financial Market as a Complex System

One of the most widely studied example of a complex socio-economic system is the financial market. Financial markets are markets where financial assets can be traded. In the most general formulation, assets are a "legal claim to a future cash flow" and include, e.g. shares of a company, bonds or derivative contracts [17]. Note that many methods developed for asset prices can also be adapted to commodities such as gold or foreign exchange markets. Whereas new assets are issued on primary markets (e.g. when a company issues its shares for the first time to selected investors), they are traded between buyers and sellers on secondary markets (i.e. stock markets) where usually so-called market-makers provide liquidity. Most research focuses on secondary markets because of their symmetric balance between buyers and sellers and the many microscopic traders contributing to the macroscopic observable - the asset's price - are a perfect example of a complex system [17]: Financial markets experience feedback loops from the traders' expectations and actions, the price time series are highly non-stationary, many interacting agents contribute to the global dynamics by adapting their behaviour, leading to an evolution of the agent population. Moreover, the observed system is only a single realisation of all possible market outcomes, thereby crucially requiring considerations about the non-ergodicity of the system and the reproducibility of its behaviour, which is additionally not a closed system, but coupled to the environment. Emergent behaviour observed in such markets include booms and bubbles (such as the famous dot-com bubble during the late 1990s) as well as crashes and crises (like the 2007/08 financial crisis) which cannot be directly understood from the microscopic behaviour of individual traders alone [16, 17, 18, 19, 20]. The high-frequency trades at the stock exchange ultimately depend on the slow development of the real economy and are "enslaved" by them, i.e. they quickly react to changes in the real economy similar to how the weather depends on the slow changes of the climate [11, 21, 22]. Perhaps most importantly, the roughly

identical profit-seeking behaviour of the individual traders leads to the prices reflecting all available information. This motivates the efficient market hypothesis (EMH) that the market is (approximately) arbitrage free and hence, no risk-free profit can be gained by consistently beating the market [23].
Insights and methodologies from physics and complexity science are crucial to study financial markets as physicists can provide numerous time series analysis and modelling methods as well as a thorough connection between the microscopic and macroscopic world. Deriving a description for the macroscopic system observable from its microscopic constituents leads to better understanding of how long memory can persist in financial markets [24] and has been empirically refined by taking the distribution of traders into account [25]. Such an approach is similar in spirit to the Ising model which explains the different magnetic behaviour on the macroscopic level of the full system by deriving it from its microscopic spins [3, 4]. Several other physics-inspired methods with a focus on time series analysis will be presented in the remainder of this introductory section.

1.2 A Primer on Quantitative Finance

The fundamental observable for each asset is its price P_t at time t and its time series $(P_1, P_2, \dots) = (P_t)_{t\in\mathbb{N}}$ or shortened as $(P_t)_t$ is given mathematically by

$$P\colon \mathbb{N} \to \mathbb{R},\ t \mapsto P_t. \tag{1.1}$$

Price time series are used by, e.g. models and applications based on the geometric Brownian motion (GBM) such as the influential Black-Scholes model for derivative pricing [26, 27, 28]. Nevertheless, most analyses are instead using the returns $R_t = (P_{t+1} - P_t)/P_t$ for two reasons. First, they measure relative changes and therefore returns from two assets with vastly different absolute prices can still be compared to each other, e.g. when constructing a portfolio. Second, the returns usually have a zero mean and an approximately stationary distribution which makes them more suitable for many statistical analysis tools [18]. Additionally, one often finds that the logarithmic returns $\mathcal{R}_t = \log(P_{t+1}/P_t)$ are used. Because $|R_t| \ll 1$ usually holds,

$$\log\left(\frac{P_{t+1}}{P_t}\right) = \log\left(\frac{P_{t+1} - P_t + P_t}{P_t}\right) = \log(1 + R_t) \approx R_t \tag{1.2}$$

and therefore R_t can be approximated by $\mathcal{R}_t$. However, while R_t is per definition confined to the asymmetric interval $[-1, \infty)$, the logarithmic returns are defined on the full range

of $\mathbb{R}$ which might be desired for some statistical models. As $\mathbb{E}[R_t] \approx 0$ is usually valid, the variance of the returns is often approximated as

$$\sigma^2 = \mathbb{V}(R_t) \approx \mathbb{E}[R_t^2] \tag{1.3}$$

and in financial vocabulary, the standard deviation σ of a return time series is usually called the volatility. For a formal definition of $\mathbb{E}$ and $\mathbb{V}$, the reader is referred to chapter 2.2.1.

The statistical properties of (logarithmic) returns show some general properties that manifest themselves across almost all assets and financial markets. These empirical observations are known as stylised facts and have motivated numerous directions of research as they cannot be reproduced by simple random walk models for asset prices [16, 17, 18]. First, the return distributions are non-Gaussian with fat tails, i.e. extreme events occur much more frequently than a Gaussian model would expect. If the returns are calculated based on longer time steps (e.g. days instead of seconds), their distribution becomes increasingly Gaussian which is referred to as aggregational Gaussianity. Johnson, Jeffries and Hui speculate in [17] that the non-Gaussianity arises because convergence to a normal distribution according to the central limit theorem is violated since the smallest time unit is that of an individual trade execution. Hence, in any given time interval Δt, there are a finite number of infinitesimal time steps (i.e. trades) and Δt is not infinitely divisible. While fat-tailed distributions such as Student's t distribution can be suitable alternatives [16], searching for accurate distribution to characterise the returns for extreme events such as during a crisis is still an ongoing field of research. Nevertheless, the Gaussian distribution is frequently used for convenience because of its useful mathematical properties.

Analysis of the tails of the return distribution uncovered scaling behaviour and power laws in the empirical data [17, 29]. Symmetric Levy stable distributions with parameter $0 < \alpha < 2$ can be used to model the return distributions because of their power-law distribution

$$p_\alpha(x) \sim |x|^{-(1+\alpha)} \tag{1.4}$$

for large deviations $|x| \gg 0$ as shown in figure 1.1. Note that $\alpha = 2$ corresponds to a Gaussian distribution which, unlike the Levy distributions, has finite variance. Hence, truncated Levy distributions can be used to model finite volatility. Because Levy distributions are stable under convolutions like Gaussians (i.e. the sum of two random variables with a Levy or Gaussian distribution is again Levy or Gaussian), if a return at a given time step follows a Levy distribution, it will remain Levy distributed on all larger

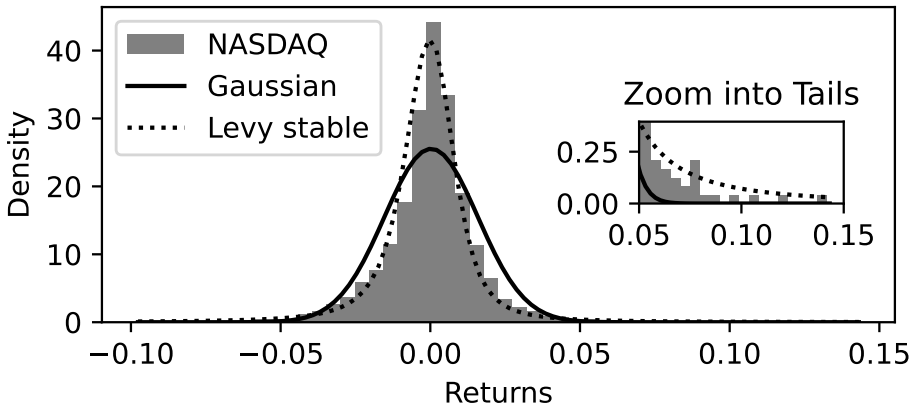

Figure 1.1: The daily returns of the NASDAQ Composite index from 2000 to 2020 are fitted with a Gaussian and a Levy stable distribution. As shown in the inset, the Gaussian distribution drastically underestimates the probability of observing extreme events in the tails of the empirical distribution.

time scales. Chapter 3 of [17] contains a pedagogical review of the Levy distribution for empirical data and, in line with [29], estimates $\alpha \approx 1.4$ and uses this value to scale the data on all time scales Δt. The scaled empirical distributions collapse to an identical shape, thereby indicating the self-similarity of the data. Such power laws $f(x) = ax^{-\gamma}$ and the lack of a characteristic scale due to the scaling law

$$f(cx) = a(cx)^{-\gamma} \sim ax^{-\gamma} = f(x) \tag{1.5}$$

are often found at self-organised criticality and indicate the onset of a phase transition [18]. Moreover, the fractal shape of many financial time series also gives rise to scaling laws [30].

Another stylised fact concerns the autocorrelation of returns: In liquid markets, returns usually have an autocorrelation function that quickly decays to zero, i.e. their memory seems to vanish very quickly which is in line with the EMH. However, nonlinear transformations of the returns show long-range autocorrelations, e.g. for the absolute $(|R_t|)_t$ or squared $(R_t^2)_t$ returns. Because of the connection between the squared returns and the volatility in equation (1.3), this phenomenon is usually referred to as volatility clustering and implies that a highly volatile or calm market will, respectively, stay volatile or calm in the immediate future. This behaviour can be measured empirically by estimating the volatility on moving windows that slide across the full time series.

1.3 Further Examples from Econophysics

Accurately predicting financial time series is one of the main challenges in quantitative finance. Statistical methods from time series analysis such as the moving average (MA) or autoregressive (AR) models or a combination of them (ARMA) are often fitted to return time series [16]. Extensions such as the (generalised) autoregressive conditional heteroscedasticity methods ARCH and GARCH explicitly model the volatility σ of the observed return process and therefore manage to replicate the volatility clustering observed in real data [18]. An alternative modelling technique is to use stochastic differential equations to describe the price movements. The famous Brownian motion

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \mu + \sigma P \epsilon \tag{1.6}$$

with standard Gaussian noise $\epsilon = \epsilon(t)$ was originally devised to model a particle's motion and has already one century ago been suggested as a tool to model stock prices P with a drift μ and fluctuations of scale σ [26]. Its extension to the geometric Brownian motion

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \mu P + \sigma P \epsilon \tag{1.7}$$

takes the non-negativity of prices into account and has been used in many more advanced quantitative finance models such as the famous Black-Scholes equation [27, 28]. Figure 1.2 shows a comparison between the BM and GBM process. Chapter 3 will further expand on the geometric Brownian motion and empirically analyse the possibility of additional deterministic terms in equation (1.7) with respect to the stability of fixed points in the price time series.

For convenience, many time series analysis methods usually assume that ensemble averages and time averages converge to the same value for large ensembles or time periods, i.e. that they are ergodic. Recently, physicists have raised the point that the market is always just a single realisation without any parallel universes and that combining this with the multiplicative growth of financial time series requires non-ergodic effects to be taken into account [31]. Long-standing economic puzzles such as the insurance paradox (Why do people sign insurance contracts if both parties expect to profit at the expense of the other?) and the St Petersburg paradox (a gamble with a seemingly infinite payoff which is nevertheless disliked by gamblers) could be resolved by careful distinction between ensemble and time average optimisation of economic agents [32, 33].

Usually, a single stock is not analysed in an isolated setting, but its co-movement with other assets has to be considered in order to create a well-balanced portfolio where risks

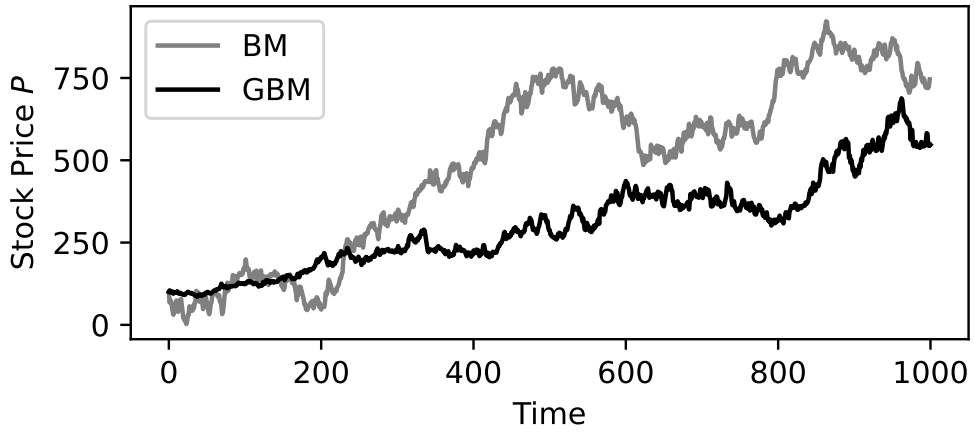

Figure 1.2: Simulation of a BM and a GBM process. Though both plots look similar, it can be noticed that the volatility is constant for the BM, but scales with the process's value for the GBM.

might cancel out each other. Correlation matrices have already been used to estimate such risk and insights from nuclear physics, where random matrix theory (RMT) has been used to model the spectra of heavy nuclei, have been applied to study financial correlation matrices [34, 35]. While most of the eigenvalues lie within the spectrum predicted for a purely random correlation matrix of uncorrelated assets, some eigenvalues lie far outside of this bulk and therefore represent structures in the market data. Their existence is especially remarkable because RMT implies that the eigenvalues should be distributed according to the Marchenko-Pastur distribution whose upper limit λ_+ should not be exceeded by any of the eigenvalues. Yet, the empirical studies in [34, 35] show that financial correlation matrices have eigenvalues $\lambda > \lambda_+$ whose corresponding eigenvectors help to interpret them: The largest $\lambda_{\max}$ often lies several orders of magnitude outside of the range predicted by RMT and its eigenvector's entries are almost constant across all assets. Hence, it represents a portfolio which is equally invested in the entire market and has thus been called the market mode. The other eigenvalues above λ_+ similarly represent individual business sectors as their eigenvectors are strongly focused on, e.g. banking or industry stocks. Figure 1.3 schematically depicts the empirical and theoretical eigenvalue distributions. Chapter 4 expands on the research on financial correlation matrices in [36] and uses an algorithm from the field of explainable artificial intelligence to understand correlation matrices which correspond to different states of the market.

Networks are a modelling approach that is suitable for different fields of research, including many socio-economic systems [37, 38]. The networks of various real systems show a

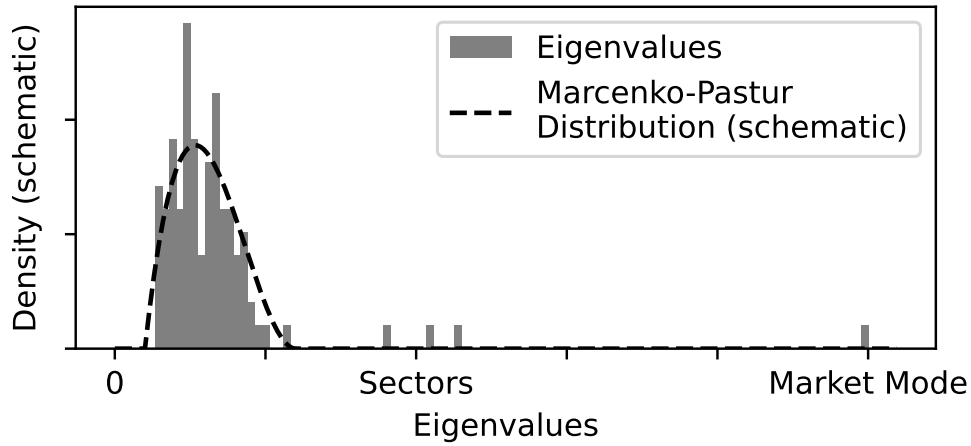

Figure 1.3: Sketch of the empirical eigenvalue spectrum of financial correlation matrix against the Marcenko-Pastur distribution for random matrices (schematic). The eigenvalues outside of the range of possible eigenvalues according to Marcenko-Pastur usually correspond to business sectors or the market mode. Note that the Marcenko-Pastur distribution is exactly 0 outside of the distribution's bulk and not just approximately 0.

scale-free topology (i.e. the structure of the network is self-similar at every length scale) and a power law tail in the degree distribution [39]. Additionally, methods to randomly generate realistic large networks have been found to display a phase transition from isolated small clusters to a giant component including almost all nodes [40]. Applications of network methods in econophysics include spanning trees to map the similarity between different financial time series [41] and understanding contagion effects during financial crises [42]. Because correlation does not imply causation, chapter 5 goes beyond the study of correlations in chapter 4 and estimates the network of Granger causality between financial time series. This network is further analysed via the Helmholtz-Hodge-Kodaira decomposition which discretises the Helmholtz decomposition from electric field theory to identify a potential and an associated hierarchy in the network [43, 44].

Network models, as the term "social network" suggests, naturally lend themselves to study social effects beyond economics and finance. For binary opinions such as the preference between two large political parties, Ising-like models have been used by physicists with different networks and "temperature" to reflect the social connections and external influences in the opinion formation [45, 46]. Inspired by this research, chapter 6 simulates an agent-based model on a lattice in which the agents can cooperate in the face of stochastic fluctuations of their wealth.

Socio-economic processes operate on vastly different time scales: Quantitative finance includes both high frequency trading with a sub-seconds frequency as well as long-term investments where portfolio adjustments are done on a daily basis or even less frequently. Coupled to the financial markets is the real economy with its annual seasonality as well as even slower cycles of economic booms and crises. From complex systems research, Haken's slaving principle gives us a methodology to analyse these different processes by assuming that the fast time scales adapt quickly to the slow-moving processes which, from the faster perspective, behave as an approximately constant and static environment [21, 22].

Underneath the economic cycles, the number of potential workers and consumers is subject to the slow demographic changes with a characteristic time scale of several decades. In human societies, population growth certainly depends on the economic development and in turn provides the human resources upon which future economic endeavours are built. The new field of cliodynamics has emerged as an effort to study structural-demographic patterns on this longest time scale in order to better understand the long-term fate of societies [47]. Starting with differential equations motivated from theoretical biology to describe human demography [48], cliodynamics has recently assembled the *Seshat: Global History Databank* on social complexity across several millennia and from all over the world in an interdisciplinary research project by combining domain expertise with statistical tools like multiple imputation [49]. The time series data from Seshat allows researchers to empirically test hypotheses on the development of social complexity [50, 51, 52], but as data records over such long time periods are scarce, data analysis is often supplemented with models from dynamical systems research [53, 54]. Chapters 7 and 8 use the Seshat data to analyse the resilience of such pre-modern states and to derive a characteristic time scale of their development from low to high social complexity.

As this brief introduction to socio-economic complexity shows, researchers in econophysics and sociophysics apply tools from complex systems science on all time scales of human interaction, from the high frequency data in financial markets to the centuries-long dynamics of demographic change.

2 Methods

"Data alone has no value — it's just masses of numbers or words."

— Steven J. Bowen, Total Value Optimization

Usually, nonlinear dynamical systems are difficult or outright impossible to solve analytically. Hence, mathematical tools like the fixed point analysis have been developed to understand these systems even if a full solution is not available [1]. Additionally, data-driven methods have become increasingly popular among scientists due to the growing computational power and data availability [55]. A brief introduction to the important methods and tools used in this dissertation will be given below with more technical details being reserved for the individual chapters or the appendix.

Some excerpts of this section are taken from the methods section of the author's publications listed in appendix A. In particular, section 2.2.5 is based on *Tobias Wand, Introduction to Artificial Intelligence, in Michael te Vrugt (Ed.), Artificial Intelligence and Intelligent Matter, Springer* (forthcoming in 2025 and reproduced with permission from Springer Nature) [56]. The figures in section 2.2.5 are adopted from [56] and were created by Tobias Wand.

2.1 Fixed Point Analysis

An ordinary differential equation (ODE) describes the dynamic behaviour of a system in continuous time via $\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x} = f(x)$. If we wish to analyse the long-term behaviour of an ODE, explicitly solving it analytically might unfortunately not be possible. Numerical solutions are of course an alternative, but even without computational power, it can be possible to infer the long-term behaviour analytically because a fixed point x^* with the property $f(x^*) = 0$ (i.e. no change in the dynamics) is often much easier to find than an explicit solution of the ODE. The following explanation of the fixed point analysis is mainly derived from [1].

For a one-dimensional system, assume the system is in a fixed point x^* of the ODE and that there is a small perturbation $\varepsilon(t) = \varepsilon$

$$x(t) = x^* + \varepsilon \tag{2.1}$$

with $|\varepsilon|$ small compared to the typical scale of the system. The resulting change of $x(t)$ is then given by $\dot{x} = f(x)$. This can be approximated via linearisation by using that $f(x^*) = 0$ per definition. Hence,

$$\dot{x} = f(x(t)) = f(x^* + \varepsilon) = f(x^*) + f'(x^*) \cdot \varepsilon + \mathcal{O}\left(\varepsilon^2\right) \approx 0 + f'(x^*) \cdot \varepsilon \tag{2.2}$$

with $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}$. In the vicinity of x^*, the ODE is therefore approximately described by $\dot{x} \approx f'(x^*)\varepsilon$. As x^* is a constant and $\varepsilon(t) = x(t) - x^*$, this results in the exponential ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(x(t) - x^*\right) = \frac{\mathrm{d}}{\mathrm{d}t}x(t) - 0 = f(x(t)) \approx f'(x^*) \cdot \varepsilon(t) \tag{2.3}$$

for the perturbation $\varepsilon(t)$ which will decay or grow depending on the sign of $f'(x^*)$. If $f'(x^*) < 0$, we have now inferred that any small perturbation will relax to the fixed point x^* and that the fixed point is stable, whereas for $f'(x^*) > 0$, any perturbation will grow and drive the system away from the fixed point (x either diverges or reaches another fixed point). Importantly, this system behaviour can be inferred without knowing the solution of the ODE.

Fixed point analysis can also be generalised to higher dimensions with $\dot{x}_i = f_i(\mathbf{x})$ and $\mathbf{x} = (x_1, \ldots, x_m)$. Linearisation around a fixed point $\mathbf{x}^* = (x_1^*, \ldots, x_m^*)$ is now done via the Jacobian

$$J_f(\mathbf{x}) = \left(\frac{\partial f_i}{\partial x_j}(\mathbf{x})\right)_{1 \leq i,j \leq m}. \tag{2.4}$$

The eigenvalues λ_i of $J_f(\mathbf{x}^*)$ are then calculated and have the property that the corresponding eigenvectors $\mathbf{v}_i$ fulfil $J_f(\mathbf{x}^*) \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i$. Because the eigenvectors usually form a basis, a decomposition is possible such that the perturbation can be expressed as $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \ldots, \varepsilon_m(t)) = \sum_i \alpha_i(t)\mathbf{v}_i$. Hence, the perturbation will develop like

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\varepsilon} = J_f(\mathbf{x}^*) \cdot \boldsymbol{\varepsilon} = J_f(\mathbf{x}^*) \sum_i \alpha_i(t)\mathbf{v}_i = \sum_i \alpha_i(t) J_f(\mathbf{x}^*)\mathbf{v}_i = \sum_i \alpha_i(t)\lambda_i \mathbf{v}_i. \tag{2.5}$$

On the other hand, applying the time derivative to the $\alpha_i(t)$ leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_i \alpha_i(t)\mathbf{v}_i = \sum_i \left(\frac{\mathrm{d}}{\mathrm{d}t}\alpha_i(t) \right) \mathbf{v}_i.$$

Sorting according to the axes spanned by the eigenvector basis, one can see that the perturbation's coefficients will develop as

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha_i(t) = \lambda_i \alpha_i(t) \tag{2.6}$$

and therefore again show an exponential growth or decline depending on the signs of the eigenvalues. Most notably, if all eigenvalues are negative, the perturbation will decay back to $\mathbf{x}^*$ and the system is locally stable. Additionally, complex eigenvalues result in an oscillatory $e^{i\omega t}$ and therefore correspond to oscillations in the phase space around $\mathbf{x}^*$.

2.2 Data Analysis

Ever since the advent of big data and new advances in machine learning, data-driven analysis has established itself at the forefront of science [57]. But even before these recent developments, statistical analysis tools have been an integral part of complex systems research [58, 59, 60]. This section will introduce the reader to some of the statistical fundamentals and basic concepts of machine learning. More details on the used methods will be given in the respective chapters of this thesis or in the appendix, as this section only serves as a brief introduction and revision to the interested reader with only basic experience in statistics. For a fundamental introduction to probability theory, the interested reader is referred to [61].

2.2.1 Nomenclature of Probability Theory

Let X denote a random variable and x the realised values that it can take. For a discrete random variable, the number of possible realisations is countable $x_1, x_2, \ldots$ and

each will be taken with probability $p_i = \mathbb{P}(X = x_i) \geq 0$ such that $\sum_i p_i = 1$. For a continuous random variable, the realisations are distributed according to a density $\rho(x) \geq 0$ with $\int \rho(x)\,\mathrm{d}x = 1$. The expectation value $\mathbb{E}[X]$ is defined as

$$\mathbb{E}[X] = \sum_i p_i x_i \qquad \text{or} \qquad \mathbb{E}[X] = \int x\rho(x)\,\mathrm{d}x \tag{2.7}$$

for the discrete and continuous case, respectively. For most random variables, the expectation value indicates a typical realisation value and the sample mean converges to the expectation value. The variance is defined as

$$\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \tag{2.8}$$

and indicates the spread of the realisations. Typically, the variance is expressed as the square of the standard deviation $\mathbb{V}(X) = \sigma_X^2$. For two random variables X and Y, the Pearson correlation

$$\rho_{X,Y} = \frac{(\mathbb{E}[X - \mathbb{E}[X]])\,(\mathbb{E}[Y - \mathbb{E}[Y]])}{\sigma_X \sigma_Y} \tag{2.9}$$

lies in the interval $[-1, 1]$ and indicates how much the measurements of X and Y depend on each other. The correlation of a random variable with itself is always 1, but it can be interesting to see how much a series of measurements of the same variable changes over time. The autocorrelation function ACF given by

$$ACF(r) = \rho_{X_t, X_{t+r}} \tag{2.10}$$

computes this by comparing the measurements of X at time t with those at time $t + r$ for all possible values of t. Note that $ACF(0) = 1$ per definition. More details on these topics can be found in [61].

2.2.2 Basic Statistical Methods

These methods are frequently applied in data analysis and require little prior knowledge beyond elementary statistics.

Kernel Density Estimation (KDE)

A KDE reconstructs a probability density function based on a sample $x_1, \ldots, x_n$ of measurement data by smoothing the histogram of the data [62, 63]. The estimated

density $\hat{\rho}(x)$ is modelled as a weighted sum of probability densities (kernels) centred around the measured x_i.

Residuals and Root Mean Squared Error

For an algorithm f which estimates values $\hat{y}$ from data X with true values y, there are several methods to evaluate the accuracy of f. One of them is the root mean squared error *RMSE*. It is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} r_i^2} \tag{2.11}$$

via the residuals $r_i = \hat{y}_i - y_i$. An *RMSE* much smaller than the range of measured values y_i means that the model shows only little deviation from the data. A roughly symmetric distribution of the residuals around 0 indicates that the model does not have a bias towards particular values.

Coefficient of Prediction ρ^2

Another method to evaluate the quality of an estimated function f is the coefficient of prediction ρ^2 used in [50]. It takes the value of $\rho^2 = 1$, if the prediction is always exactly true, and $\rho^2 = 0$, if the prediction is only as accurate as always using the mean $\bar{y}$. It is defined by

$$\rho^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y} - y_i)^2}. \tag{2.12}$$

Bootstrapping

Bootstrapping is used to estimate standard deviations and confidence intervals (CIs) in a model-free approach. A sample $z_1, \ldots, z_n$ is re-sampled *with* replacement, i.e. a new sample $\tilde{Z} = z_{i_1}, \ldots, z_{i_n}$ is created that for some $j \neq k$ fulfils $i_j = i_k$. This procedure is repeated N times so that there are $\tilde{Z}_1, \ldots, \tilde{Z}_N$ bootstrapped samples. For N large enough, e.g. the mean $\tilde{\mu}(z)$ of the re-sampled data will converge to the true mean of the original sample, but the empirical distribution of the re-sampled means $\tilde{\mu}_1(z), \ldots, \tilde{\mu}_N(z)$ enables the calculation of the confidence interval of the empirical mean [64]. This approach can be adapted to make inference on the standard deviation and CIs of any statistical property of the original sample.

Principle Component Analysis

For multivariate data, principle component analysis (PCA) can be used to transform the data into a different set of coordinates [65]. PCA constructs an orthonormal base of uncorrelated linear combination of the original features. Additionally, the PCA dimensions are ranked according to how much variance of the full data is explained by them. Hence, a low-dimensional reduction to the first PCAs is often used to visualise and analyse multivariate data.

Information Criteria

The performance of two models can be compared via information criteria to balance model complexity and accuracy, if the model formulation includes a likelihood function for the observed data. The Akaike information criterion (AIC) for a model with p parameters is given by

$$AIC = -2\mathcal{L} + 2p \tag{2.13}$$

with $\mathcal{L}$ as logarithmic likelihood of the model [66]. Lower AIC values indicate a better or more parsimonious fit and the AIC differences between two models can be interpreted similar to the characteristic values of hypothesis tests [67]. Variations such as the Bayesian information criterion for a sample size n given by

$$BIC = -2\mathcal{L} + p \log n \tag{2.14}$$

put an even greater emphasis on the sparsity of the model.

2.2.3 Stochastic Differential Equations

As many time series in physics and finance are described by stochastic differential equations (SDEs), it is prudent to briefly review some concepts related to them. We note that this section is rather heuristic and simplifies the formal mathematical definitions. For many SDEs, a Wiener process or Brownian motion $(W_t)_t$ is the reason for their non-deterministic behaviour. Such a process is defined for continuous time t and fulfils that the increment distributions $W_t - W_s$ and $W_s - W_q$ with $t > s > q$ are independent and distributed according to a Gaussian density

$$W_t - W_s \ \sim \mathcal{N}(0, t-s). \tag{2.15}$$

As the above equation holds true for infinitesimal intervals $(t-s)$, the time derivative $\frac{\mathrm{d}W}{\mathrm{d}t}$ is a Gaussian white noise. An SDE for a time series $(X_t)_t$ can then be expressed as

$$\mathrm{d}X = a(X,t)\,\mathrm{d}t + b(X,t)\,\mathrm{d}W \tag{2.16}$$

with deterministic drift function a and diffusion b or alternatively

$$\frac{\mathrm{d}X}{\mathrm{d}t} = a(X,t) + \tilde{b}(X,t)\epsilon \tag{2.17}$$

with $\epsilon \sim \mathcal{N}(0,1)$ and any multiplicative scaling being absorbed by $\tilde{b}$. This thesis will use the notation in equation (2.17) and refer to it as a Langevin equation, but both notations appear in the literature. Simulating SDEs on a computer needs a discretisation of time with time steps Δt. The Euler-Maruyama simulation scheme for simulating X_t via

$$X_{t+\Delta t} = X_t + a(X_t,t)\Delta t + b(X_t,t)\Delta W_t \tag{2.18}$$

can be deduced from equation (2.16) with $\Delta W_t = \sqrt{\Delta t}$ from equation (2.15).
If we are interested in the differential of a function $f(X)$ of a stochastic process X, we need Ito's lemma to calculate $\mathrm{d}f$. In contrast to deterministic problems, the linearisation of f has to take the second order derivatives into account because the second order of the infinitesimal Brownian motion fulfils $(\mathrm{d}W)^2 = \mathrm{d}t$ and is therefore on the linear order of $\mathrm{d}t$. With Taylor expansion of f, we get Ito's lemma

$$\begin{aligned}
\mathrm{d}f &\approx \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial X}\,\mathrm{d}X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\mathrm{d}X)^2 \\
&= \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial X}\,(a\,\mathrm{d}t + b\,\mathrm{d}W) + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\,(a\,\mathrm{d}t + b\,\mathrm{d}W)^2 \\
&= \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial X}a + \frac{1}{2}b^2\frac{\partial^2 f}{\partial X^2}\right)\mathrm{d}t + \frac{\partial f}{\partial X}b\,\mathrm{d}W + \mathcal{O}\left(\mathrm{d}t^{1.5}\right).
\end{aligned} \tag{2.19}$$

Further details can be found in [16].

2.2.4 Bayesian Statistics and Markov Chain Monte Carlo

Classical frequentist statistics is mainly concerned with the likelihood $f_L(x|\theta)$ of observing a measured event x given some underlying parameters θ. For example, the maximum likelihood principle

$$f_L(x|\theta) \stackrel{!}{=} \max \tag{2.20}$$

is used to estimate the optimal parameters θ to fit a model to the observed data x. Also, hypothesis tests are used in frequentist statistics to test, e.g. whether the true value of the parameter θ is exceeding a certain threshold θ_0 [61]. In contrast, Bayesian statistics does not treat the parameters θ as fixed, but rather assigns an uncertainty to them via a probability distribution [55]. Without taking the observed data x into account, one assigns a prior distribution $f_{prior}(\theta)$ to the parameters based on previous knowledge or expert judgement. Generalising Bayes's theorem for conditional probabilities $\mathbb{P}(\cdot|\cdot)$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)} \tag{2.21}$$

to conditional probability densities, the posterior distribution f_{post} can now be defined as a combination of the prior knowledge and observational data via

$$f_{post}(\theta|x) = \frac{f_{prior}(\theta)f(x|\theta)}{f(x)} \sim f_{prior}(\theta)f(x|\theta). \tag{2.22}$$

The normalisation term $f(x) = \int f_{prior}(\theta)f(x|\theta)\,\mathrm{d}\theta$ is usually omitted because it does not depend on θ and therefore is just a scaling constant. Thus, Bayesian statistics allows to formulate a probability density for the unknown parameters θ which is at the heart of the many methods derived from this methodology [55, 68]. Usually, with increasing number of observations, the posterior distribution becomes less dependent on the prior density and instead approximates the likelihood. Notably, Bayesian methods can naturally be extended to incorporate new knowledge in a statistical approach to learning as whenever new data is available, one can always use a previous posterior density as a prior distribution and incorporate the new data in an updated posterior distribution.

Markov chain Monte Carlo algorithms (MCMC) are often used to calculate the posterior distribution by numerically solving the integral in equation (2.22). Therefore, even though they are not strictly a Bayesian technique, MCMC methods often get taught alongside Bayesian statistics [55, 68]. MCMC generates a sample $\theta_1, \theta_2, \ldots$ via a Markov chain (i.e. choosing θ_t depends only on θ_{t-1}, not on θ_{t-2}) that is supposed to approximate a distribution function $g(\theta)$ according to Monte Carlo methods. The standard MCMC algorithm used is the Metropolis-Hastings algorithm [69], but all MCMC methods share the same core idea to sample an unknown distribution efficiently by mostly exploring areas of the distribution's support where the probability distribution is relatively high instead of wasting time on areas with almost no contribution to the distribution. Following the notation from [70], a new candidate s for the sample is suggested by a candidate

distribution $Q(s|\theta_t)$ which is usually a multivariate Gaussian as it is easy to implement and its probability mass is concentrated around the previous sample member θ_t. The Metropolis ratio

$$\rho = \frac{g(s)Q(\theta_t|s)}{g(\theta_t)Q(s|\theta_t)} \tag{2.23}$$

is computed and s is accepted ($\theta_{t+1} = s$) with probability $\min(1, \rho)$. If it is rejected, the Markov chain stays in its previous position $\theta_{t+1} = \theta_t$. This decision rule ensures that the process will tend to move towards regions of high probability density but still allows it to make a downhill move and thereby escape a local maximum and enter another local maximum of the density. Usually, several Markov chains are run in parallel as so-called walkers and the first n_{burn} elements of each walker are discarded as burn-in because they depend too strongly on the initial position θ_1 where the chain was started. The resulting ensemble of sampled $(\theta_k)_k$ can then be used to e.g. directly derive credible intervals (CIs) of 95% or any other accuracy from the quantiles of the sample. Moreover, if θ is multidimensional, the samples can also reveal correlations between the different components of θ. MCMC sampling is implemented in various programming languages, e.g. in Python via the package *emcee* [71].

2.2.5 Machine Learning

The field of machine learning (ML) and artificial intelligence (AI) is a fast growing discipline and it cannot be summarised in a single section of this thesis. Instead, a brief introduction to the fundamental concept of *learning* and an introduction to neural networks as the current standard method of machine learning will be given. Parts of this chapter will be taken from [56] which provides a more detailed, yet still short overview on machine learning. Of course, the interested reader is recommended to take a look at the standard textbook on machine learning for further details [72].

Training and Testing in the Machine Learning Methodology

The fundamental aspect which differentiates machine learning from classical statistical methods is to include evaluation on unknown data as part of the fitting process. The observed data is initially split into training data and test data and the algorithm – whether it is a linear regression, a decision tree or a deep neural network – is fitted to the training data. Then, its performance is evaluated on the test data which has not been used to tune the parameters before. The steps of evaluation and improvement are repeated until the model finally achieves a satisfying level of accuracy on the unknown data and, depending on this performance, the algorithm can be changed by e.g. adding

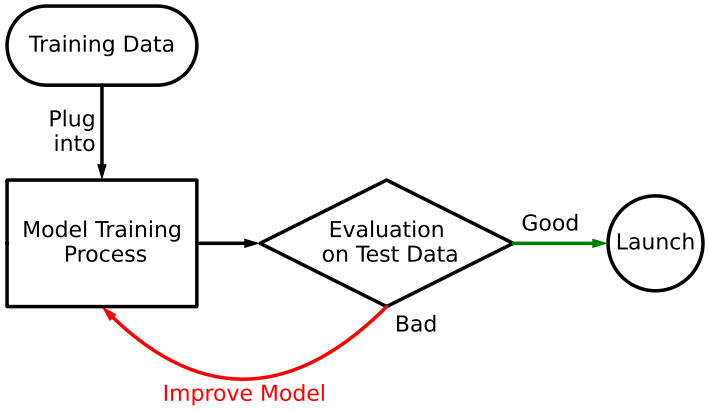

Figure 2.1: Visualisation of the machine learning methodology as a flowchart via [73]. Figure taken from [56].

or removing a layer in a neural network which is illustrated in figure 2.1. Because of modern computing power, it is possible to fit a huge model with many parameters such that it perfectly hits every observed instance in the data. This, however, is not a stable form of learning that can be generalised to unknown data, but rather just learning the observed data by heart without gaining any kind of knowledge from it and hence, it will usually fail to perform well on the unknown test data. Sometimes, an additional dataset is used as the validation data and the fine-tuning of the algorithm is done on the validation data whereas the test data is used only at the very end of the pipeline to get a final performance evaluation on new data.

Neural Networks

Many of the most impressive recent achievements of machine learning and artificial intelligence have a neural network at the core of their architecture. Often, engineering solutions have been inspired by biology (the lotus effect is a famous example for this) and as the name suggests, artificial neural networks are no exception to this. The fact that its basic structure is modelled after our brain probably helps to evoke strong associations with human-like intelligence whenever artificial intelligence based on neural networks is discussed. However, a look at the mathematics behind neural networks helps to demystify them.

The fundamental unit of a neural network, a neuron, is directly inspired by the neurons in our brain. A biological neuron receives a signal and if the signal is strong enough, it "fires" and sends out another signal as an output. Artificial neurons or perceptrons work similarly [74]: they receive input signals $i_1, \ldots, i_n$ and use a weighted sum to combine them to one value $\sum_{j=1}^{n} w_j i_j$. If this sum is large, the neuron produces a strong output according to an activation function $h(\cdot)$, otherwise its output is close to zero or exactly zero. Often, a bias term b is also added to the sum. The neuron's output is therefore given as

$$h\left(\sum_{j=1}^{n} w_j i_j + b\right) \tag{2.24}$$

where i_j are the inputs and the tuneable parameters are the weights w_j and bias b. Initially, h was usually given by a step function with $h(x) = \max(0, x)$, but because of its constant derivative, this function is problematic for gradient-based fitting methods which are used for parameter optimisation. Common activation functions that are used alternatively include

$$\begin{gathered}\mathrm{ReLU}(z) = \max(0, z) \qquad \mathrm{LeakyReLU}_{\alpha}(z) = \max(\alpha z, z) \text{ with free parameter } \alpha \\ \mathrm{SeLU}_{(s,a)}(z) = sz\mathbb{1}_{z\leq 0} + sa(\exp\{(z)\} - 1)\mathbb{1}_{z<0} \text{ with } a = 1.67326324 \text{ and } s = 1.05070098.\end{gathered} \tag{2.25}$$

However, one neuron alone is not a neural network. Instead, there are usually several hidden layers of neurons and each layer contains many neurons. The input values (i.e. the features of an instance) are plugged into each neuron of the first hidden layer and these neurons compute an output according to their activation function like (2.25). Their output values are then passed into the next deepest layer where they are used as inputs for the next neurons and so on. Finally, all outputs from the deepest layer are plugged into one or more output nodes which then give the prediction of the network. For a regression task, one output node is usually enough as it will give the numerical value that the algorithm predicts. For classification tasks with k classes, you usually have k output nodes (one for each class) whose outputs are then normalised to form a probability distribution. Output O_κ will then give you the probability that the instance is in class κ. Such networks are called multilayer perceptrons or MLPs and a sketch of such a network is shown in figure 2.2.

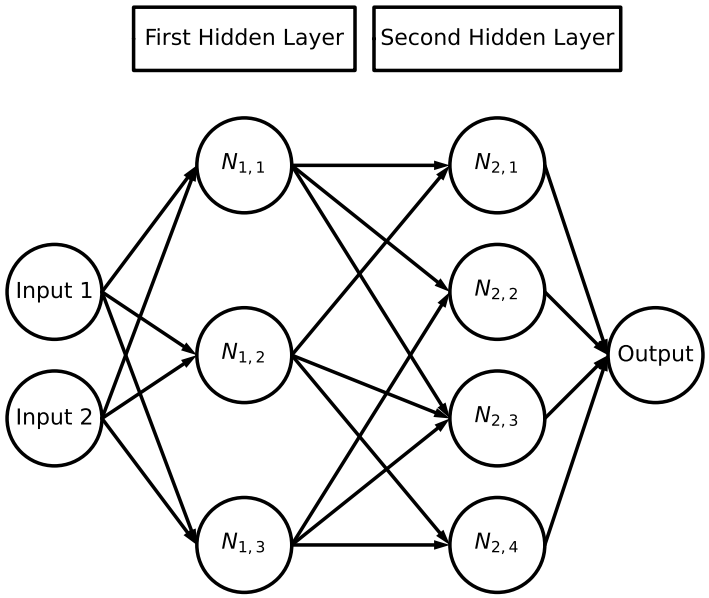

Figure 2.2: Sketch of a neural network's architecture made via [73]. Each $N_{i,j}$ is a neuron with an activation function according to (2.24) and the arrows show input/output relationships. This figure was taken from [56]

Implementation

The programming language Python has established itself as the premier language to implement machine learning models because its rather simple syntax makes it accessible for people without coding experience and its multitude of packages include many popular machine learning algorithms. The package *scikit-learn* (usually abbreviated as sklearn) provides a good general machine learning toolkit alongside some data wrangling tools [75] and also an implementation of MLPs. However, for large-scale neural networks, the packages *tensorflow.keras* and *Pytorch* are much more frequently used [76, 77]. These packages also include very straightforward implementations of more advanced architectures for neural networks beyond the MLPs.

2.2.6 Interpretability

Machine learning algorithms become increasingly more important for many industries and even day-to-day activities, yet they also become increasingly complicated and difficult to understand. Modern algorithms like the large language models from the GPT series can easily have billions of parameters which means there is no chance for a single human

to actually comprehend the algorithm's decision making. Hence, there have been some calls for a better understanding of statistical and AI algorithms. This is not only driven by the desire to better understand an algorithm, but regulatory agencies have also started to demand more transparent explanations for deployed AI algorithms [78]. As an example of the "garbage in, garbage out"-paradigm, one image classifier learnt to recognise pictures of horses only because they all had the same photographer's tag in one corner of the image [79]. This situation is quite amusing, but what if something like this occurs if the AI is supposed to make high stakes decisions?

Explainable AI

Explainable AI (XAI) has been suggested to increase trust in AI by making the "AI black box" more transparent [80, 81]. A complicated algorithm (e.g. a deep neural network) is analysed by another algorithm (the XAI method) in order to give a heuristic explanation for the AI's decision. For example, the LIME algorithm explains the the original AI algorithm's prediction for one specific instance by looking at the instances in a local environment around the instance of interest. It locally fits an easily interpretable model (e.g. linear regression with only few coefficients) to the AI's predictions and this simple surrogate model is used to locally explain the predictions. Based on the game-theoretical Shapley values [82], another method computes the SHAP values by iterating over all explanatory features and computing how much explanatory power a feature adds to the predictions based on coalitions of other features [83]. SHAP can be used to both locally explain a single instance and to globally give importance scores to the features based on the full algorithm and all available data. Unlike LIME and SHAP, the Layer-wise Relevance Propagation (LRP) is a method that cannot be applied to all algorithms, but is specifically designed for neural networks or problems that can be reformulated as a network. It essentially runs backwards through the network and distributes the relevance of each neuron to the neurons in the previous layer whose inputs affected it [84]. Finally, it distributes the relevance on the level of the input features and can, e.g. produce a heatmap to show which parts of an image were most important for the network's prediction.

Causality

While XAI methods can give the user of AI algorithms a heuristic explanation for the AI's decisions, the algorithm itself still remains a black box and one has to trust that the XAI did not introduce a source of error or bias. Instead, some researchers have

argued that modern AI systems should be designed as inherently interpretable and understandable algorithms with the help of domain experts and that these systems can, due to modern computing power, nevertheless achieve high accuracy [85]. Additionally, physical constraints can be explicitly taken into account as an additional condition in the parameter optimisation of physics-informed neural networks [86]. Following a similar philosophy, the practice of simulating scientific problems via models has recently been criticised for straying too far away from the real world in an attempt to produce "interesting" phenomena rather than accurately modelling reality [87, 88].

As humans and scientists, we tend to think in relations of cause and effect and wish to model reality in these terms. While standard statistical and machine learning models tend to remain at a purely correlational description, a new school of thought has emerged in statistics which attempts to uncover causal relationships in the data [89]. Paramount to causal inference is that there must be a time lag between the cause and the effect: If X causes Y, the relationship between X_t and Y_t is irrelevant, but a time lag τ should yield a significant predictive power of $X_{t-\tau}$ on Y_t.

One of the standard methods of causal inference is the Granger causality of X on Y which measures how much the uncertainty of a future prediction of Y_{t+1} increases, if $X_{t,t-1,\dots}$ are excluded from the set of predictor variable $\mathcal{U}$ used to estimate Y_{t+1} [90]. For dynamical systems, convergent cross mapping CCM has established itself as a frequently used method to infer causal relations. CCM analyses the phase space M_X of time-lagged observations of X, i.e. $(X_t, X_{t-\tau}, \dots, X_{t-k\tau})$. If variables X and Y are causally coupled, a local environment of an element in M_X should be possible to be mapped bijectively to M_Y such that $Y(t)$ can be predicted via the next neighbours of $X(t)$ in M_X [91, 92]. Oscillator systems in which both the change in momentum depends on the position as well as the change in position on the momentum prove to be challenging, yet interesting case studies for causal inference [93, 94].

Beyond these methods originating from data science, Judea Pearl has introduced structural causal models (SCMs) as an alternative framework to model causal relationships [89]. SCMs are restricted to variables whose causal relationships can be encoded in a directed acyclic graph and Pearl introduces the do-operator as a mathematical tool to model interventions on the system (e.g. the air pressure should not rise just because someone moved the pointer on a barometer). This might prove to pave the way for a new mathematical understanding of causality in the future.

Part II

Data-Driven Analysis of Financial Markets

The abundance of data from financial markets has sparked the interest of physicists in these systems and has allowed them to contribute to financial research with data-driven tools [16, 18]. Naive approaches involving Gaussianity assumptions and independence fail to describe many phenomena found in financial markets and insights from physics help, e.g. to describe financial crashes or the correlation matrices of financial time series. The following chapters analyse financial time series of different time scales to better understand the market dynamics:
In chapter 3, daily and intraday price time series of individual companies are modelled as stochastic differential equations. Extensions to the geometric Brownian motion (GBM) allow us to identify regimes of stable fixed points in the price dynamics in chapter 3. We then focus on the interactions between different assets and analyse the collective correlation of all assets via two methods in chapter 4: First, explainable AI (XAI) is used to better understand the market states which are found in correlation matrices of daily sector returns. Second, the time series of the market's weekly mean correlation is modelled as a stochastic differential equation which reveals memory effects in the correlation's dynamics. Finally, going beyond correlation, Granger causality is used to build a network of cause and effect between the daily return time series of business sectors in chapter 5. With the help of the Helmholtz-Hodge-Kodaira decomposition (HHKD), this network is split into a rotational and gradient-based subgraph and the latter reveals that the precious metal and pharmaceutical sectors are dominant causal drivers in the financial market during the Covid crisis.

3 Fixed Points and Langevin Potentials for Single Assets

"It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

— Sir Arthur Conan Doyle, Sherlock Holmes

Abstract The geometric Brownian motion (GBM) is a standard model in quantitative finance, but the potential function of its stochastic differential equation (SDE) cannot include stable nonzero prices. This chapter generalises the GBM to an SDE with polynomial drift of order q and shows via model selection that $q = 2$ is most frequently the optimal model to describe the data. Moreover, Markov chain Monte Carlo (MCMC) ensembles of the accompanying potential functions show a clear and pronounced potential well, indicating the existence of a stable price.

This chapter is based on *Tobias Wand, Timo Wiedemann, Jan Harren and Oliver Kamps, Physical Review E 109, 024226* (2024) [95]. Tobias Wand conceptualised and carried out the research for this publication and wrote the data analysis code. Timo Wiedemann and Jan Harren curated the data and Oliver Kamps supervised the project. All figures in this chapter are taken from [95] and were created by Tobias Wand.

3.1 Introduction

Ever since Bachelier's seminal work on the Brownian Motion to describe stock prices and its extension to the geometric Brownian motion (GBM) in the Black-Scholes-Merton model [26, 27, 28], differential equations have been an important tool to analyse financial data. Econophysicists have used (stochastic) differential equations to analyse currency exchange data [96], interest rates [97] and stock market crashes [98] or derived (partial) differential equations from microscopic trading models [99, 100]. A recent empirical study tried to model price time series with a harmonic oscillator ODE to reconcile the randomness of financial markets with the idea of a fair price [101]. The GBM, still widely used as a standard model for price time series, presents the researchers with a subtle difficulty with regards to its interpretation: Its deterministic part implies either an unlimited exponential growth or an exponential decline to a price of 0 as pointed out in [102, 103, 104]. While traditional finance models have tried to improve the GBM by changing its stochastic component, the deterministic part has largely been left unchanged (cf. the discussion in section 1 of [104]). While [104] used a constrained model with regularisation via strong prior information to fit parameters to their model, the goal of the present chapter is to estimate model parameters and to select the best model without any of these restrictions, i.e. letting the data speak for itself.
The estimation of Langevin equations from data via the Kramers-Moyal coefficients [105, 106] sparked a family of methods to estimate nonparametric drift and diffusion coefficients to model the observed system as a stochastic differential equation which have also been applied to financial data [107]. A particularly interesting expansion of this method is given by the maximum-likelihood-framework (ML) in [108]: for each time step t_i and observed data x_i, the transition likelihood $L_i = p(x_{i+1}|x_i)$ from x_i to x_{i+1} is calculated and the joint likelihood $L = \sum_i L_i$ is maximised by the estimation algorithm. This approach takes the inherent stochasticity of stochastic differential equations into account and can be performed with a parametric model to recover algebraic equations to increase its interpretability. Similarly, the SINDy algorithm recovers a sparse functional form of the underlying algebraic equations by fitting the data to a candidate function library, but struggles with noisy and stochastic data [109, 110].

This chapter uses a combination of the ML framework of [108] with the candidate function library in [109] for a robust method to estimate stochastic differential equations from data similar to [111]. As an extension, the presented method can be used to estimate data from time series with non-constant time increments $\mathrm{d}t_i \neq \mathrm{d}t_j$. We use stock

market prices at daily and 30-minute intervals as described in section 3.2 to estimate their stochastic differential equations. In particular, we estimate the potential in which the dynamics take place to evaluate the stability of the dynamical process with the overall goal to distinguish between periods with a stable fixed point and unstable dynamics as explained in section 3.3. The results for the different polynomial orders of the model and their implication for the stability are shown in section 3.4 and discussed with respect to possible applications for risk assessment in section 3.5.

3.2 Data

We analyse stock market data from the companies listed in table 3.1 to cover a range of different business sectors. Our analysis covers two distinct market conditions: (i) a calm period from early 2019 through early 2020 which was characterised by low overall volatility and (ii) the Covid selloff beginning in March 2020 which was accompanied by a spike in market volatility. We analyse two sampling intervals: daily end-of-day price changes (for which our data availability covers the whole of 2019 and 2020) and 30-minute intervals (for which our data is limited to the period between January 2019 up to and including July 2020).
Note that we are directly analysing the price time series P_t instead of the returns $r = \log(P_{t+1}/P_t)$. Although analysis of the price data is also an important contribution to research [112], the returns are often chosen as an observable because of their approximately stationary distribution which allows the application of several time series analysis methods. However, our focus is explicitly on the non-stationary behaviour of stock prices: We estimate the potential of the differential equation's dynamics for different time intervals to differentiate between dynamics with and without a stable fixed point (see section 3.3). Similarly, the work in [102, 103, 104] also uses prices to determine the position of the fixed points (or, equivalently, the wells of the potential): Although a return of 0 also indicates a fixed point, it is not clear whether the price associated with it is the same as in the previous time window under observation. In particular, the research in [102, 103, 104] stresses the important difference between fixed points at a nonzero price $P > 0$ (normal behaviour of a stock) and at a price of $P = 0$ (crash of the stock). Both phenomena correspond to a return of $r = 0$, but describe vastly different situations of the stock.

TAQ Database: We use intraday data from the TAQ (Trade and Quote) database. To account for microstructure related issues, such as the bid-ask-bounce or infrequent

Company	Business Sector	Ticker
Apple	Technology	AAPL
Citigroup	Banking	C
Walt Disney Co.	Media	DIS
Evergy Inc.	Energy	EVRG
General Electrics	Industry	GE
Pfizer	Pharmaceutics	PFE
Walmart Inc.	Retail	WMT

Table 3.1: The companies whose data has been analysed in our chapter.

trading, we rely upon quoted prices that we re-sample to a 30-minute frequency.[1] For that, we first remove all crossed quotes, i.e., all quotes where the bid price exceeds the ask, require the bid-ask-spread to be below 5$, and finally use the last valid available quote within every 30-minute interval.[2] We further account for dividend payments and stock splits, which mechanically influence stock prices, and create a performance price index using quoted mid-prices.

CRSP: We also consider lower-frequency (daily) data from the Center of Research in Security Prices (CRSP) which is one of the most widely used databases in economics and finance. We again calculate a performance price index for each stock using the daily holding period return provided by CRSP. Note that, while we use quoted mid-prices for the 30-minute high-frequency data, CRSP uses trade prices to calculate the holding period return. However, as the trading volume has increased considerably over the last decade, this should not be an issue [114].

3.3 Theoretical Background and Model

The standard stochastic differential equation to describe a stock price P is the geometric Brownian motion given by

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \mu P + \sigma P \epsilon \tag{3.1}$$

with standard Gaussian noise $\epsilon = \epsilon(t) \overset{iid}{\sim} \mathcal{N}(0,1)$, constant drift μ (typically $\mu > 0$) and volatility σ. As pointed out in [104], the physical interpretation as a particle's trajectory

[1] When we talk about quotes, we refer to the National Best Bid and Offer (NBBO) where the national best bid (offer) is the *best* available quoted bid (offer) price across all U.S. exchanges. See [113] for an overview.

[2] We forward fill quotes if there is no valid entry for a given time interval. However, this does almost never happen for very liquid stocks such as the ones chosen in this chapter.

P in a potential $V(P)$ transforms equation (3.1) to

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\mathrm{d}V}{\mathrm{d}P}(P) + \sigma P \epsilon_t \quad \text{with} \quad V(P) = -\frac{\mu}{2}P^2 \tag{3.2}$$

and an arbitrary potential constant C (set to zero for simplicity). However, analysing this potential V in terms of its linear stability (cf. section 2.1 or [1]) leads to the problematic result that the only fixed point in the data with $\frac{\mathrm{d}V}{\mathrm{d}P}(P_0) = 0$, namely $P_0 = 0$, is an unstable fixed point for $\mu > 0$. Without a stable fixed point, trajectories are expected to diverge away from $P_0 = 0$ towards infinity. As this is - at least for limited time scales - a highly unrealistic model, the authors of [102, 104] have suggested higher order polynomials in the potential V of (3.2). From the assumption that the rate of capital injection by investors should depend on the current market capitalisation, they derive a quartic potential

$$\begin{aligned} V(P) &= -P\left(\frac{\alpha_1}{2}P + \frac{\alpha_2}{3}P^2 + \frac{\alpha_3}{4}P^3\right) \text{ with} \\ -\frac{\mathrm{d}V}{\mathrm{d}P}(P) &= \alpha_1 P + \alpha_2 P^2 + \alpha_3 P^3, \end{aligned} \tag{3.3}$$

i.e. the drift term $-\frac{\mathrm{d}V}{\mathrm{d}P}(P)$ is a polynomial of order $q = 3$. With a suitable choice of parameters α, this potential can adopt the shape of a double-well potential with stable fixed points at $P_0 = 0$ and $P_1 > 0$, thereby predicting both the presence of a bankruptcy state at the stable fixed point $P_0 = 0$ and an additional stable state with nonzero price $P_1 > 0$. In [104], however, major constraints to the parameters during the estimation process were necessary to achieve this.

3.3.1 Numerical Implementation

If a time series $(s_n(t_n))_n$ with observations s_n at time t_n has been recorded, a maximum likelihood approach can be used to estimate the most likely parameters σ and α_i such that a stochastic differential equation according to equations (3.2) and (3.3) may have produced the observed time series. For any two adjacent points $s_n(t_n)$ and $s_{n+1}(t_{n+1})$ and any given parameters $\phi = (\sigma^2, \alpha_i)$, the likelihood L of observing the transition from $s_n(t_n)$ to s_{n+1} at t_{n+1} can be explicitly calculated as

$$\begin{aligned} L(s_{n+1}|t_{n+1}, s_n, t_n, \phi) = &\left(2\pi\left(\sigma s_n \sqrt{t_{n+1} - t_n}\right)^2\right)^{-\frac{1}{2}} \\ &\cdot \exp\left\{\left(-\frac{\left(s_{n+1} - \left(s_n + \left(-\frac{\mathrm{d}V}{\mathrm{d}s}(s)\right)(t_{n+1} - t_n)\right)\right)^2}{2\left(\sigma s_n \sqrt{t_{n+1} - t_n}\right)^2}\right)\right\} \end{aligned} \tag{3.4}$$

or as the logarithmic likelihood

$$\begin{aligned}\mathcal{L}(s_{n+1}|t_{n+1}, s_n, t_n, \phi) &= \log L(s_{n+1}|t_{n+1}, s_n, t_n, \phi) \\ &= -\tfrac{1}{2}\log\left(2\pi\left(\sigma s_n\sqrt{t_{n+1}-t_n}\right)^2\right) \\ &\quad -\frac{\left(s_{n+1}-\left(s_n+\left(-\frac{\mathrm{d}V}{\mathrm{d}s}(s)\right)(t_{n+1}-t_n)\right)\right)^2}{2\left(\sigma s_n\sqrt{t_{n+1}-t_n}\right)^2}.\end{aligned} \tag{3.5}$$

Because we assume Markovian dynamics, the complete log-likelihood for the full observed time series is then simply the sum over the stepwise log-likelihoods

$$\mathcal{L}\left((s_n)_n|(t_n)_n, \phi\right) = \sum_{i=0}^{n-1}\mathcal{L}(s_{i+1}|t_{i+1}, s_i, t_i, \phi). \tag{3.6}$$

For given observations (s_i, t_i), the likelihood $\mathcal{L}$ can be maximised by varying the parameters ϕ to estimate the optimal parameters ϕ^*.

According to Bayes's theorem and Bayesian statistics [55], the likelihood of observing the measured data conditional on some parameter values $L\left((s_n(t_n))_n|\phi\right)$ is combined with an a-priori distribution $f_{prior}(\phi)$ to calculate a posterior distribution of the parameters given the observed data:

$$f_{post}(\phi|(s_n(t_n))_n) \sim f_{prior}(\phi)L\left((s_n(t_n))_n|\phi\right). \tag{3.7}$$

For an uninformed flat prior, this transformation is mathematically trivial, but allows us to calculate f_{post} as a probability density of the parameters ϕ conditional on the observed data. Hence, the distribution of the parameters ϕ can be explored via Markov chain Monte Carlo (MCMC) methods [69] by drawing samples $\left(\phi^{(j)}\right)_j$ from the posterior distribution as implemented in the Python package *emcee* [71]. MCMC can uncover correlations between different parameters and also explore local maxima of the probability density. It therefore gives a more complete view of the underlying distribution than summary statistics like the mean or standard deviation. In particular, we will use the sampled parameters $\left(\phi^{(j)}\right)_j$ to construct an ensemble of potentials $V(P)$ and evaluate whether their shapes are roughly consistent with each other.

3.3.2 Synthetic Data

To test our method, synthetic time series $(s_n)_n$ are simulated via the Euler-Maruyama scheme [115] as

$$s_{n+1} = s_n + \left(-\frac{\mathrm{d}V}{\mathrm{d}s}(s)\right)(t_{n+1} - t_n) + \sigma s_n \sqrt{t_{n+1} - t_n}\epsilon_n \tag{3.8}$$

with $\epsilon_n \overset{iid}{\sim} \mathcal{N}(0,1)$ for any parameters α_i for the potential in (3.3). The following paragraphs discuss how well our model can then identify the underlying dynamcis and parameters from the observed data (s_n, t_n). Note that for the synthetic data, non-equidistant time steps have been used.

Estimating the Correct Order

Generalising the potential from (3.3) to a potential with arbitrary polynomial order q leads to

$$V(P) = -P\sum_{i=1}^{q} \frac{\alpha_i}{i+1} P^i. \tag{3.9}$$

For given q, random parameter values α_i and a random noise level σ are sampled and the resulting time series is simulated with these parameters. If the sampled parameters result in numerical errors (i.e. the time series diverges towards infinite values), the time series is discarded from the ensemble. This is done repeatedly until the ensemble includes 100 time series with a length of 1000 time steps each. For the synthetic time series, the best order is estimated via the Akaike information criterion AIC [66] given by

$$AIC = -2\mathcal{L}_{\max} + 2(q+1) \tag{3.10}$$

where $q+1$ is the total number of model parameters (q monomials' prefactors α_i and σ). Hence, q is varied, the maximum likelihood $\mathcal{L}_{\max}$ for the chosen q is estimated and the resulting AIC is calculated. The model with the lowest AIC is chosen as the best model. The results shown in figure 3.1 show that for polynomial orders $q = 1$ (geometric Brownian motion), $q = 2$ and $q = 3$ (Halperin's suggestion as in equation (3.3)), the correct order is usually identified as such. An even higher order $q = 4$ shows very unreliable results, but will be included for completeness for the further data analysis.

Estimating the Parameters

Instead of sampling repeated trajectories with different parameters, now for order $q = 3$, the parameters $\phi = (\sigma^2, \alpha_1, \alpha_2, \alpha_3)$ are kept constant as $\phi = (0.05, 2, -1, 0.01)$. 100 time series with 1000 time steps are sampled and their parameters are estimated by fitting a model with $q = 3$. Despite the constant parameters, the randomness of the ϵ_n in equation (3.8) nevertheless ensures that the time series are different from each other. The

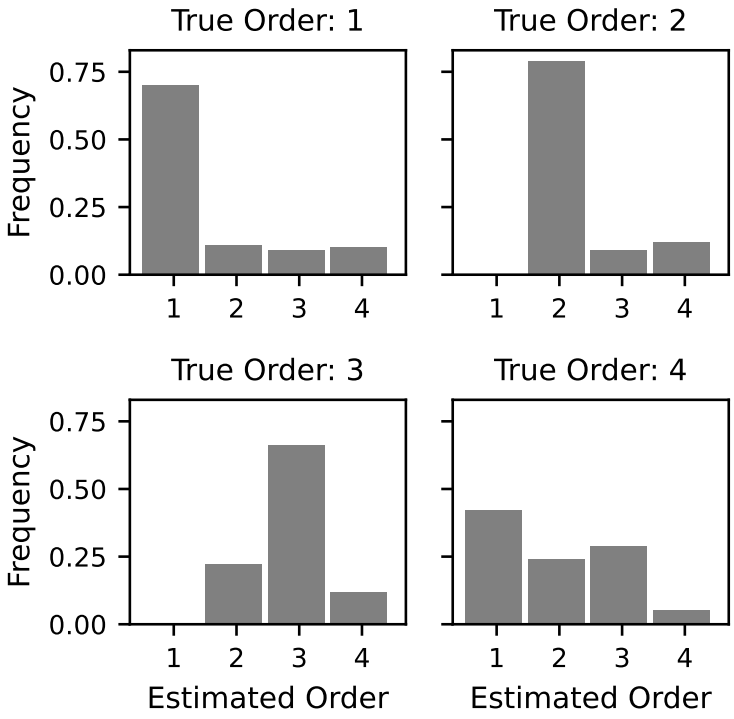

Figure 3.1: For each true polynomial order q, 100 trajectories are randomly sampled and then their best order is estimated. The histograms show that the method successfully estimates trajectories with order $q = 1, 2, 3$, but struggles with the higher order $q = 4$.

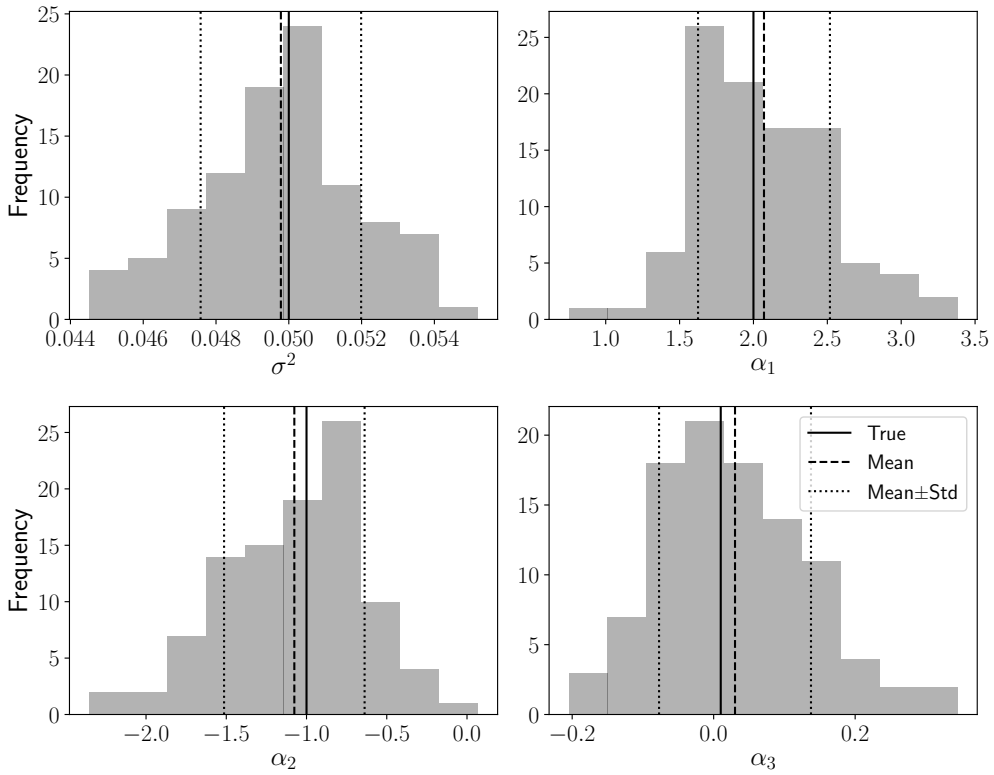

Figure 3.2: Parameter estimations for 100 trajectories with the same true parameters $\phi = (\sigma^2, \alpha_i)$ as given by the solid lines. The interval of Mean ± Standard Deviation of the estimated ensembles always includes the true parameters.

histograms of the estimated parameters, their means and standard deviations are shown together with the true parameter values in figure 3.2. Note that the true parameter value is always within the one-standard-deviation-interval around the mean and that the parameter α_3 has a distribution virtually indistinguishable from that of a parameter with mean zero: The parameter estimation correctly shows that α_3 is so low that it is a superfluous parameter for model inference. Note that if the same data is estimated by a model with $q = 2$, the results are fairly consistent with the depicted histograms. However, fitting the data to a model with $q = 4$ results in the one-standard-deviation-interval of α_2 also containing the value 0, which is a consequence of overfitting the model.

3.4 Results

While [104] analyses time periods of one year to estimate parameters, we believe that because of the assumption of constant volatility in equation (3.4), it is prudent to restrict

Optimal Order (30 Min)	Opt. Order (Daily)			
	1	2	3	4
1	24	6	3	1
2	6	41	10	3
3	1	10	8	1
4	7	7	3	2

Table 3.2: Comparison of the estimated orders q for the same company and month with the price time series in daily and 30-minute intervals.

the date to short intervals of one trading month. Hence, we divide the given data into non-overlapping monthly intervals and estimate the polynomial order q of the underlying stochastic differential equation via the *AIC*. Note that the time difference between each observation is taken as a constant interval of 1 time step in trading days or 30-minute steps, respectively, including the overnight return.

3.4.1 Polynomial Orders

The distributions of the estimated polynomial orders q are shown in table 3.2 and in figure 3.3: On both time scales and in all market periods, the order $q = 2$ is the most frequently estimated order with the GBM model at $q = 1$ being the second most frequent estimation. The suggestion $q = 3$ from [104] as well as the even higher-order $q = 4$ are only rarely estimated as the most accurate model. Interestingly, calm and turbulent periods (as defined in section 8.1.2) show essentially identical distributions, whereas the order $q = 4$ seems to be a bit more frequently estimated for the shorter time scale of 30 minutes than for the daily data. Table 3.2 shows that there is high consistency between the estimated orders for both time intervals for orders $q = 1$ and $q = 2$, but increasing disagreements for orders $q = 3$ and $q = 4$.
Overall, this suggests that a polynomial order of $q = 1$ and $q = 2$ can be a reasonable modelling assumption for the time series data. Moreover, the identification of these two orders is consistent for the two sampling intervals under consideration, whereas the choice of calm or turbulent periods does not seem to influence our results.

3.4.2 Potentials

For the optimal polynomial order q of each month, we then sampled the parameters $\phi = (\sigma^2, \alpha_1, \ldots, \alpha_q)$ from their posterior probability distribution to get an ensemble of parameters $\left(\phi^{(j)}\right)_j$. From that, we calculate the corresponding ensemble of potentials

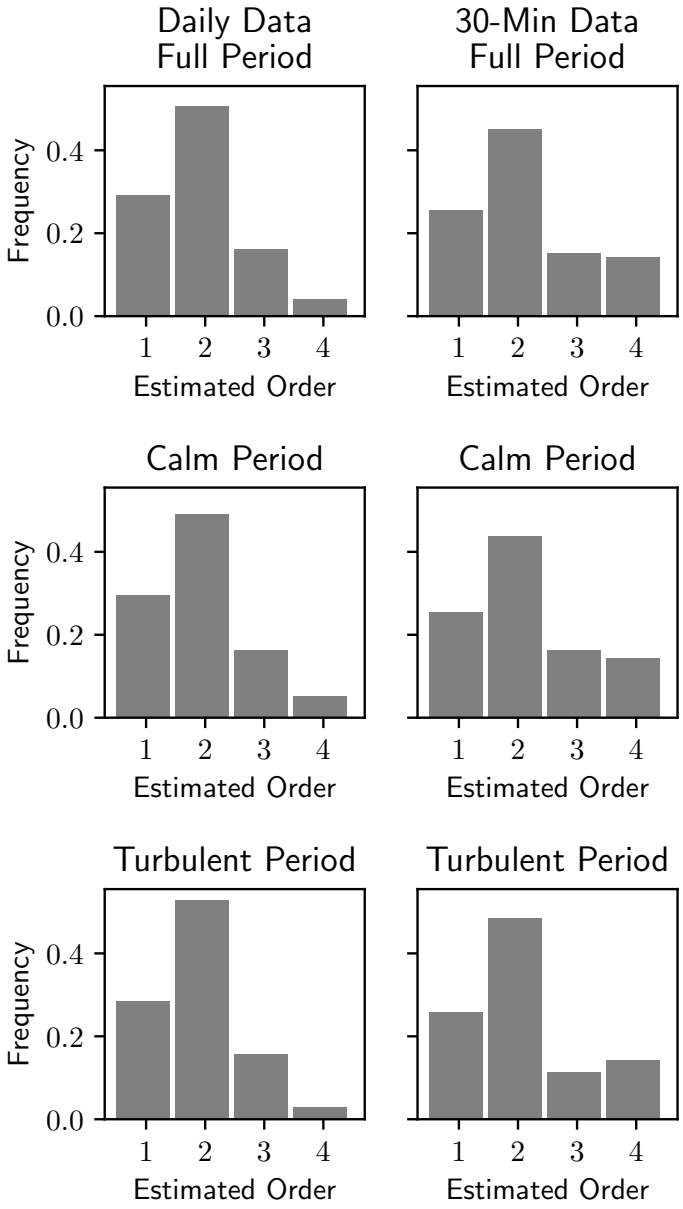

Figure 3.3: Estimated polynomial orders q of the monthly real-world price time series of the companies in table 3.1 for the daily (left) and 30-minute intervals (right). Differences between the calm (up to February 2020) and turbulent period (starting in March 2020) are only small.

$V(P)$ according to (3.9) and plot them, their pointwise centred 68% and 95% credible intervals (CIs) and the potential corresponding to the maximum likelihood estimation. The zero horizontal is shown in these plots as the y-axis position of the potential at $P = 0$ to indicate where the potential is above or below the potential energy at the zero price (and if the price "particle" would therefore prefer or not prefer to be at the potential level of $P = 0$). A couple of generic features can be observed for these potentials and do not depend on the chosen sampling rate:

Order 2

As the order $q = 2$ is the most frequently identified polynomial order according to the results in figure 3.3, it is quite insightful to focus on its associated potentials. They virtually always look like the potential depicted in figure 3.4 and show a potential well as a pronounced minimum. Close to this minimum, the 68% CI is usually also below 0 and sometimes (as depicted in figure 3.4) even the 95% CI. The MCMC-sampled potential ensembles thus support the existence of a potential well as they clearly show the potential well for a large majority of trajectories. Following the interpretation of the potential wells from [104], this supports the existence of a locally stable price within this potential minimum.

Order 1: GBM

The GBM with $q = 1$ is the second most frequently estimated order. As shown in the subplots a and b in figure 3.5, the MCMC samples show two general types of ensembles: in a, the maximum likelihood estimation of the potential is always very close to 0 and the 68% CIs therefore envelope the zero horizontal. This makes it difficult to gauge a clear direction of the potential and hence of the movement of the price time series. Contrary to that, potential ensembles like the one shown in b have a clear direction. In b, the potential is increasing for higher prices (and hence a restoring force pulls the price to the minimum at 0), but decreasing potential ensembles can also be found for other time intervals. That means some time intervals like b show a predominant direction of price movement, whereas others like a have no predominant direction, but rather a random movement.

Order 3

The potential with $q = 3$ is the one suggested in [104] and the typical shape of their MCMC samples are shown in subplot c of figure 3.5. Note that some of these potentials

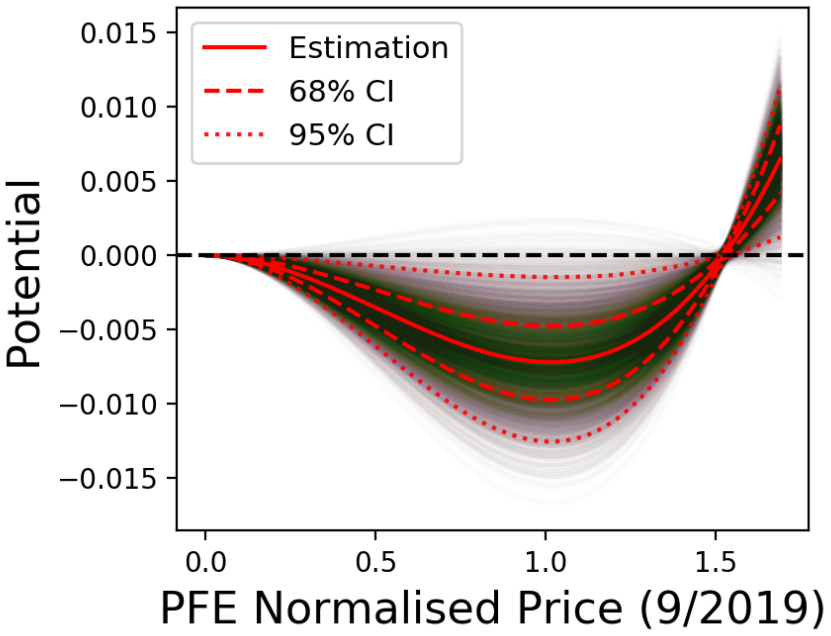

Figure 3.4: MCMC-sampled potentials of order $q = 2$ for 30-minute intervals for Pfizer in September 2019. The maximum likelihood estimation (Estimation) lies firmly within the ensemble of potentials and even the 68% credible interval shows a clear potential well.

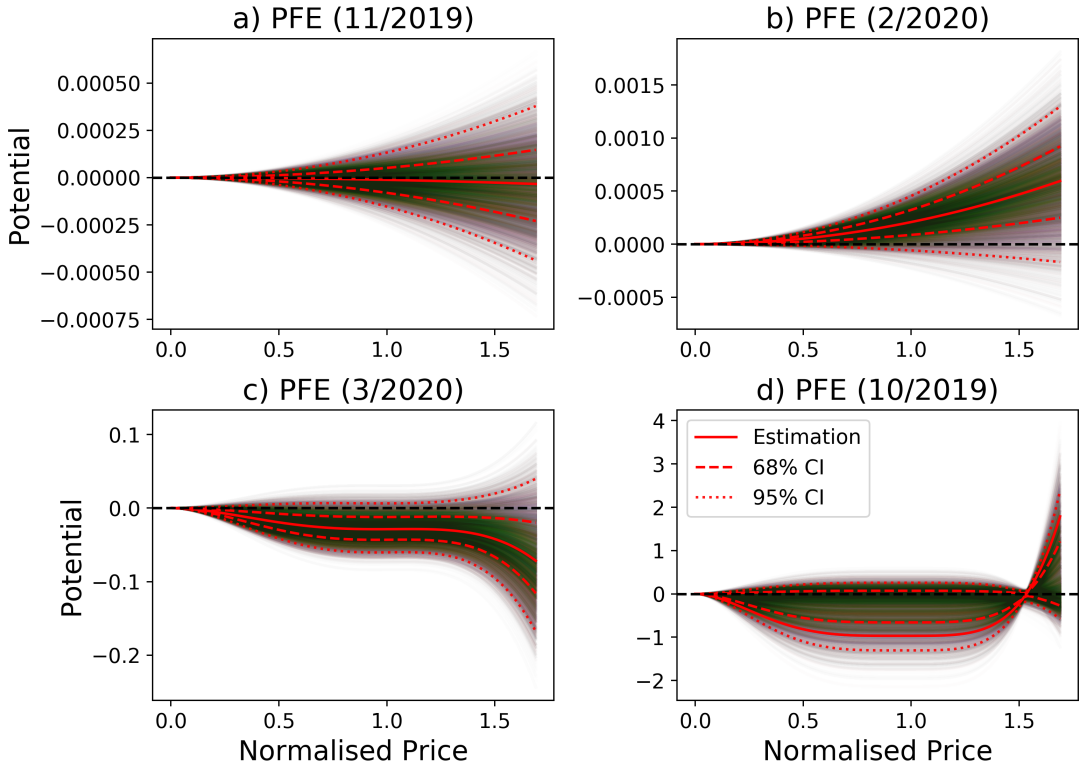

Figure 3.5: MCMC-sampled potentials of order q for 30-minute intervals for Pfizer in time intervals with $q = 1$ (a and b), $q = 3$ (c) and $q = 4$ (d). Note that the maximum likelihood estimation lies within the 68% credible interval for all orders q except for $q = 4$.

are also mirrored along the x-axis. Similar to the potentials in b, they also show a predominant direction, but also often a bistable saddle point. Notably, they do not show the pronounced double-well potential predicted in [104].

Order 4

Potentials with $q = 4$ usually correspond to very wide potential wells as can be seen in subplot d of figure 3.5 (compare its full width at half maximum to that of the potential in 3.4). In the depicted MCMC ensemble, the maximum likelihood estimation does not lie within the 68% credible interval. This indicates a multimodal posterior distribution and was found in surprisingly many time intervals. Similar to the ensemble shown in subfigure a, these potentials also often envelope the x-axis with their 68% CIs and therefore show no clearly predominant direction.

3.5 Conclusion

3.5.1 Summary

We use a maximum likelihood estimation to analyse price time series of stocks. Via the AIC model selection, we find that a second order polynomial for the drift term often offers a suitable description of the data. While the standard GBM model with a first order polynomial is not selected as frequently as the second order model, it still appears often enough to be considered a valid candidate model. Higher order polynomials are rarely estimated. Sampling the posterior density of the parameters via MCMC reveals that the potentials of second order polynomials show pronounced potential wells (i.e. stable minima) for nonzero prices which is mathematically impossible for the GBM's potential as pointed out in [104].

3.5.2 Discussion

Our research question is heavily inspired by [104], but differs from it in a key factor: The model presented in [104] always has a drift polynomial of order $q = 3$ and uses the credit default swap rates (CDS) to estimate the probability of a considered company going bankrupt. This probability is then used as a constraint in the parameter estimation such that the jump from a potential well with nonzero price to another potential well at price zero (i.e. the stock collapsing) has a jump probability (Kramers's escape rate) which is quantified by the CDS. Thus, the work in [104] combines the price dynamics of stochastic differential equations with the CDS data as additional constraints

to estimate a stochastic differential equation with a probability of the stock crashing. In contrast to this, our estimation scheme uses no additional constraints or external data, but purely the price time series. Our $q = 3$ estimations (subfigure c in 3.5) do not show the double-well potential postulated by [104]. However, our model selection via the AIC indicates that $q = 2$ is instead the most frequently observed polynomial degree and for $q = 2$, MCMC shows a clear potential well that is consistent for the whole MCMC ensemble. In short, because we do not use additional constraints, we cannot reproduce the double well potential with default probabilities, but instead show via our fully unconstrained approach that the potential wells arise naturally just from the price time series alone. Estimating potentials for stochastic financial dynamics and analysing the stability of their fixed points has also been done in [116, 117], but two key differences exist between them and our approach: While we estimate explicitly analytical potentials for the time series of individual stocks, the work in [116, 117] estimates potentials purely numerically without a closed-form analytical expression and does so for the collective market movement instead of treating individual assets. Interestingly, [116, 117] observe transitions between the different minima of the potentials and therefore a non-stationary market behaviour, similar our study, because the AIC selection means that the stocks are not described by the same polynomial degree q for all time series. Instead, our model selection implies that the potential itself is time-dependent.

A possible explanation for this is that due to external effects, an order parameter such as the capital influx into the financial markets is changed. Then, the underlying potential might change due to these effects and experience a bifurcation which changes the price dynamics. Whether a bifurcation is a suitable description of the dynamics under such conditions requires further analysis on the transitions between the different models. As an example, imagine that a price initially starts as being in a stable fixed point with $q = 2$ like in figure 3.4, but external news change the potential to that of subfigure b in figure 3.5 with $q = 1$. Now, the price has a predominant direction of movement and is no longer experiencing a restoring force back to the price at the previous fixed point and can therefore explore new areas of the phase space (figure 3.6 illustrates the transition between different regimes). The market finally manages to process the news and their implications and finally, the price reaches a new fixed point with $q = 2$. Thus, the price at the potential minimum can be interpreted as a fair price similar to the discussion in [101]. However, further research into the transition between the different potentials is necessary to verify this interpretation. Note that GBMs with potentials such as subfigure a in figure 3.5 have essentially no predominant direction of movement and show a random walk without a restoring force. This is a different behaviour to $q = 2$ which also does

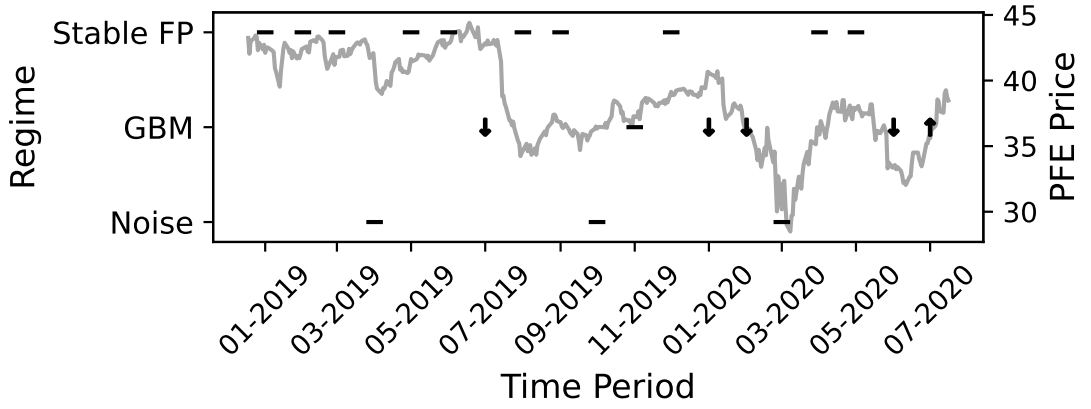

Figure 3.6: Evaluation of the different regimes for the dynamics of the Pfizer stock for 30-minute intervals. the rare orders $q = 3$ and $q = 4$ have been summarised under the label "Noise", the label "Stable FP" indicates order $q = 2$ with a potential well and the GBM of order $q = 1$ is further split up into periods of growth ($\uparrow$), random stagnation ($-$) and decline ($\downarrow$): according to their MCMC-sampled potentials (cf. subfigures a and b in 3.5), a growth or decline is only assumed if the 68% CIs do not include the 0 horizontal. The regimes of growth and decline show a good correspondence with the price time series of the associated stock.

not show a predominant direction, but instead has such a restoring force that restricts the price to the potential well.

One might have assumed that stable fixed points ($q = 2$) should occur significantly less frequently during the turbulent period because of the overall instability of the market. But interestingly, our results do not seem to show a difference between the calm and turbulent market period (cf. figure 3.3), perhaps indicating that the market can quickly adjust to such turbulent behaviour.

Finally, it is a reassuring result that the AIC selection still frequently suggests $q = 1$ (the widely used GBM model) as the best polynomial order. The standard GBM model still appears rather frequently in our data and therefore nevertheless manages to provide a reasonably accurate model.

3.5.3 Further Research and Applications

As discussed in the previous subsection, our method can be used to distinguish between different regimes (stable fixed point or growth/decay) of the dynamics of a price time series. One could use our methodology to continuously model a given time series, update

it with new data and pay attention to when the potential is changing such that the system is transitioning from a stable (resilient) state to an unstable one or vice versa. This point of view can be used to judge the system's resilience against noise and anticipate critical transitions to a qualitatively new system behaviour [111, 118]. In the non-stationary system of a free market, such monitoring might support risk management decisions.

While this chapter focused on the drift term like in [104], there are of course possible extensions of the diffusion/volatility that can also be taken into account. Stochastic volatility and local volatility models have been widely accepted in finance [119, 120], but other modelling possibilities exist, too: While this chapter used the volatility parametrisation from the GBM in (3.1) via $\sigma P \epsilon_t$, one can also imagine e.g. a polynomial model here given by

$$\text{Diffusion}(P) = \sigma \left(\sum_j \beta_j P^j \right) \epsilon_t. \tag{3.11}$$

However, from the author's experience, the maximum likelihood estimation can become troublesome if the diffusion term has several free parameters as the estimator can then attempt to essentially attribute the whole observed dynamics to the diffusion. Strong regularisation might be necessary if one wishes to expand the diffusion model. A multi-stage estimation procedure might provide another alternative: First estimate the GBM model with drift parameters $\phi_{D,1}$, then keep these parameters fixed to estimate the parameters $\phi_{V,1}$ of a more complicated volatility model (e.g. Heston's stochastic volatility). Then keep the parameters $\phi_{V,1}$ of the volatility model fixed and vary the drift parameters according to the scheme presented in this chapter in order to find the optimal order q and its associated parameters $\phi_{D,2}$. For fixed order q, iteratively use fixed $\phi_{D,n}$ to estimate $\phi_{V,n+1}$ and fixed $\phi_{V,n+1}$ to estimate $\phi_{D,n+1}$ until the parameter values converge. Developing and fine-tuning this procedure, however, is beyond the scope of the present work whose main aim was to investigate the existence of stable fixed points in the drift potential.

Another model extension might be the incorporation of memory effects. Generalised versions of the Langevin equation include non-Markovian memory terms by e.g. an explicit memory kernel [121] or by assuming the existence of a second hidden process that has not been observed [122]. Such a hidden component might correspond to the traders' knowledge or belief which certainly influences the stock prices, but is not explicitly recorded. Although we believe that there is some virtue in having a simple model as evidenced by the widespread use of the GBM, a more complex analytical model than a polynomial

approach can of course be used in the maximum likelihood framework to expand our rather simple model. Combining all these extensions and using a strict regularisation procedure to discard superfluous terms might ultimately help to develop a model that not only differentiates between the different regimes of stability (as shown in the present chapter), but also reproduces the well-known stylised facts from the empirical literature. It is noteworthy to point out that although we used equidistant time intervals between the observations of the data, the model has been tested on synthetic data with non-equidistant time intervals in section 3.3.2. Such a situation arises naturally in the context of tick-by-tick data which is the highest resolution of trading data. Here, instead of sampling the price at a high frequency, every single trade is recorded at the exact time that it occurred. As the time between two subsequent trades can be arbitrarily short or long, the application of a robust method without the need for equidistant time steps might prove useful here.

While this chapter analysed the complex dynamics of an individual asset, chapters 4 and 5 will take the interactions between different time series into account. Correlations will be the focus of analysis in chapter 4 and chapter 5 will use Granger causality to estimate the causal hierarchy of the financial market.

4 Collective Effects of the Market Correlation

"When we try to pick out anything by itself, we find it hitched to everything else in the Universe."

— John Muir, My First Summer in the Sierra

Abstract Taking correlations between financial assets into account is crucial for optimal portfolio selection and risk management. While it is well-known in the scientific literature that the market correlations show collective behaviour, the dynamical time evolution of these correlations is still underexplored. Moreover, memory effects are usually neglected for convenience, but might include important information on the system's behaviour. This chapter uses daily data from the S&P500 index to analyse the different market states with methods from explainable artificial intelligence (XAI) and to model the one-dimensional mean correlation time series with explicitly non-Markovian memory effects.

This chapter and the associated appendices D and E are based on *Tobias Wand, Martin Heßler and Oliver Kamps, J. Stat. Mech. 043402* (2023) [123] and *Tobias Wand, Martin Heßler and Oliver Kamps, Entropy 2023, 25(9), 1257* (2023) [124]. Unless otherwise stated, the figures in section 4.1 and 4.3 and appendix E are taken from [124] and the figures in section 4.2 and appendix D are taken from [123]. All figures shown in this chapter were created by Tobias Wand. In [123], Tobias Wand conceptualised the research, curated the data, wrote the code to analyse the data and carried out the research while Martin Heßler assisted with the implementation of the code and Oliver Kamps supervised the project. For the material in this thesis based on [124], Tobias Wand curated the data, wrote the code to analyse the data and carried out the research while Oliver Kamps supervised the project and all authors conceptualised the project.

4.1 Introduction

4.1.1 Correlations in Econophysics

The S&P500 is an aggregated index of the stocks of the 500 largest companies traded at US stock exchanges and therefore serves as an indicator of the overall US economic performance. Research on financial correlations has been a cornerstone of risk management ever since Markowitz's portfolio theory [125]. Although Pearson's correlation coefficient only detects linear dependencies, it is used ubiquitously in data-driven inference on financial markets [126]. Due to the increasing availability of financial data, elaborate statistical methods can be applied in this field of research. Many researchers have chosen to focus on the correlations between the relative price changes of different stocks, i.e. the correlations of the stocks' returns in their correlation matrix [34, 35, 125, 127, 128, 129]. While the majority of the eigenvalue spectrum follows the Marcenko-Pastur distribution that is expected for purely random matrices according to RMT, several eigenvalues are far outside of this spectrum as shown in figure 1.3. This hints at a strongly non-random effect and the corresponding eigenvectors often represent distinct real-world business sectors and therefore describe collective behaviour of these stocks in the market [18, 34, 35, 126]. Most importantly, the eigenvector of the largest eigenvalue has an almost equal contribution of all stocks and is therefore called the market mode as it represents the market's global behaviour. Note that [126] also points out that this property is not only found in US stock markets, but is a general observation across markets.

However, the market is not a stationary system, but experiences different states: Chapter 3 has shown this for individual assets, but even the collective motion of the market can experience states such as booms and crashes like the dot-com crisis of the US technology market shown in figure 4.1. The time evolution of many non-stationary complex systems can be described as a sequence of transient states, i.e. the system passes through a sequence of states where it remains for a short time, before transitioning to the next state. Since this represents an enormous reduction of the dimensionality of the system, the identification of these states might open the possibility for a better prediction and control of the system. Examples can be found in such diverse fields as, e.g. neurodynamics [130] and fluid dynamics [131]. Since in some cases an underlying mathematical description is not available, one has to rely on data-driven methods to find such states, e.g. via clustering or similar methods [132, 133]. Due to the increasing availability of financial data, such methods can also be applied in this field of research and can be used to gain further insights into the market's correlation matrix. In [36]

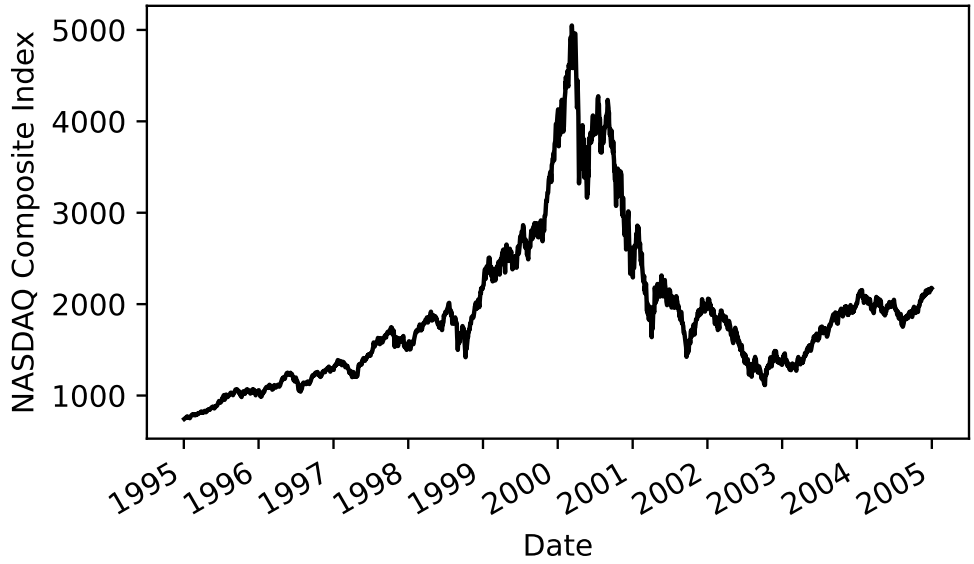

Figure 4.1: The NASDAQ Composite index is heavily influenced by the US technology companies. Around 2000, the emergence and burst of the dot-com bubble is visible as a spike.

the authors tried to identify transient states in the economy by analysing correlation matrices of daily S&P 500 data. Here, the entries $C_{i,j}(t)$ of the correlation matrix are defined as the correlation coefficient from equation (2.9) between two return time series i and j for a time window around the midpoint t. In [36, 116, 117], the correlation matrices of the daily S&P 500 data could be sorted into eight clusters, which can be interpreted as states of the economy corresponding to, e.g. economic growth phases or crises [36] and show locally stable behaviour [116, 117]. Further research showed that subtracting the dominant market mode identified in [34, 35] from the correlation matrix leads to a reduced-rank correlation matrix that highlights the differences between different market states and allows researchers to identify exogenous events, precursors for crises and collective dynamics of industrial sectors [134, 135, 136].
Moreover, as shown via Principle Component Analysis in [117], the mean correlation

$$\bar{C}(t) = \frac{1}{N} \sum_{\mathrm{i,j}} C_{i,j}(t) \tag{4.1}$$

already describes much of the variability in the data: Because the largest eigenvalue or market mode of the correlation matrix has an eigenvector with almost uniform contributions from all time series, the trajectory of the mean correlation $\bar{C}(t)$ for a time interval

with midpoint t approximates the behaviour of the collective market mode. This restricts the analysis of transient states to this one-dimensional time series $\bar{C}(t)$ and hence reduces the computational complexity of any further analysis as shown in [117]. Additionally, $\left(\bar{C}(t)\right)_t$ can be analysed with time series models to gain further insights into the market's collective dynamics.

4.1.2 Data Preparation

Daily stock data from the companies in the S&P 500 were downloaded via the Python package *yfinance* [137] for the time period between 1992 and 2012 which is the same time period as in [116, 117]. Although this limits our analysis to slightly outdated data, it ensures that we can compare our methodology and results to the findings in [117]. Only stock data of companies that were part of the S&P 500 index during at least 99.5% of the time were used for this analysis, leaving us with 249 time series for 5291 trading days. Note that while this data selection is in line with [36, 116, 117, 123], it introduces a survivorship bias by disregarding the information about companies that went bankrupt during this time period. However, it ensures that all considered companies experience the same macroeconomic events, which would not be the case for those companies that vanish during the considered time period due to bankruptcies or mergers. If a company's price time series P_t is not available for the full time period, the remaining 0.5% of the price time series data are interpolated linearly with the *.interpolate()* method in *pandas* [138, 139]. Each company's returns $R_t = (P_{t+1} - P_t)/P_t$ are locally normalised to remedy the impact of sudden changes in the drift of the time series with the method introduced in [140], as

$$r_t = \frac{R_t - \langle R_t \rangle_n}{\sqrt{\langle R_t^2 \rangle_n - \langle R_t \rangle_n^2}}. \tag{4.2}$$

Here, $\langle \dots \rangle_n$ denotes a local mean across the n most recent data points, i.e. r_t is subjected to a standard normalisation transformation with respect to the local mean and standard deviation (i.e. volatility σ). Following [36], $n = 13$ was chosen for the daily data.

Because of computational restrictions, it can be useful to limit the analysis to a lower-dimensional representation of the data instead of using the full correlation matrix with more than 10^4 entries. Two of these representations will be discussed here and analysed in this chapter: the correlation matrix of industrial sectors and the one-dimensional mean correlation.

Abbreviation	Sector
CD	Consumer Discretionary
CS	Consumer Staples
E	Energy
F	Financial Services
HC	Health Care
I	Industrials
IT	Information Technology
M	Materials
T	Telecommunication Services
U	Utilities

Table 4.1: Sectors abbreviations according to the GICS .

Industrial Sectors according to the Global Industry Classification Standard

Following the approach in [117], the data is aggregated according to the 10 global industry classification standard (GICS) sectors shown in table 4.1 representing individual parts of the economy [141]. Stock prices from companies of the same GICS sector were added to calculate representative time series $(S_t)_{t=1,\dots,T}$ for each of the 10 sectors by adding the price time series of the companies in the same economic sector, which corresponds to the price of buying one stock of each. This drastically reduces the number of free parameters $N(N-1)/2$ in the $N \times N$ correlation matrix from an order of 10^4 to exactly 45 to ensure that there is enough data to estimate each correlation, i.e. that the estimation is not underdetermined.

For each time step t and each pair of sectors i, j the local correlation coefficients

$$C_{i,j} = \frac{\left\langle r_t^{(i)} r_t^{(j)} \right\rangle_\tau - \left\langle r_t^{(i)} \right\rangle_\tau \left\langle r_t^{(j)} \right\rangle_\tau}{\sigma_\tau^{(i)} \sigma_\tau^{(j)}} \tag{4.3}$$

are calculated over a time period of $\tau = 42$ trading days like in [36] (the 42 trading days correspond to two trading months) with the local standard deviations $\sigma_\tau^{(i)}$. The window is shifted by one trading day each, resulting in overlapping windows, and hence, the window to calculate $\langle \dots \rangle_n$ in equation 4.2 is also shifted by one day. This approach is similar to the sector analysis in [116], where the correlations between the individual companies' stock time series are averaged within their GICS sectors. However, our procedure of first computing the GICS sector time series and then calculating the correlation ensures that larger companies are given more weight in the resulting correlation matrix than smaller ones.

It has to be stressed that the GICS sector classification has changed over the considered time period. For instance, when [116] was published, there were only 10 GICS sectors, whereas nowadays, there are 11. During the data period from 1992 to 2012, the exact composition of sectors and subsectors has also changed which poses additional problems in sorting a company into a sector. Here, the historical data is sorted into the sectors according to the GICS classification of the companies at the end of the time period in 2012, reflecting the same sectors as the ones used in [36].

Mean Correlation

The mean correlation $\bar{C}$ of the sectors correlations from equation (4.3) is calculated for each time t according to (4.1). Figure 4.2 shows the resulting time series for two different choices of $\tau = 42$ and $\tau = 5$ for the window width on which the pairwise correlations C_{imj} are calculated in equation (4.3). Of course, the time series based on the smaller windows is more noisy, but both time series show similar positions of peaks and the same long-term trend of increasing correlation. The time t is here selected as the central value of the time window of length τ in order to have a symmetrical window. The pre-processed time series is available via [142].

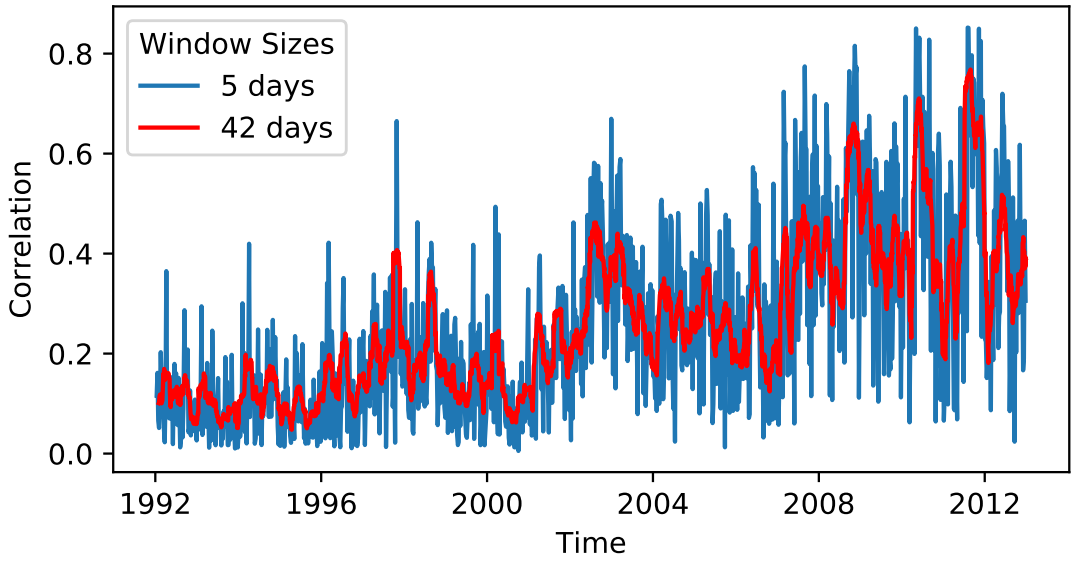

Figure 4.2: The mean correlation of the S&P500. The depicted time series are the mean values of correlation matrices, which were calculated based on moving windows of length τ days plotted against the centre of the τ-days-interval. This figure depicts $\tau = 5$ (red) and $\tau = 42$ (blue).

4.1.3 Outline of this Chapter

The analysis of the S&P500 data and its collective behaviour is split into two parts in this chapter. First, the correlation matrix of the GICS sectors is analysed with the help of explainable artificial intelligence (XAI) in order to better understand the differences between the market states in section 4.2. Second, the mean market correlation is modelled as a stochastic differential equation with non-Markovian memory effects which will be juxtaposed with the Markovian model suggested in the literature in section 4.3. Finally, a brief summary of the main results from both analyses will be presented in section 4.4.

4.2 Financial Market States Seen through the Lens of Explainable Artificial Intelligence

Understanding and forecasting changing market conditions in complex economic systems like the financial market is of great importance to various stakeholders such as financial institutions and regulatory agencies. Based on the finding that the dynamics of sector correlation matrices of the S&P 500 stock market can be described by a sequence of distinct states via a clustering algorithm, we try to identify the industrial sectors dominating the correlation structure of each state via XAI and a Bayesian change point analysis.

4.2.1 Introduction to Explainable AI for Market States

The first study on the dynamic behaviour of the market's correlation matrix in [36] identified eight market states via k-means clustering, qualitatively discussed the differences between the market states by comparing their centroids to each other and found different regimes of correlation structure: some states correspond to strong overall correlations throughout the market, whereas other states exhibit weak correlations outside of companies from the same industry sector or anticorrelations between some sectors. However, these comparisons were purely qualitative via plotting and comparing heat maps of the clusters' centroids and thus, there has not yet been a quantitative analysis of what differentiates these market states from each other. In this section, our goal is to identify the dominant factors within the different economic states in the framework of [36, 116, 117] by using a method [143] from the emerging field of XAI. Methods from XAI shed light on the black box that most AI methods seem to be from the user's per-

spective [80, 81]. XAI allows a quantitative ranking of how much a given observable has influenced the decision algorithm (here: clustering algorithm) and therefore gives us deeper insights into the differences between the market states than the analysis in [36]. Most importantly, this methodology can allow us to perform a dimensionality reduction by identifying the most important sector correlations without any qualitative ambiguity. XAI is nowadays considered for applications in physics [144, 145, 146] and finance alike [147] and in particular, financial regulators have already embraced the methods of XAI to increase transparency for customers with regards to AI-supported business decisions of financial institutions [78].

In an alternative approach to identify market states in [148], the market states are characterized and interpreted in terms of the clustered returns, whereas we consider the correlations of different market sectors. Such an analysis is similar to the qualitative analysis of the states in [36, 116, 117], but XAI can give us a more detailed analysis: The observables in [36, 116, 117] and in our study are the correlation matrices and therefore, the clustered variables are the correlations between different sectors. These correlations, however, are strongly interdependent; not only because of the economic real-world situation that they model, but also because of mathematical constraints of the correlation matrix (its eigenvalues are non-negative, meaning that its matrix entries cannot be chosen independently from each other; see [149] for the problems that can arise from this fact in practical applications). The XAI algorithm from [143] is well-suited to deal with such dependencies because of its holistic approach: it starts with the prediction of the k-means algorithm and goes backwards through the algorithmic structure until it arrives at the input layer, where it distributes the prediction relevance on all input variables (here: correlations) simultaneously. This means that it does not look at an isolated input variable in a vacuum, but always takes the values of other (potentially highly interdependent) variables into account.

Here, we use the k-means clustering method to identify the clusters corresponding to market states. While most of the above clustering analyses use a hierarchical k-means clustering, we use regular k-means clustering. This has a big advantage because it can be reformulated to allow the use of XAI which can analyse the decisions of the clustering algorithm [143]. This allows us to reveal the dominant factors of the different market states and to compare the market states quantitatively, thereby improving the qualitative comparison depicted in figure 2 of [36]. Moreover, as shown later in the chapter, the XAI selection allows for a dimensionality reduction of the system by selecting the most relevant correlations.

The remainder of this study is structured as follows: Section 4.2.2 introduces the clustering method and the XAI method before interpreting the XAI results. Also, it discusses a reduced model which only uses the features (i.e. sector-sector correlations) with the best XAI values to learn the clustering classification. This reduced model is used to test the usefulness of XAI for reducing the dimensionality of the data. Finally, section 4.2.3 summarises the results of these methods, compares them to previous works and suggests follow-up research. The appendix includes further information on the XAI algorithm in appendix D.1, some additional figures of results in appendix D.2 and technical details on neural networks in appendix D.3.

4.2.2 Identifying the most relevant sector correlations within a market state

In this section we first give an overview of the clustering method with further technical details explained in the appendix D.1, before using XAI to shed light on how the clustering algorithm arrives at its results. Then, the XAI results are discussed and interpreted. Finally, a reduced surrogate model that only uses the highest-XAI features is used to see whether the reduction to features with high XAI preserves most of the clustering algorithm's information.

Identifying Market States via K-Means Clustering

Because the analyses [36, 116, 117] unanimously revealed eight market states for the time period analysed in this chapter, we set the number of clusters for the k-means algorithm to eight. In contrast to the procedure in the literature, regular k-means clustering is used instead of a hierarchical binary k-means clustering in order to make the cluster centres amenable to the XAI method suggested in [143]. The regular k-means clustering method starts with a predefined number of k clusters and randomly selects k initial cluster centres (also called centroids). It then iteratively assigns each daily correlation matrix to its nearest cluster centre and updates the centres by setting them as the cluster's centre of mass . Assignment and calculating the new centres are repeated until the centroid positions converge [150]. The market correlation matrix $\mathbf{C}$ belongs to cluster j, if for the cluster centroids $(\mathbf{C}_l)_{l=1,\dots,k}$ the inequality

$$||\mathbf{C} - \mathbf{C}_j|| < \min_{l \neq j} ||\mathbf{C} - \mathbf{C}_l|| \tag{4.4}$$

holds for a suitable distance metric $||\dots||$. Here, the Euclidean distance is used.

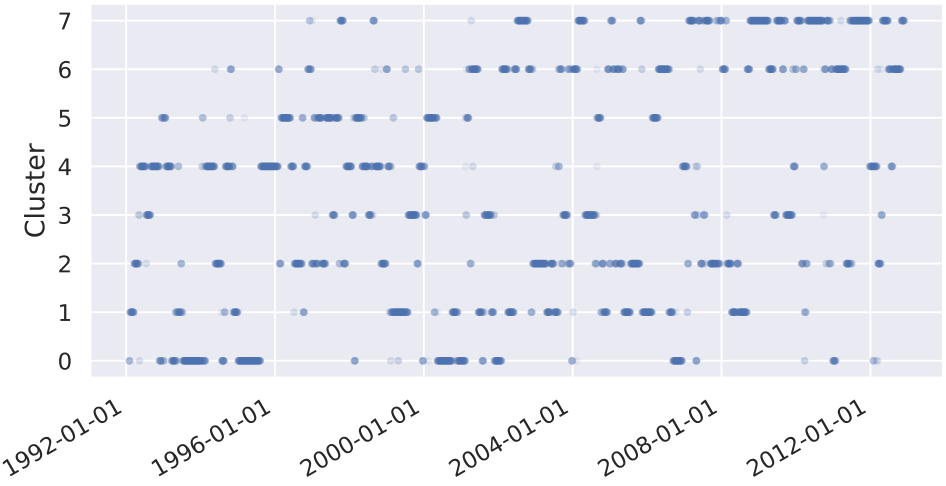

Figure 4.3: The dynamics of which cluster state is occupied by the market at which point in time. Individual data points are depicted with low opacity and the slight overlap between adjacent points increases the opacity and may appear like a solid line.

The time series of the resulting clusters is depicted in figure 4.3 and shows a qualitatively similar behaviour to the clusters in [36, 116, 117]. In the observed time period, some clusters cease to appear more than sporadically in the data after some years (clusters 0, 1 and 5) and some only start to appear more than sporadically years after the start of the observation time period (cluster 6, 7 and again 5). However, figure 4.3 shows some quantitative differences compared to the aforementioned literature, in particular because the clusters in figure 4.3 often appear isolated and for short time periods apart from their main emergence interval. This may be attributed to the slightly different clustering technique used for this chapter and the three clustering analyses [36, 116, 117] also show differences in their plots of clusters against time.

Explainable AI for Clustering

We now want to explore whether the clusters are dominated by certain industrial sectors and, if this is the case, which sectors are significant for the identification of the market states represented by the clusters. To this end we use the method developed in [143] to analyse k-means through the lens of XAI. Therefore two steps are necessary [143]: First, the clustering classifier has to be neuralised, i.e. rewritten as a neural network. Note that the network is not trained, but its weights are derived from the centroids

estimated via k-means. Second, an XAI method for neural networks is used for analysing the neuralised clustering algorithm, namely the layer-wise relevance propagation (LRP) [151]. For each instance (data point), the XAI algorithm estimates how much the values of its features contributed to the neural network's decision for this instance. It answers the question which particular feature values led to $\mathbf{C}(t)$ being sorted into cluster j instead of any other cluster. The contribution of the i^{th} feature is quantified in a relevance score ρ_i which is calculated according to equations (2) and (4) from [143] for each instance with respect to its cluster assignment j for all features i. Appendix D.1 gives further details on the implementation of this algorithm.

From Local to Global Relevances This method produces relevance scores ρ_i for each feature i (for the correlation between sectors) for every instance (i.e. for each trading day), as depicted in the inset of figure 4.4 for the first day of cluster 2. However, we wish to focus on the aggregated relevance of one specific feature to understand which sector correlations are more important for the market state than others. As an example, figure 4.4 shows the histogram of XAI relevances for the correlation Energy/Materials for all trading days in cluster 2. Are such values particularly high? Does the correlation between Energy and Materials matter more for cluster 2 than other correlations? To answer such a question, one has to use the XAI values of each instance (local values) for a global aggregation to compare the aggregated influence of the Energy/Materials correlation with other correlations. Using the mean relevances of each sector seems reasonable, but the mean is distorted by some instances with very high absolute values for some of their relevances. Hence, one can instead focus on the median of each feature's relevance for each cluster's instances. Alternatively, one can identify each instance's feature with the highest ρ_i and then concentrate on the features that are most frequently the most important feature of an instance within each cluster. Because it regards the statistical mode (i.e. the most frequently observed value) and calculates the mode of the mode of each instance's ρ_i, we chose to name this method "mode-mode". The next sections addresses the question, which of these aggregation methods is more suitable for the data.

Determining Relevant Correlations For each individual cluster, the XAI relevances of each instance are aggregated to analyse the importance of any feature for the clustering decision. The median method can be represented by the median $\hat{\rho}_i$ and the mode-mode method by how often (n_i times) the i^{th} feature is the most relevant feature of an instance. Both the $(\hat{\rho}_i)_i$ and $(n_i)_i$ values are sorted ascendingly as shown in a sample plot in figure 4.5. These plots can be approximated by piecewise linear segments and sometimes, they

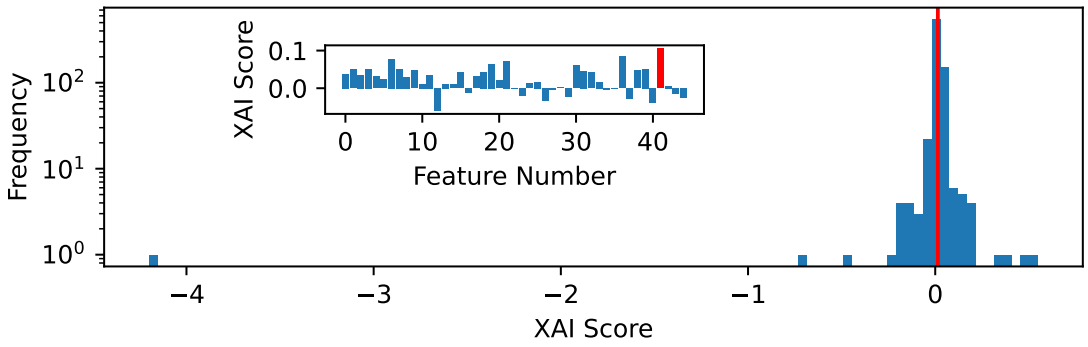

Figure 4.4: Main figure: Histogram of the XAI importance of the correlation Energy/Materials for all instances (trading days) in cluster 2. Outliers like the one visible on the far left of this histogram distort the mean. The median value is depicted in red. Inset: Barplot of the XAI importance scores for the the first day in cluster 2, showing how much each feature contributed to the classification. The mode (feature with highest XAI score) is depicted in red.

show a characteristic elbow shape: They start at a long plateau of low values and later experience a drastic increase, resulting in an almost rectangular shape (cf. figure 4.5, right). This shape allows to divide the features into two groups. Therefore, a change point analysis can identify the most drastic change in the slope of these plots and help us to achieve a clear threshold between important features with high $\hat{\rho}_i$ or n_i and features with little importance [55]. Here we use the Bayesian change point analysis implemented in the *antiCPy* package [111] based on the ideas in [152].
Like depicted in figure 4.5, the mode-mode values usually show a much more pronounced elbow that makes change point selection easier and their estimated change point is often at a much higher feature number than for the median method. Therefore, if features on the right-hand side of the change point are to be interpreted as highly relevant features, then the mode-mode method allows for a more exhaustive feature reduction than the median method.

We use the mode-mode method to differentiate between relevant and irrelevant correlations, because the approach usually leads to a more parsimonious feature reduction (i.e. higher interpretability) and often much clearer change points (cf. elbow-like relevance plot in the right subfigure of 4.5). Note that a relevant correlation is not necessarily unusually strong or weak, but rather particularly useful to differentiate between this cluster and the others. As an example, figure 4.6 shows for cluster 2 the centroid (left), the aggregated XAI values (centre) and the change point selection of relevant features (right). The relevant features cannot be detected trivially by, e.g. just identifying par-

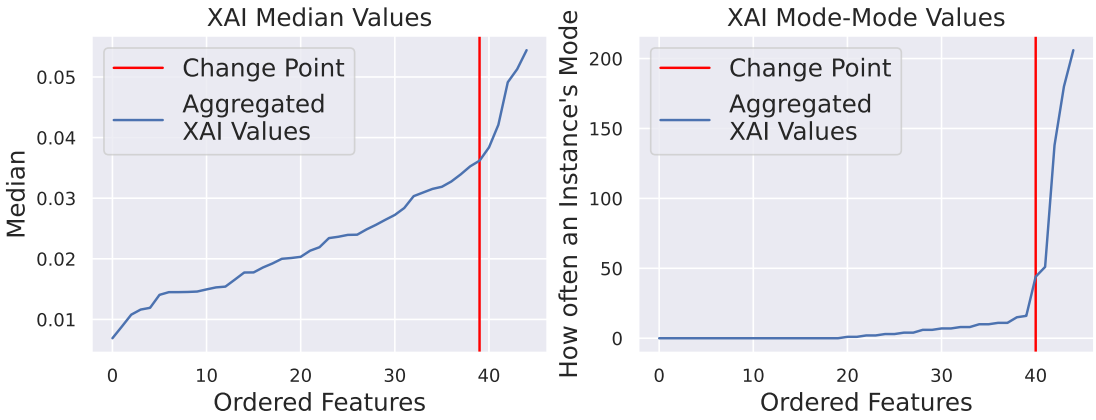

Figure 4.5: Comparison of ascendingly plotting the features according to the median values XAI (left) and mode-mode XAI (right) in blue and their respective change point estimation with *antiCPy* in red. Plots for the other clusters are included as supplementary material in the appendix D.2.

ticularly high or low values in the centroid, but follow from a comparison to the other centroids. Hence, the change point analysis of relevant values does not follow trivially from the cluster centroids, but uncovers new information about the data and yields a parsimonious selection of relevant economic sector correlations.
A plot of all relevant correlations for each cluster is shown in figure 4.7. The correlations of the Energy sector appear in states associated with clusters 1 and 6, the Utilities sector in 1 and 4 and the Telecommunication sector in 3 and 4. The IT sector dominates cluster 5, influences cluster 0 and, alongside Telecommunication Services, also influences cluster 2. States in cluster 7 only have a high relevance for the correlations between Energy/Financial Services and IT/Telecommunication. Other clusters show a high influence of Utilities and Telecommunication Services, which are both tied to the Energy or IT sector, respectively: the Utilities sector includes companies which distribute gas and electricity, while Telecommunication Services are related to the physical infrastructure of information transportation.

Dimensionality Reduction for all Clusters

It makes sense to assume that features with a high relevance should be particularly useful for predicting the original classifier. In particular, this might reveal that only few features are necessary to predict the majority of the cluster assignments. If we only wish to classify one cluster vs. all the others, then the relevant features selected by the cutoff in figure 4.5 are a reasonable choice for reducing the dimensionality of the data.

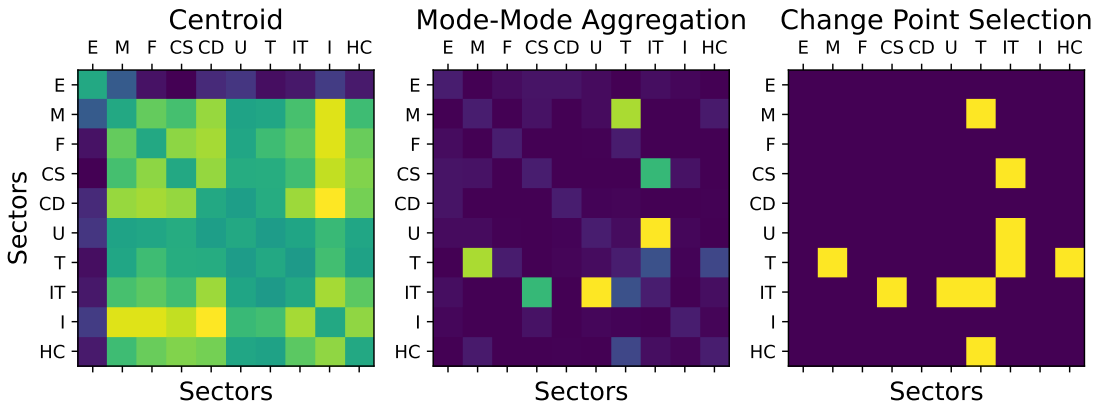

Figure 4.6: Visualisation of different matrices that represent cluster 2: Comparison between the cluster's centroid, the mode-based aggregated XAI values and the binary antiCPy-selected relevant XAI values. Brighter colours denote higher values and the left and central figure show continuous colours while the right figure is binary. Because the diagonal's correlation values are always 1, their values have been set to the matrices' averages. Table 4.1 explains the sector abbreviations.

But for an all vs. all classifier, choosing all relevant features of all clusters would barely reduce the number of features at all, because clusters 1 and 4 (cf. figures D.3 and D.5) would already each contribute more than ten features to this list of relevant features. Instead, one can restrict the analysis to the single best feature for each cluster, resulting in eight features instead of the original 45. This also constitutes a second chance to compare the mode-mode selection to the median selection. If a model trained on such a dimensionality reduction predicts most of the cluster assignments correctly, it will serve as a self-consistency check by indicating that features with high XAI relevance are also highly useful for predicting.

To test whether a dimensionality reduction is possible, we train a neural network as a surrogate model that should predict the original clustering decisions, but restrict its data to only some of the original features (cf. [80, chapter 8.6]). For each of the eight clusters, only the feature that has the highest XAI relevance score is used to train a neural network as a surrogate model. To check if this surrogate model has a high accuracy with respect to the original classifier, we also train neural networks with eight features randomly selected from the leftover features. Accuracy is defined as the percentage of correctly predicted cluster labels. A short introduction to neural networks can be found in section 2.2.5 and details on the architecture of the used neural network are given in the appendix D.3. Training, test and validation data each contained $\frac{1}{3}$ of the full data. The networks

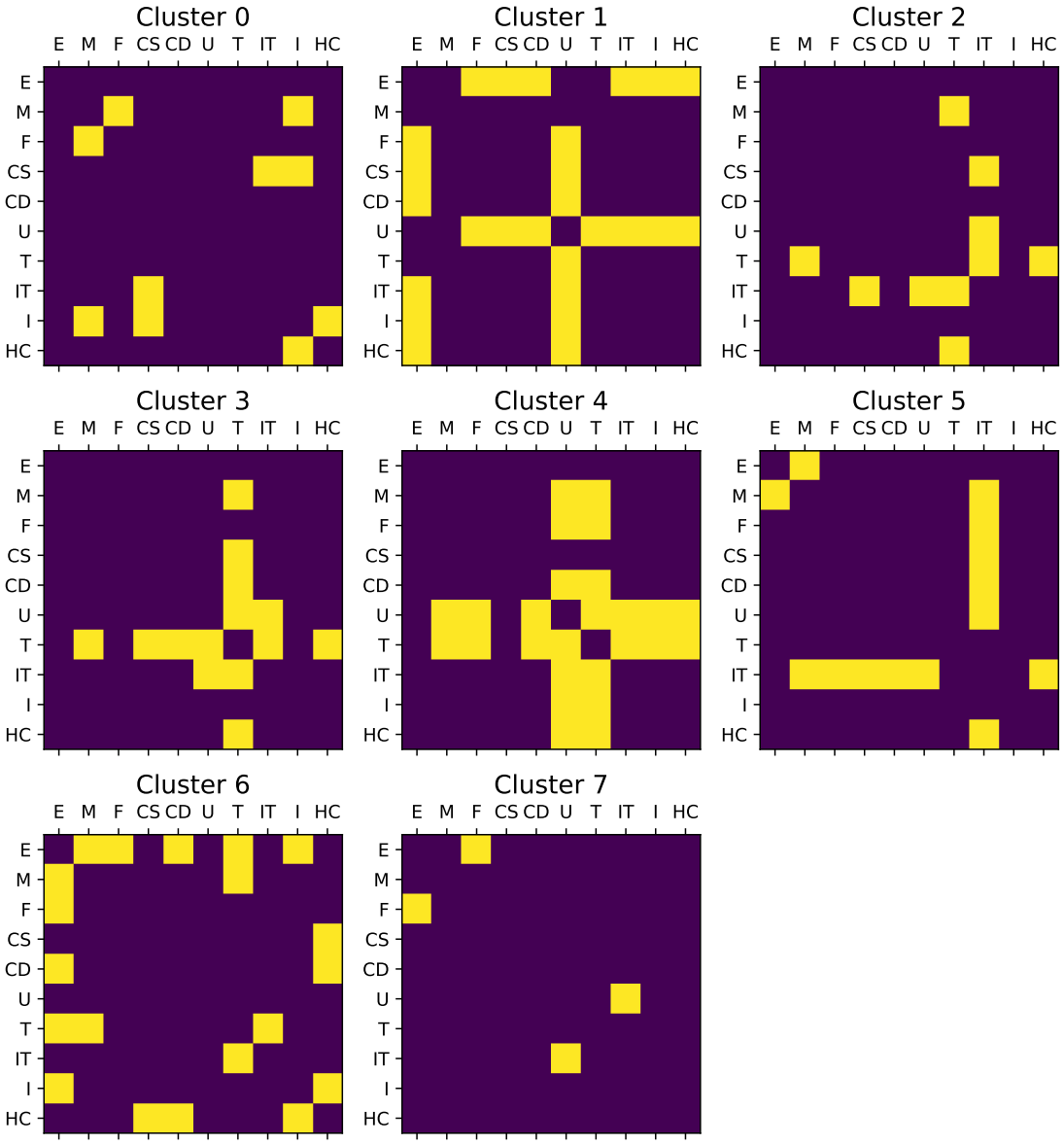

Figure 4.7: Relevant correlations (yellow) of the sectors for the 8 clusters in figure 4.3. Relevance selection was done via the change point analysis explained in section 4.2.2 on the mode-mode XAI values. Note that the diagonal has been set to zero relevance by default because its correlations are always 1. Table 4.1 explains the sector abbreviations.

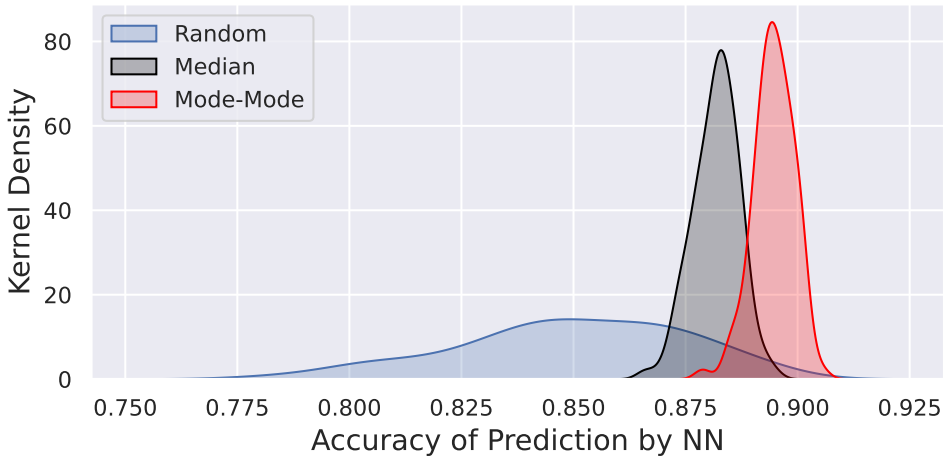

Figure 4.8: With the eight mode-mode features, the eight best median features and eight random features, 100 neural networks were trained each with a different seed. Kernel density estimations show that the mode-mode and median networks achieve a much higher accuracy than the randomly selected suboptimal features.

were trained on (a) the eight optimal mode-mode features, (b) the eight optimal median features and (c) eight randomly selected features. For each group, 100 networks were trained, each of which was initialized with a different random seed to marginally change the performances in the groups (a) and (b). Also, each network in group (c) has a different selection of random features. This results in a distribution of network accuracies for each of the three groups. The eight correlations used for the mode-mode method are: E/M, E/F, E/HC, M/U, F/IT, CS/IT, U/T and U/IT (see table 4.1 for the abbreviations). The accuracy of predicting the original clustering label out of the eight clusters was chosen as a goodness criterion for the different feature selections. The results are depicted in the kernel density plots in figure 4.8, which are smoothened histograms of the 100 accuracy values for each neural network. The random feature selection achieves noticeably worse results than the two XAI based selection methods. The mode-mode ($\mu_{\text{mode}} \pm \sigma_{\text{mode}} = 0.895 \pm 0.004$) is slightly better than the median method ($\mu_{\text{median}} \pm \sigma_{\text{median}} = 0.882 \pm 0.005$) and their confidence intervals (CIs) do not overlap at the 1σ level. Combined with its more parsimonious XAI selection for the individual clusters as demonstrated in figure 4.5, the mode-mode criterion is shown to be the more useful criterion.

4.2.3 Intermediate Discussion

The goal of the research presented in this section was to answer the question whether it is possible to identify sector correlations that dominate the assignment of a trading day to a market state. To this end we used a clustering procedure that allows for an interpretation by a method from the field of XAI. We designed an aggregation procedure to combine these local explanations to global feature relevances for the entire data. The resulting relevance curves show pronounced elbow plots for the mode-mode aggregation like in figure 4.5 which can be analysed with a change point identification method to define cutoffs between irrelevant and relevant features.

Using the described methodology, it is possible to show that the assignment of a certain trading day of the S&P 500 market to a cluster (i.e. to a market state) is dominated by only a few sector correlations. The correlations involving the Energy or IT sector each make up three of the eight best features selected by the mode-mode XAI aggregation, respectively, but E/IT is not one of the eight best mode-mode features. Note that in figure 4.7, clusters 1, 2, 5 and 6 also show a high influence of IT and/or Energy. The frequent occurrence of the IT sector among the highly relevant correlations may be traced back to the fact that the observation period includes the build-up, the emergence and burst of the dot-com bubble which saw the rise and rapid decline of many IT companies as shown in figure 4.1. This is also reflected in the analyses in [36] and [117]. The high importance of the energy sector implies that this sector has major influence in determining whether the economy is in a state of crisis or growth, probably because all other economic activity requires energy, e.g. for transportation or production. While the high importance of the IT sector might have to be attributed to the dot-com bubble, the continued influence of the energy sector on the economic state might be a more reasonable conclusion from this analysis. These results support previous studies that have identified the energy sector as a key influence in economic activity [153, 154, 155].

We have also shown that a neural network which is only trained on the features with the highest XAI relevance scores can predict the trading day's cluster with a high degree of accuracy. In general, our results imply that a further dimensionality reduction for the description of the dynamics of the financial market is possible. This can be particularly useful in high frequency trading applications, where the time span of interest lies within the life expectancy of a cluster. In this case, when data needs to be processed very quickly, it may be too slow to work with the complete correlation matrix. Nevertheless, our findings suggest that a reduced matrix only containing the most informative correlations for this cluster can lower the computation times without sacrificing too much accuracy.

Note that the quantification of each feature's relevance was only possible due to the holistic nature of the used XAI algorithm.

Comparison to other Clustering Algorithms

Our analysis follows the work in [36, 116, 117] and treats the correlation matrix of the returns as the observable that is to be sorted into clusters (market states) via k-means. A different approach to the identification of market states is given in [148], where the clustering algorithm in the maximum-likelihood-framework of [156] is used. The author of [148] not only uses the clustering algorithm to identify market states, but also to identify economic sectors. These show strong overlap with the sectors in the SIC classification which have been used as industry standards and recently been replaced by the GICS sectors. Because of this strong overlap, we use the GICS classification as a starting point for our analysis. In [148], daily prices from 1990 to 1999 have been analysed, whereas our work uses data from 1992 to 2012 in line with [116, 117] and therefore has more recent data. While [148] gauges the importance of business sectors for the market states by showing selected scatter plots of the companies' average returns for two different market states, we use an all vs. all comparison of all market states with the methodology of XAI. Despite the increased time frame and the different approach to clustering, our analysis confirms some of the key findings in [148], namely the pronounced role of the High-Tech companies (roughly corresponding to the IT sector in our analysis) and the sectors of Oil & Gas and Petroleum companies (corresponding to the Energy sector). While [148] also finds a peculiar behaviour of the Gold & Silver mining companies, our analysis does not have such a high resolution of detailed business sectors and instead these companies are part of the Materials sector. While [148] does not find any peculiar behaviour for the Finance sector (or Commercial Banks in [148]), the Finance sector appears twice among our top 8 correlations, indicating an above average importance. This probably corresponds to the fact that our data, unlike the time interval used in [148], includes the 2008 financial crisis.

An interesting difference between our study and the analysis of market states in [36] can be found by regarding the last state (8 in [36] and 7 in our chapter): the qualitative state comparison in [36] simply finds an overall high degree of correlation across the different sectors during this state, whereas our analysis still identifies that the two sector correlations Energy/Finance and Utilities/IT (cf. figure 4.6) have a dominant effect on this cluster. The identification of these two dominant correlations was only possible due to the quantitative approach of the XAI methodology.

Outlook on Market States

While this chapter used a standard k-means clustering algorithm and therefore deviated from the hierarchical clustering in [36, 116, 117], this opens up another possibility for follow-up research by combining these methods: using a hierarchical clustering algorithm and the XAI on each split in the hierarchy may increase our understanding even further via XAI as an addition to the intrinsic interpretability of hierarchical algorithms. This would potentially tie together the explainability of XAI with the paradigm of interpretable models presented in [85].
Despite the insightful analyses of financial correlation matrices in [36, 116, 117] and in this chapter, one has to keep in mind that the analysis of the states is based on the matrix of Pearson's correlation coefficients between financial time series. Although this method is frequently used in financial research [34, 35, 125, 129], it has been subject to criticism: As noted in [157], Pearson's correlation only provides information on linear relationships, but not on nonlinear correlations. This problem is related to the stylised fact that returns have no autocorrelation, but squared (i.e. nonlinearly transformed) returns show a strong autocorrelation [18]. Instead, Spearman's correlation may provide another view on correlations and include nonlinear effects [158] and therefore can be used as another starting point for follow-up research. Additionally, as already mentioned in section 4.1.2, the GICS classification has been updated over time to reflect changes in the economy's structure. Because these updates also happened during the regarded 20 years time period, they may have affected the results of this analysis. Moreover, there is a more subtle problem about the data that can be referred to as a survivorship bias: Like [36, 116, 117], we only regard companies whose stock prices are available for the entire time period. Companies that went bankrupt or were merged into another company during the time period are therefore completely disregarded by these analyses, although their bankruptcies or takeovers certainly contain valuable information about the state of the economy.

4.3 Memory Effects and non-Markovianity in the Market Mode

In this section, we analyse the mean market correlation of the S&P500, which corresponds to the market mode in principle component analysis. We fit a generalised Langevin equation (GLE) to the data to uncover memory effects in the market dynamics. Moreover, a Bayesian resilience estimation (carried out by Martin Heßler in [124]

and discussed in his PhD thesis [159]) provides further evidence for non-Markovianity in the data and suggests the existence of a hidden slow time scale that operates on much slower times than the observed daily market data. Because of the connection between these analyses, this thesis will briefly sketch the main methods and results behind the resilience estimation while leaving the details to [159].

4.3.1 Langevin Equations in Finance

As explained in the introductory section 1.3, the "market mode" (the largest eigenvalue of the correlation matrix) reflects the collective behaviour of the market and is usually given by the mean correlation of all time series of the system [34, 35]. The study in [117] models the mean market correlation as a stochastic differential equation. Its results suggest that it is justified to focus on the mean market correlation as a low-dimensional observable which nevertheless contains much of the information about the market dynamics.

Ever since Bachelier's seminal work introduced the random walk model to describe stock prices [26], the inherent stochasticity of financial data has been taken into account by researchers and practitioners alike. The Langevin equation is a model for stochastic differential equations that includes a deterministic drift and a random diffusion and can be used to describe such stochastic data. Much research has been devoted to estimating Langevin equations from data [105, 106, 108, 122, 160, 161, 162] and Langevin equations have found applications in various fields of research such as fluid dynamics [163], molecular dynamics [164] and meteorology [165] (see [58] for a review of applications). The generalised Langevin equation (GLE) expands the Langevin model from [117] by introducing a kernel function to include memory effects [166].

The remainder of this section is structured as follows: In 4.3.2, the GLE estimation procedure is explained in detail and additionally, a rough sketch of the resilience estimation is given. In 4.3.3, the goodness-of-fit and the estimated memory parameters for the GLE model are evaluated and the results of the resilience estimation are briefly touched upon. Finally, the findings of these analyses are interpreted and compared to the results from [117] in dection 4.3.4.

4.3.2 Time Series Models for the Mean Market Correlation

Fitting a Generalised Langevin Equation with Memory Kernel

Sections 4 and 5 of [117] analyse the one-dimensional mean correlation time series by fitting a Langevin model

$$\frac{\mathrm{d}\bar{C}}{\mathrm{d}t}(t) = D^{(1)}\left(\bar{C}, t\right) + \sqrt{D^{(2)}\left(\bar{C}\right)}\Gamma(t) \tag{4.5}$$

with independent Gaussian noise Γ, a deterministic time-dependent drift function $D^{(1)}$ and a time-independent diffusion function $D^{(2)}$. Note that the Langevin equation is Markovian, i.e. it has no memory. As an alternative model, a Generalised Langevin Equation (GLE) includes previous values of the time series in a memory kernel $\mathcal{K}$, but has only time-independent parameters with a functional equation

$$\frac{\mathrm{d}\bar{C}}{\mathrm{d}t}(t) = D^{(1)}\left(\bar{C}\right) + \int_{s=0}^{t} \mathcal{K}(s)\bar{C}(t-s)\mathrm{d}s + \sqrt{D^{(2)}\left(\bar{C}\right)}\Gamma(t). \tag{4.6}$$

A Bayesian estimation of the parameters of Equation (4.6) is implemented in [121]. It discretises the memory kernel $\mathcal{K}(k)$ to $\mathcal{K}_k$ and discretises the observed data into n_B different bins to significantly speed up the estimation procedure. Because an overlap of the windows used to calculate $\bar{C}$ can lead to artefacts in the memory effects (cf. Appendix E.1), we choose to calculate $\bar{C}$ with different parameters than $\tau = 42$ and $s = 1$ from [117] (the mean correlation of every sth day is retained for the analysis; hence, for $s = 1$, the full time series is used). Instead, we set $\tau = 5$ and $s = 5$, meaning that the mean market correlation $\bar{C}$ is calculated over one trading week ($\tau = 5$ trading days) for the end of the trading week in disjoint windows (i.e. we create a time series whose length is $1/s = 20\%$ of the original time series). Because using disjoint intervals automatically leads to a thinning of the time series, this seemed like a useful trade-off with a real-world interpretation as weekly correlation. It is worth noting that while the calculation of $\bar{C}$ only considers data from five trading days, the correlation matrix $C_{i,j}$ encompasses approximately 10^4 correlation values. Even though each (i, j) correlation pair is determined over a brief time window, the large ensemble of $C_{i,j}$ values, with their average being $\bar{C}$, underscores our confidence in the accuracy of the mean $\bar{C}$. The resulting time series is shown in figure 4.2 together with the time series used in [117] and although it obviously becomes much noisier due to the shorter window size, it still retains some of the features of the less noisy time series such as, e.g. the position of spikes with high correlation.

As described in [121], the memory effects are aggregated to a quantity K_k (named κ_k in [121]) that measures the strength of all memory effects from 1 up to k time steps ago. If K_k saturates towards a plateau at step k_0 or starts to decrease, then k_0 is the maximum length of any reasonable memory kernel. This often also coincides with kernel values of $\mathcal{K}_{k_0} \approx 0$ at k_0. The estimated values for the model parameters are chosen by calculating the marginal posterior distribution, exploring it with MCMC and choosing the mean or maximum a posteriori (MAP) parameter estimation. The goodness-of-fit is also evaluated by using MCMC to simulate an artificial time series $\bar{C}_t^{(a)}$ with the estimated model and comparing the autocorrelation function and the distributions of $\bar{C}_t^{(a)}$ and the increments $\bar{C}_t^{(a)} - \bar{C}_{t-j}^{(a)}$ to those of the original data. Finally, the inferred model structure can be trained on only a subset of the data (training data) and be used to predict the remaining test data to evaluate its predictive performance.

Resilience Estimation

Although this is beyond the scope of this thesis and based on the work carried out by the second main author of [124], the results of the resilience estimation in [124] relate to the memory effects analysed in this section. Hence, a short introduction to modelling resilience in non-Markovian stochastic differential equations will be presented here. More details will be provided in [159].

Using the approach in [122], the Langevin model of the mean correlation time series from (4.5) is extended by proposing a hidden process λ

$$\begin{aligned} \frac{\mathrm{d}\bar{C}}{\mathrm{d}t}(t) &= D_C^{(1)}\left(\bar{C}, t\right) + \sqrt{D_C^{(2)}\left(\bar{C}\right)}\lambda(t) \\ \frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) &= D_\lambda^{(1)}\left(\lambda, t\right) + \sqrt{D_\lambda^{(2)}\left(\lambda\right)}\Gamma(t). \end{aligned} \tag{4.7}$$

For both processes $\nu \in \{\bar{C}, \lambda\}$, a characteristic time scale is defined via

$$\tau_\nu = \left| \left(\frac{\mathrm{d}D_\nu^{(1)}(\nu)}{\mathrm{d}\nu} \right)^{-1} \right|_{\nu=\nu^*} \tag{4.8}$$

around the respective fixed point ν^*. Also, the stability of the observed process $\bar{C}(t)$ can be gauged via the stability parameter

$$\zeta = \frac{\mathrm{d}D_C^{(1)}(\bar{C}}{\mathrm{d}\bar{C}}\bigg|_{\bar{C}=\bar{C}^*} \tag{4.9}$$

according to linear stability analysis such that $\zeta < 0$ corresponds to a stable process [1]. As the theory of locally stable and quasi-stationary market states discussed in Stepanov et al. [117] assumes the stability of the observed process $\bar{C}(t)$, the resilience estimation should yield a model with $\zeta < 0$.

4.3.3 Uncovering Memory Effects in the Data

Estimated Generalised Langevin Equation

To select a model with a reasonable kernel length, we first fit a model with a rather large kernel length and evaluate its combined memory aggregation K_k. The time lag k_0 for which K_k no longer increases is then chosen as an upper bound of the memory kernel length for the subsequent model fitting. The data are split into $n_B = 10$ equally wide bins and an initial modelling attempt with a rather long kernel $k_{\max} = 10$ is tested, i.e. $\mathcal{K}_q = 0$ for all $q > k_{\max}$. It shows that the aggregated memory stops to increase after $k = 6$ in figure 4.9. Hence, kernels with $k > 6$ do not meaningfully contribute new information and only models with $k_{\max} \leq 6$ are a reasonable choice. We fit a model with $k = 6$ for the memory kernel to evaluate its goodness-of-fit.

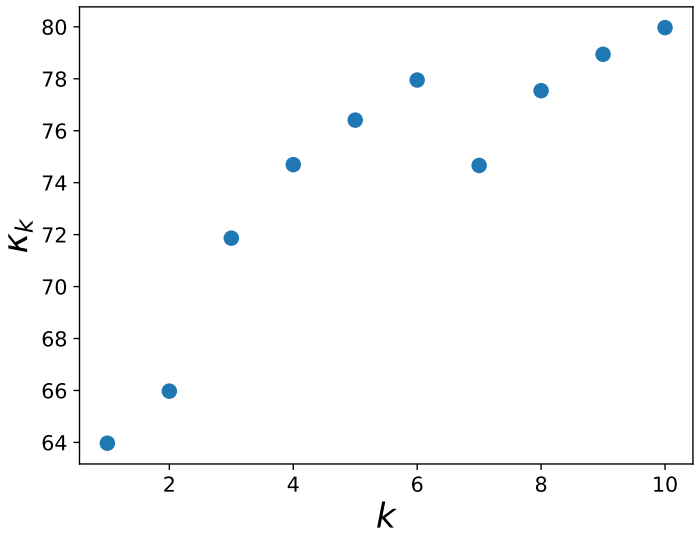

Figure 4.9: For a model with a long kernel $k_{\max} = 10$, the combined memory effect K_k decreases after $k = 6$ and no longer increases meaningfully afterwards.

Goodness-of-Fit A time series of length 10^5 is simulated via Euler–Maruyama integration of the GLE model to compute the autocorrelation function (ACF) from equa-

tion (2.10) and to compare it to the ACF of the original time series. The best model is chosen via MAP estimation in the Bayesian framework of [121], but selecting the mean estimation yields almost identical results. Both ACFs are computed via the function *statsmodels.tsa.stattools.acf* from the Python package *statsmodels* [167]. Figure 4.10 shows that the two ACFs show very good agreement up to lags of 10 trading weeks and decent agreement up to lags of 20 weeks. The distributions of the time series and the first two increments for the original and the simulated data are shown in figure 4.11 for the MAP estimation and also for the mean estimation. Figure 4.11 shows an almost perfect overlap between the two estimation procedures and a good agreement between the estimated time series and the original time series, especially for the increment distributions. Overall, these diagnostics indicate that the estimated model with memory kernel length 6 manages to reproduce these important statistical properties of the original time series.

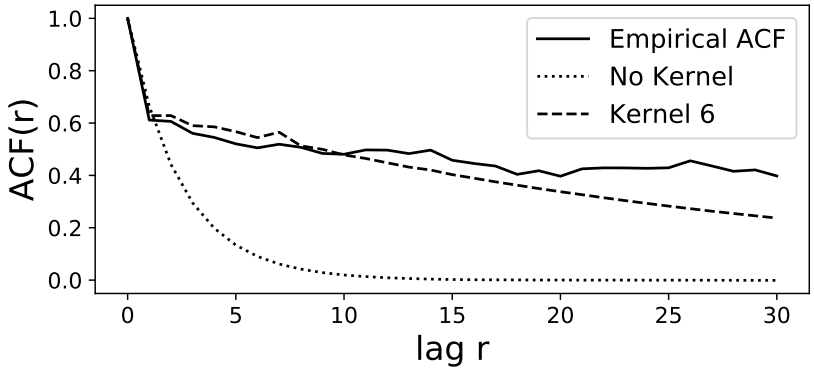

Figure 4.10: Comparison between the autocorrelation functions of the original time series (solid line) and the simulated data from the model with kernel length 6 (dashed line) simulated via MAP estimation of the MCMC-integrated marginal densities. Up to lag 30, the ACFs show good agreement. The alternative simulation via mean estimation of the density yields an almost identical ACF. However, the ACF based on the model with no memory kernel (dotted line) fails to capture the empirical ACF.

Because the realised values of the time series $\bar{C}$ are not distributed uniformly, we also use a modelling procedure with unequal bin widths so that each of the 10 bins has the same amount of data. However, this barely changes the model diagnostics shown in figures 4.10 and 4.11. The only noticeable change is that for long kernels with length ≥ 10, the ACF seems to be captured a little more poorly in the shown range of figure 4.10, but manages to fit more closely to the empirical ACF for large lags at around $r \approx 100$.

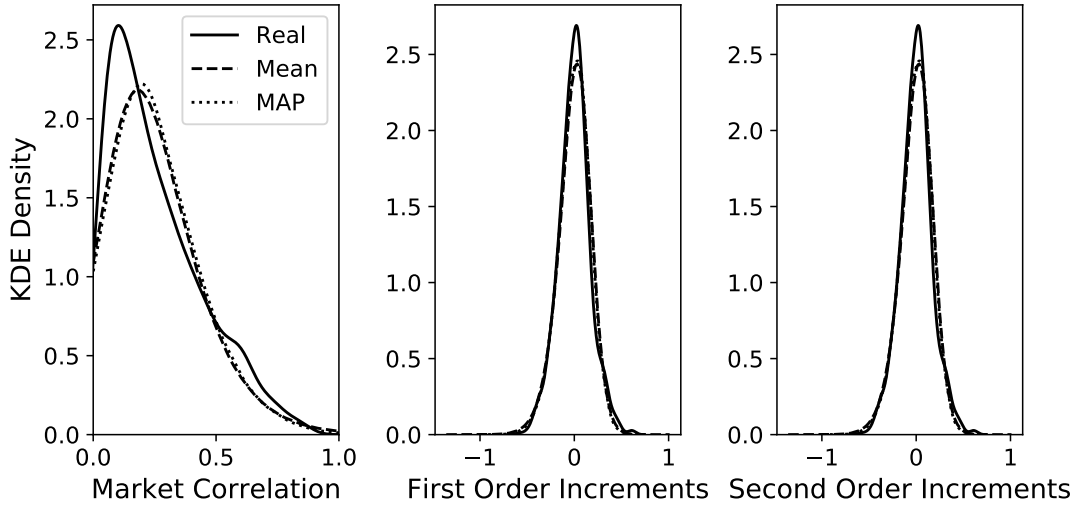

Figure 4.11: KDE plots to compare the distributions between the real data and the MCMC samples for the two estimated GLE models with kernel length 6 for the time series data $\bar{C}_t$ (left), the increments $\bar{C}_t - \bar{C}_{t-1}$ (centre) and $\bar{C}_t - \bar{C}_{t-2}$ (right). Differences between the two simulated models are hardly visible and overall, there is a good overlap with the real data. For the LE without memory, the respective distributions are depicted in Appendix E.2.

While the regular LE without memory has a very similar increment distribution, the ACF is considerably less accurate than the ACF of the GLE.

Estimated Memory Kernel To make further inference on the memory kernel and to estimate its quantitative effect, the posterior distribution of the model with memory length 6 is sampled via MCMC. 100 walkers are simulated for 10^5 time steps and after a burn-in period of 450 initial time steps is discarded, the chains are thinned by only keeping every 450th step to obtain uncorrelated samples. Bayesian CIs can now be calculated at the 95% level for the associated parameter $\mathcal{K}_k$ of each time lag $k \leq 6$ in the kernel function and the results are shown in figure 4.12.

The CIs of $\mathcal{K}_4$ are already close to the zero line and the CIs of $\mathcal{K}_5$ include zero, meaning that there is no evidence for a memory term at five weeks distance. However, the results in figure 4.12 clearly imply a nonzero memory effect for four time steps with a clearly nonzero effect strength for memories up to $k = 3$ trading weeks. Hence, $k = 3$ will be used as a maximum kernel length for the prediction via the GLE.

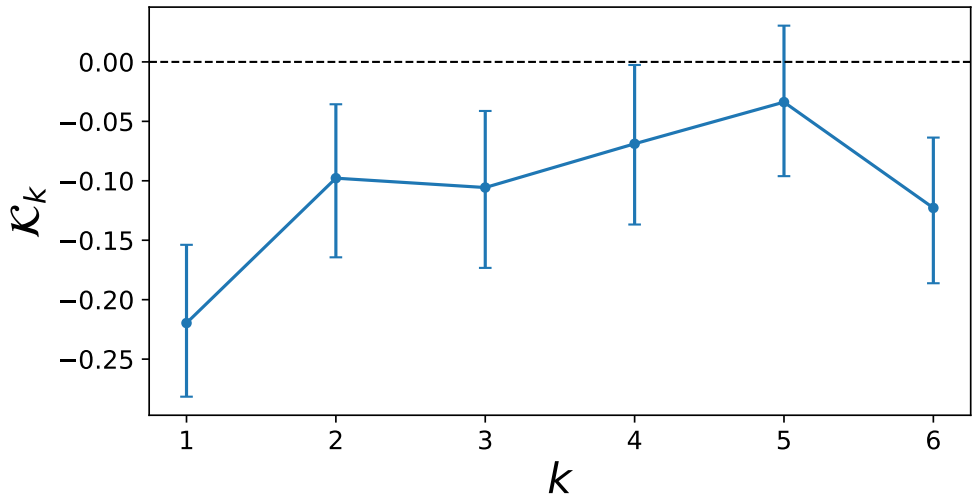

Figure 4.12: Values for the memory kernel $\mathcal{K}_k$ in the model with memory length 6. Credible intervals were estimated via MCMC at the 95% level. The inclusion of 0 in the credible interval of $\mathcal{K}_5$ implies that the nonzero memory effect for $\mathcal{K}_6$ may be misleading. All previous memory kernels $\mathcal{K}_{1,2,3,4}$ have nonzero values at the 95% credibility level.

Prediction via the GLE with Kernel Length 3 With a conservative interpretation of the results from the previous section, a model with memory length $k = 3$ is used to evaluate the GLE's power to predict a future value y_{t+1} by forecasting an accurate prediction $\hat{y}_{t+1}$. It is tested against a regular Langevin equation (LE) model without any memory effects (corresponding to $k = 0$ and also estimated with the code in [121]) and against the naive benchmark of predicting the next time step y_{t+1} by simply setting it to the last previously known value: $\hat{y}_{t+1} = y_t$. To test these three methods, the Langevin models are trained on the first α% of the time series (the training data) and the predictions of the GLE, the LE and the naive forecast are evaluated on both the training data (as in-sample predictions) and on the remaining $1-\alpha$% test data (as out-of-sample predictions). The coefficient of prediction ρ^2 from equation (2.12) is used to evaluate their predictive accuracy. It takes the value of $\rho^2 = 1$, if the prediction is always exactly true, and $\rho^2 = 0$, if the prediction is only as accurate as always using the mean $\bar{y}$, and $\rho^2 < 0$, if it is less accurate than using the mean. ρ^2 is computed via the function *sklearn.metrics.r2_score* from [75]. The results in Table 4.2 show that the GLE model consistently achieves the highest accuracy on in-sample and out-of-sample predictions for the three chosen test data sizes. Notably, the negative ρ^2 of the naive method for the test data indicates that the out-of-sample prediction is by no means trivial, meaning that the low, but positive

ρ^2 of the GLE on the test data nevertheless constitutes a good performance. The GLE achieves slightly better results than the LE, indicating that the memory effect should be taken into account for prediction tasks. Figure 4.13 shows the predictions of the LE and GLE for the in-sample and out-of-sample predictions with $\alpha = 90\%$. The same visual comparison between naive forecast and the GLE can be found in the Appendix E.3.

Table 4.2: Comparing the predictions of the naive forecast $\hat{y}_{t+1} = y_t$, the Markovian LE and the GLE with memory kernel length $k = 3$ via their ρ^2 score on in-sample training data and out-of-sample test data ($\alpha\%$ training data and the last $1 - \alpha\%$ of the time series as test data). Note that the partially negative ρ^2 for the test data indicates that the test data is quite difficult to predict, which corresponds to the high fluctuations in the final part of the time series in figure 4.2.

Model α	**In-Sample**			**Out-of-Sample**		
	80%	**85%**	**90%**	**80%**	**85%**	**90%**
Naive	0.07	0.15	0.19	−0.21	−0.15	−0.11
LE	0.29	0.32	0.35	−0.14	−0.01	0.08
GLE	0.39	0.42	0.45	0.01	0.08	0.10

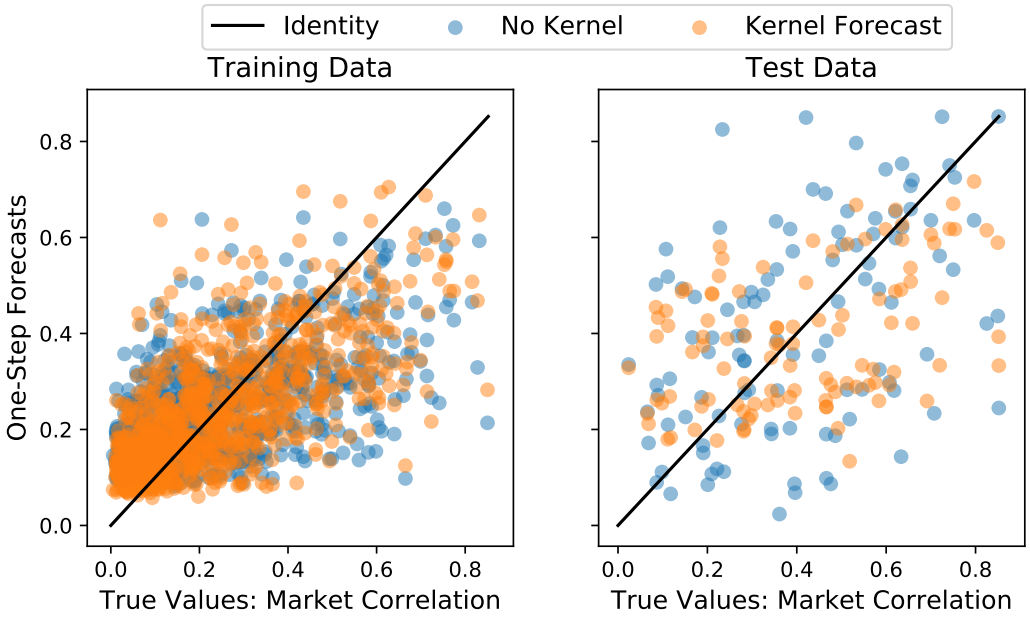

Figure 4.13: Comparison between one-step-ahead forecasts of the GLE model with Kernel against a regular Langevin estimation without memory effects on the training data (left) and the test data (right). The identity $f(x) = x$ is given as a benchmark for perfect predictive accuracy. Here, $\alpha = 90\%$ of the time series were used as training data.

Hidden Slow Time Scale and Non-Markovianity

As mentioned before, the results from [124] concerning the hidden process should not be considered as part of this thesis, but serve as an interesting additional point of comparison. Assuming that the observed process is stable with $\zeta < 0$, it is revealed that the hidden process in the non-Markovian model has to have an unobserved *slow* time scale, i.e. $\tau_\lambda > \gamma \cdot \tau_{\bar{C}}$ with $\gamma = 2$. The simple Markovian model with neither memory terms nor a hidden process cannot reproduce the stability of the observed process $\bar{C}(t)$.
Although one usually tends to regard the hidden time scale as a high-frequency noisy process, the evidence for a hidden slow time scale is supported by socio-economic intuition: Various processes with fast and slow time scales contribute to the economic development such as the fast high-frequency trading, the slow grow and fast burst of speculative bubbles, the gradual emergence of new technologies or cultural shifts which slowly change the investors' behaviour. Compared to the daily frequency of the financial data used for this research, many of these processes are much slower and could serve as candidates for such a hidden process.

4.3.4 Discussion of the Memory Effects

The estimated GLE model manages to reproduce the statistical properties of the original data for the end-of-week correlations, as shown in figures 4.10 and 4.11. The estimated memory kernel parameters show that even with a highly conservative interpretation of the 95% credible level, there are clearly nonzero memory effects for memory terms for all lags as far back as a lag of three weeks. Therefore, it is advised to use a model with memory to describe the correlation of the S&P500 market, which is an improvement to the Markovian Langevin model estimated in [117]. Moreover, the GLE estimation presented in this chapter achieves a high goodness-of-fit for the entire time series, whereas Stepanov et al. used a time-dependent Langevin model by splitting the time series into different intervals and estimating Markovian Langevin equations for each of them. Our work shows that this procedure can be circumvented by using a model with memory of three trading weeks. The major advantage of our method is its possible application in predicting future market correlation: The time-dependent drift estimation in [117] has no clear or smooth functional dependence on time and therefore, little information can be inferred about future values of the correlation time series. Our method needs no time dependency, generalises over the entire time series and can be used to predict future correlation values, which can be used for portfolio risk assessment. As shown in Section 4.3.3, the memory kernel helps the GLE to achieve better prediction accuracy than the regular Langevin equation and much better results than the naive forecasting method of using the last observation as the predicted value.

Notably, the existence of memory effects in the market's correlation structure can be interpreted in the context of volatility clustering. It is a well-known stylised fact from empirical research on financial markets that the volatility of a stock's returns tends to cluster: periods of high volatility are often followed by periods of high volatility and vice versa for low volatility [126, 168, 169]. The correlation $\rho_{X,Y}$ between two asset returns X and Y as defined in equation (2.9) directly includes the volatility values σ_X and σ_Y. Because the time series of volatility estimators $\sigma_X(t)$ shows a well-known memory effect, it is not far-fetched to assume a similar memory effect in the correlations $\rho_{X,Y}$ between two assets or, as we have discussed in this chapter, in the mean correlation of the entire market. Note that other memory effects have already been described and discussed in the literature, e.g. originating from large-scale traders splitting their strategy into many small orders over a longer period of time [25], via a cascading process along multiple time scales [126] or in the transition between market states [148].
These findings are supported by the results of the resilience analysis from [124] which is described in the PhD thesis of Martin Heßler in greater detail [159]. The resilience

analysis implies that a hidden and very slow time scale is coupled to the observed process $\bar{C}(t)$. In a metaphor from climatology, we can refer to these fast-scale processes as "economic weather" against the slower "economic climate" (cf. also the distinction in climate-like and weather-like tasks in [87] and the discussion of multiple time scales in [170, 171]). Haken's slaving principle might provide an explanation for the interaction between these time scales [21, 22]. According to this principle, the slowest dynamics enslave the faster dynamics which react so fast that they can adapt to the slowest dynamics in order to ensure the stability of the system. Note that the largest eigenvalues of the correlation matrix have already been interpreted in the context of different time scales with respect to Haken's slaving principle: In [126], their time scale is found to be slower than that of the sub-dominant eigenvalue processes. Interestingly, our study hints at the existence of a hidden process operating on an even slower time scale which is at least several years. More accurate estimation was unfortunately not possible due to limitations of the data, c.f. the discussion in [124, 159] for further details. On the other hand, the long-term changes of the economic climate might be driven by technological innovations, relate to the theories of economic cycles as explained by, e.g. Schumpeter or Kondratieff [172, 173] or need to be analysed in the broader context of cultural evolution and demographic changes proposed by the recently established research field of cliodynamics [48, 174].

4.4 Conclusion

This chapter analysed daily return time series aggregated to time windows that correspond to weekly data for the GLE and two months for the XAI algorithm. The GLE results expand on the results of [117] and imply a strong memory effect whose origin might be the hidden slow time scale from [124, 159]. As speculated by [159], this hidden process might be related to the field of cliodynamics which will be explored further in part IV of this thesis. The XAI algorithm identifies the energy and IT sector (possibly related to the dot-com bubble) as key factors in determining the state of the economy. The next chapter will use a different database for sector returns and analyse the Granger causality between them, thereby going beyond the correlation analysis of this chapter. Unlike the symmetric correlation $\rho_{X,Y} = \rho_{X,Y}$, this will include a direction of the interaction between the time series.

5 Causal Hierarchy in the Financial Market Network – Uncovered by the Helmholtz-Hodge-Kodaira Decomposition

"I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact."

— Winston Churchill, The Story of the Malakand Field Force

Abstract Granger causality can uncover the cause and effect relationships in financial networks. However, such networks can be convoluted and difficult to interpret, but the Helmholtz-Hodge-Kodaira decomposition can split them into a rotational and gradient component which reveals the hierarchy of Granger causality flow. Using Kenneth French's business sector return time series, it is revealed that during the Covid crisis, precious metals and pharmaceutical products are causal drivers of the financial network. Moreover, the estimated Granger causality network shows a high connectivity during crises which means that the research presented here can be especially useful to better understand crises in the market by revealing the dominant drivers of the crisis dynamics.

This chapter is based on *Tobias Wand, Oliver Kamps and Hiroshi Iyetomi, Entropy 2024, 26(10), 858* (2024) [175]. Tobias Wand and Hiroshi Iyetomi conceptualised the research project and wrote its software. Tobias Wand curated the data and carried out the research. Oliver Kamps and Hiroshi Iyetomi supervised the project. All figures in this chapter are taken from [175] and were created by Tobias Wand.

5.1 Introduction

One of the most important messages in many introductory lectures to statistics is that correlation does not imply causation [176]. However, it begs the question: What, then, is causality? And how can it be quantified? One of the first and most widespread attempts to formalise causality was proposed by Granger [90]. Granger causality takes the time ordering into account as the cause needs to happen before the effect. The field of causal inference has developed several tools to probe time series data for causal interactions [89] and has been used to analyse dynamical systems [92, 177]. Especially for high dimensional multivariate time series, it is difficult to infer the network of causality because one has to carefully distinguish between the different possible causal drivers [178, 179, 180, 181].

While such networks can be difficult to interpret, especially if they contain cyclic substructures [89, 182], the Helmholtz-Hodge-Kodaira Decomposition (HHKD) can disentangle them. As a reformulation of Helmholtz-Hodge field theory for discrete graphs, the HHKD can split a directed network into a cyclic and gradient-based graph [43, 44]. The latter will then provide a ranking of all nodes according to whether they are upstream or downstream in the hierarchy of causal drivers. The HHKD has been used in econophyscis to understand the dynamics of cryptocurrencies [183] as well as the network of shared ownership of companies [184]. Capturing economic and financial interactions in a network has been a standard approach of econophysics and complexity science in the past [37, 38, 39] and causal inference has been applied to such networks of companies or countries [185, 186, 187].

This chapter analyses time series data of business sectors from [188] to investigate the Granger causality between different sectors of the economy. Using the HHKD on this Granger network then reveals if a sector is rather driven by other sectors or if it is a causal driver of the whole system. This chapter will first explain the database, the algorithm from [180] to estimate Granger causality networks and the HHKD for graphs in section 5.2. The results of the HHKD will then be presented in section 5.3 for different time periods before interpreting them and discussing further extensions to this research in section 5.4.

5.2 Materials and Methods

5.2.1 Data

To analyse the interactions between different sectors of the economy, we use the database of Ken French which contains the return time series of representative portfolios for 49 different business sectors [188]. These portfolios are constructed as value-weighted averages of all stocks in a business sector listed on NYSE, AMEX and NASDAQ and the data consists of the daily returns $R_t^{(i)} = \left(P_t^{(i)} - P_{t-1}^{(i)}\right)/P_{t-1}^{(i)}$ of these portfolios' prices. The calculation of returns normalises the time series and removes trends in the data. Augmented Dickey-Fuller tests [189] performed via the Python package *statsmodels* [167] moreover indicate that the given time series are stationary for each window period of interest in section 5.3 with the exception of one sector each in 2006 and 2008 (though that does not impact our results). This database is updated continuously and further details on the data curation are given in [190]. Although the assignment of companies to a sector was done manually, a comparative study with modern statistical tools shows the high agreement between French's classification and data-driven methods [191]. While the data was originally used for capital asset pricing modelling in [192], it has found numerous applications in various fields of economic and financial research as a data resource (see [191] for an overview).

5.2.2 Granger Causality

The intuition behind Granger causality is that the cause X should happen before the effect Y and that knowing the cause should improve the future prediction of the effect. The latter can be measured by fitting autoregressive linear models with and without X and comparing their accuracy. By including possible alternative causes Z for Y, this concept is extended to the conditional Granger causality *CGC*. First, a full model is estimated that measures how well the past of X, Y and background variables Z predict the future of Y_{t+1} via

$$Y_{t+1} = \sum_{\tau=0}^{\tau_{\max}} \left(\alpha_\tau Y_{t-\tau} + \beta_\tau X_{t-\tau} + \gamma_\tau Z_{t-\tau}\right) + \epsilon \tag{5.1}$$

with i.i.d. Gaussian errors $\epsilon \sim \mathcal{N}(0, \sigma_F^2)$ and a maximum time lag $\tau_{\max}$ to limit how much of the past should be considered for predicting Y_{t+1}. Note that Z might contain more than one background variable $Z = \left(Z^{(1)}, \ldots, Z^{(s)}\right)$ with $\gamma_\tau \in \mathbb{R}^s$. Then, a reduced

model is estimated without the proposed cause X as

$$Y_{t+1} = \sum_{\tau=0}^{\tau_{\max}} \left(\alpha_\tau' Y_{t-\tau} + \gamma_\tau' Z_{t-\tau}\right) + \epsilon \tag{5.2}$$

with $\epsilon \sim \mathcal{N}(0, \sigma_R^2)$. The conditional Granger causality of X on Y is then given by how much the reduced model's variance increases compared to the full model and is defined as

$$CGC_{X\to Y} = \log \frac{\sigma_R^2}{\sigma_F^2} \tag{5.3}$$

to measure how much X causes Y.

For multivariate data with many time series, the estimation of the full model in (5.1) can easily fall into the regime of overfitting [193]. Hence, it is of utmost importance to carefully construct the full model. A comparative study of multivariate Granger networks [181] indicates that the restricted conditional Granger causality index (RCGCI) from [180] is the most suitable Granger causality estimation scheme for the financial data analysed in this chapter. At the heart of RCGCI lies the construction of the full model by starting with an empty regression model and sequentially adding variables $X_{t-k\tau}^{(i)}$ with a lag of k time units to it, if they reduce the Bayesian information criterion of the regression given by equation (2.14). Hence, the resulting full model may not contain lagged representations of all possible explanatory variables but only some selected $\left(X^{(i_1)}, \ldots, X^{(i_I)}\right)$ and therefore guarantees sparsity to prevent overfitting in the estimation process. For the selected variables, CGC can be computed by removing them from the full model and fitting the reduced model, whereas the remaining selected variables are conditioned on as the background information Z. For the other X_j which have not been included in the full model, the CGC is set to zero as no causal relationship had been estimated.

Details on the Estimation

Because the financial returns analysed in this chapter are known to have an almost nonexistent autocorrelation [18] and to avoid overfitting, we restrict our models to maximum lags of one time step which is supported by an exploratory analysis. Previous studies have shown that principle component analysis (PCA) can be used to distinguish between noise and collective effects in financial time series [34, 35]. Hence, we perform PCA on the raw data, only keep the principle components with the largest eigenvalues so that their sum describes 90% of the total variation in the data and discard the remaining principle components as noise before performing the inverse transformation back into the

original feature space. Note that the sparsity of the RCGCI algorithm additionally limits the influence of noise on the results. Averaging over all sectors and all time periods under consideration, the typical ratio between the variance explained by the full regression model and the variance of the data is $\sigma_F^2/\sigma_{\text{Data}}^2 \approx 96\%$ indicating a good fit of the full regression model and a high signal to noise ratio.

5.2.3 Helmholtz-Hodge-Kodaira Decomposition

The reconstructed network of the causality flux between multivariate time series might not be ad hoc trivial to interpret. Circular causalities (A causes B, B causes C and C causes A) can be present and inspecting the network with the naked eye may not be sufficient to understand its structure. The Helmholtz-Hodge-Kodaira Decomposition (HHKD) is a tool to analyse the flux in networks and to disentangle the flow into upstream and downstream directions [43, 44].

Mathematical Formulation of the Unidirectional HHKD

The Helmholtz decomposition theorem states that any well-behaved vector field $\mathbf{F}(\mathbf{r}) \in \mathbb{R}^n$ can be decomposed into two components $\mathbf{F}(\mathbf{r}) = \mathbf{G}(\mathbf{r}) + \mathbf{R}(\mathbf{r})$, a gradient field $\mathbf{G}(\mathbf{r})$ and a divergence-free field $\mathbf{R}(\mathbf{r})$. The rotation-free field $\mathbf{G}(\mathbf{r})$ can be expressed as the gradient of a potential $\mathbf{G}(\mathbf{r}) = -\nabla_{\mathbf{r}}\Phi(\mathbf{r})$ such that the potential determines the direction of a flux in the space of $\mathbf{r}$. The divergence-free or solenoidal field $\mathbf{R}(\mathbf{r})$ has the property that no point $\mathbf{r}$ is a source or sink of the observed flux as $\forall \mathbf{r} : \nabla_{\mathbf{r}} \cdot \mathbf{R}(\mathbf{r}) = 0$. Note that a third component can exist and represents a background flux into and out of the system, but is usually ignored as one assumes that the system of interest is sufficiently closed.

The same reasoning can be applied to a flow network on a discrete graph [43, 44]. Let J_{ij} be the observed flow from node i to j with the antisymmetric property $J_{ij} = -J_{ji}$. It can be shown that a unique decomposition $J_{ij} = J_{ij}^{(g)} + J_{ij}^{(c)}$ exists such that $J_{ij}^{(g)}$ is the gradient and $J_{ij}^{(c)}$ the circular flow from i to j. In this decomposition, the gradient flux fulfils $J_{ij}^{(g)} = G_{ij}\,(\Phi_i - \Phi_j)$ for some background potential Φ assigned to each node and with the standard choice for weights between two nodes being $G_{ij} = 1$. The circular flow fulfils $\forall i : \sum_j J_{ij}^{(c)} \stackrel{!}{=} 0$, i.e. that for each node the total influx is equal to its total outflow. For a simple network with three nodes, this decomposition is illustrated in figure 5.1. The potential and its associated gradient flow can be obtained from the least square estimation

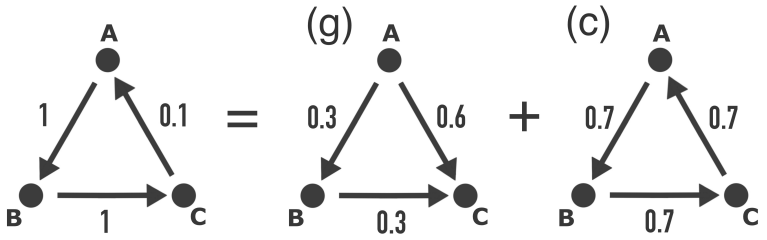

Figure 5.1: Example of the Helmholtz-Hodge-Kodaira decomposition for a single graph into a gradient-based graph (g) and a circular graph (c). Note that direction of the flux between A and C is different in (g) and (c) which is the same as changing the sign $J^{(g)}_{AC} = -J^{(g)}_{CA}$ and hence their sum is given by $J^{(g)}_{CA} + J^{(c)}_{CA} = -0.6 + 0.7 = 0.1$ and reconstructs the original flux J_{CA}. Also note that the total flux between two nodes is path independent for (g) as $J^{(g)}_{AC} = J^{(g)}_{AB} + J^{(g)}_{BC}$.

$$\min_{\mathbf{J}^{(g)}} (I) \quad \text{with} \quad I = \frac{1}{2} \sum_{i<j} \frac{1}{G_{ij}} \left(J_{ij} - J^{(g)}_{ij} \right)^2 = \frac{1}{2} \sum_{i<j} \frac{1}{G_{ij}} \left(J_{ij} - G_{ij} \left(\Phi_i - \Phi_j \right) \right)^2 \quad (5.4)$$

and the circular flow is then simply the difference $J^{(c)}_{ij} = J_{ij} - J^{(g)}_{ij}$. For the standard choice $G_{ij} = 1$, this formulation also has the useful property that the net gradient flux is the same along all paths between any two nodes. Because the gradient flow only depends on the potential difference $\Phi_i - \Phi_j$, the same gradient flow can also be obtained if the potentials have a constant offset $\Phi_i \to \Phi_i + \Phi_{\mathcal{O}}$. Hence, the minimisation of equation (5.4) needs an additional constraint to produce unique results for Φ such as $\Phi_n = 0$ or $\sum_i \Phi_i = 0$. We note that the mathematical formulation of the HHKD tends to rather minimize large errors than small ones and that it treats large flows between two nodes as more informative than small ones. This means that any necessary errors during the minimization procedure are distributed across many edges. Because of the sparse RCGCI algorithm, the weakest causal links are discarded before the application of the HHKD and hence, the errors will be relatively weak compared to the strength of the estimated network links. Yet, this shortcoming should be kept in mind. We additionally note that, as proven for a trait-performance model in [194], the uncertainty in the estimation of $CGC_{X \to Y}$ introduces a bias towards more cyclical structures.

Bidirectional Flows

Whether it be through noise or feedback loops, it is in general possible that the RCGCI algorithm estimates that $CGC_{X\to Y} > 0$ and $CGC_{Y\to X} > 0$, i.e. that two time series are estimated to Granger cause each other and that the flux cannot be defined as antisymmetric $J^{(b)}_{ij} \neq -J^{(b)}_{ji}$ (where the superscript b denotes the bidirecionality) and $J^{(b)}_{ij} \geq 0$. Naively, computing the net flow $CGC_{X\to Y} - CGC_{Y\to X}$ seems like a reasonable choice, but this discards the information about the relative strength of the net flux. Consider a system in which $CGC_{A\to B} = 0.6$, $CGC_{B\to A} = 0.1$, $CGC_{C\to D} = 5.5$, $CGC_{D\to C} = 5$. The net flux $A \to B$ and $C \to D$ is 0.5, but scaling this with the total flux between the node pairs shows that this difference is much less significant for the flux $C \to D$.

A bidirectional version of the HHKD that reflects these considerations is presented in [184]. The authors argue to interpret G_{ij} as an analogy to the conductance in electrical circuits and that a high flux between both nodes should correspond to a high conductance whereas a low total flux indicates a high resistance. Hence, they propose to split the original bidirectional network into two graphs which are then used to perform the HHKD: The difference of the flux in both directions between two nodes in the bidirectional network $J^{(b)}$ is defined as the net flux $J_{ij} \coloneqq J^{(b)}_{ij} - J^{(b)}_{ji}$ and forms a unidirectional network with the antisymmetry $J_{ij} = -J_{ji}$ to which the HHKD can be applied. The sum of the absolute values of the flux in both directions is used as the conductivity $G_{ij} \coloneqq J^{(b)}_{ij} + J^{(b)}_{ji}$ which is the same in both directions $G_{ij} = G_{ji}$. The minimisation in equation (5.4) can then be applied to the two networks (J, G) to receive the HHKD ranking of the nodes.
This generalisation also gives rise to a helpful interpretation of the potential differences of the nodes: Consider only the flux between two nodes i and j isolated from the rest of the graph. Let J_{ij} be the net flow from i to j and G_{ij} the total flow. The contribution of this connection to the minimised functional I is then given by

$$I_{ij} = \frac{1}{2}\frac{1}{G_{ij}}\left(J_{ij} - J^{(g)}_{ij}\right)^2 = \frac{1}{2}\frac{1}{G_{ij}}\left(J_{ij} - G_{ij}\left(\Phi_i - \Phi_j\right)\right)^2 = \frac{1}{2}\frac{1}{G_{ij}}\left(J_{ij} - G_{ij}\Delta_{ij}\right)^2 \quad (5.5)$$

where Δ_{ij} expresses the potential difference between the two nodes. Minimisation of I_{ij} with respect to Δ_{ij} leads to

$$\frac{\partial I_{ij}}{\partial \Delta_{ij}} = -J_{ij} + G_{ij}\Delta_{ij} \overset{!}{=} 0 \Leftrightarrow \Delta_{ij} = \frac{J_{ij}}{G_{ij}} \equiv \frac{\text{Net Flow}}{\text{Total Flow}}. \quad (5.6)$$

Hence, this rule-of-thumb approximation which disregards all other edges shows that the potential difference can be interpreted as the ratio between net and total flow between the two nodes. In particular, if the flux in one direction is much larger than in the other direction $J_{ij}^{(b)} \gg J_{ji}^{(b)}$, then net and total flow are almost identical $J_{ij} \approx G_{ij}$ so that $\Delta_{ij} \approx 1$. Therefore, as described in [184], one unit of potential difference can be interpreted as a separation of approximately one layer between the nodes i and j. Our implementation of this algorithm is available at [195].

Circularity and Hierarchy

Once the flow network is decomposed into gradient-based and circular flux, one can compare their respective contributions to the net flux [184, 196, 197]. It is possible to quantify the contribution of the gradient-based flux via the L^2 norm as

$$\Gamma = \frac{1}{2} \sum_i \Gamma_i = \frac{1}{2} \sum_i \sum_j G_{ij} \left(J_{ij}^{(g)} \right)^2 \tag{5.7}$$

and that of the circular flux as

$$\Lambda = \frac{1}{2} \sum_i \Lambda_i = \frac{1}{2} \sum_i \sum_j G_{ij} \left(J_{ij}^{(c)} \right)^2 \tag{5.8}$$

where Γ_i and Λ_i denote the contribution of the respective i^{th} node. Normalising them with the total flux

$$N = \frac{1}{2} \sum_i \sum_j G_{ij} \left(J_{ij} \right)^2 \tag{5.9}$$

leads to the definition of

$$\gamma = \frac{\Gamma}{N} \text{ and } \lambda = \frac{\Lambda}{N} \tag{5.10}$$

which fulfil $\gamma + \lambda = 1$. A completely hierarchical network will have $\gamma = 1$, a completely circular network $\lambda = 1$. A high $\gamma \gg \lambda$ indicates that the underlying potential and its corresponding hierarchy have been cleansed from noise and insignificant loops and now accurately reflects the true structure of the underlying dynamics.

5.2.4 Test on Synthetic Data

To test the pipeline of RCGCI and HHKD, we simulate 50 realisation of a network of 49 time series with the network structure given in figure 5.2. Because the RCGCI-HHKD pipeline is supposed to uncover the hierarchy of time series, the synthetic network in 5.2 was chosen to represent a hierarchical structure and we evaluate how accurately the

algorithm identifies the nodes at the top of the hierarchy: One node X_0 is an independent stochastic process and at the top of the hierarchy. It drives the nodes in the next layer $X_{1,\dots,8}$ as their causal parent pa and each of them drives five nodes from $X_{9,\dots,48}$ in the final layer at the bottom of the hierarchy. Each node $X^{(i)}$ of the network is simulated for 250 time steps in a vector autoregression according to

$$X_{t+1}^{(i)} = -0.5X_t^{(i)} - 0.5X_t^{pa(i)} + \epsilon \tag{5.11}$$

where $pa(i)$ denotes the parent node of i (if it exists) and $\epsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma)$ with $\sigma = 1$. The goal of the RCGCI-HHKD pipeline will be to accurately identify the nodes with index $\mathcal{I} = (0, \dots, 8)$ at the top of the causal hierarchy. Note that 49 time series for 250 time steps correspond to one year of trading day data in the database described in section 5.2.1 and that the standard deviation $\sigma = 1$ is roughly equal to the standard deviation of the data. Hence, these synthetic time series provide a realistic artificial version of the observed data. The HHKD is used on the Granger causality network estimated by RCGCI and it is evaluated how many of the top nine nodes in the estimated hierarchy actually belong to the set $\mathcal{I}$. For the 50 realisations of the network, the average of the ensemble of detection rates is 94% and the median is even 100%. Though not shown here, for synthetic cyclic networks, the RCGCI also successfully estimates the loop topology. Hence, the combination of RCGCI and HHKD accurately estimates the network structure. Even after adding observation noise $\mathcal{N}(0, \sigma_{\text{obs}})$ to the simulated data via

$$X_t^{(i)} \to \tilde{X}_t^{(i)} = X_t^{(i)} + \zeta \text{ with } \zeta \sim \mathcal{N}(0, \sigma_{\text{obs}}), \tag{5.12}$$

the top nodes of the networks hierarchy are still estimated with a high accuracy beyond the random expectation for noise up to $\sigma_{\text{obs}} \approx 1$.

Additionally, we also create ensembles of uncoupled time series by taking a random subset of trading days $t_{i_1}, \dots, t_{i_{250}}$ without any chronological ordering. This represents the null hypothesis that no causal coupling is present in the data. We then check if the RCGCI erroneously reconstructs a network even though none is present. We define the network connectivity as the percentage of nodes that are estimated to have a causal coupling to another node. These 50 percentage values give us a confidence interval (CI) of the network connectivity under the null hypothesis that no causal coupling is present in the data. If the network connectivity of a system exceeds the CI, we can therefore conclude that we have identified a system with meaningful causal connections.

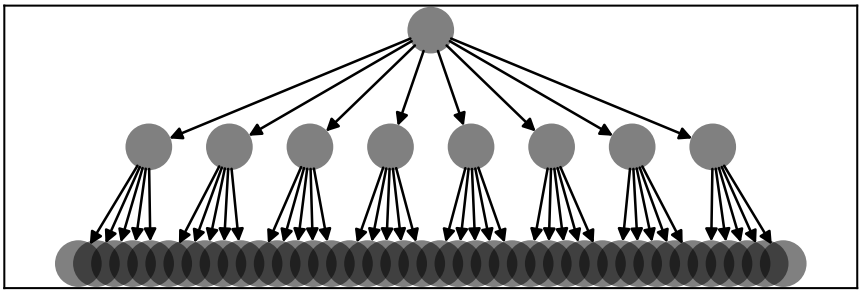

Figure 5.2: The network structure used for the vector autoregression which generates synthetic time series. One node is at the top of the hierarchy without any causal parent whereas 8 nodes are in the second layer and 40 are in the final layer. Each node in the second layer is the parent node of 5 nodes in the final layer and have the node in the first layer as their parent node. Sketched via the software [198].

5.3 Results

5.3.1 Year by Year

To gain the most insight from the HHKD, it is necessary to have a connected network in which all sectors are included. By defining the network connectivity as the percentage of nodes which are coupled to the network, the RCGCI-HHKD analysis is performed on the annual data during the last 20 years from 2004 to 2023 to identify periods with a connectivity of 100%.
The results in figure 5.3 show that the market mostly remains within the range expected from random time series, but some periods exhibit a spike to a significantly high level of network connectivity. It only reaches a connectivity of 100% (i.e. that all sectors are coupled to the network) during the year 2020 and reaches a connectivity of almost 100% (only one sector is decoupled) in 2007. Hence, the 2020 period will be the focus of the latter half of this section.

To gain more insights into general structures of the RCGCI networks across the years, the sum of Granger causality influx and outflux is recorded for each of the 49 sectors as well as during how many years they are connected to the network by influx or outflux links. After rescaling all of these quantities to the same scale in figure 5.4, kernel density estimation (KDE, cf. section 2.2.2) shows that for inwards and outwards directions, the total flux and the linkage rate are distributed similarly. The influx is a Gaussian bell

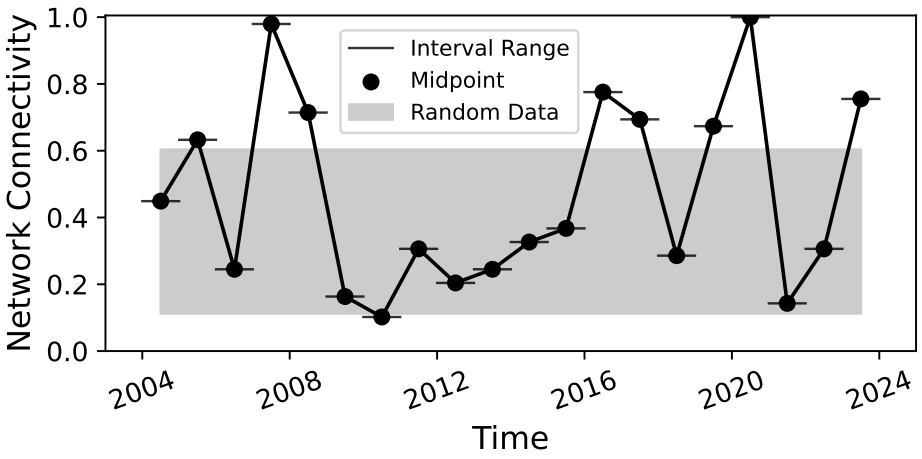

Figure 5.3: Results of the RCGCI-HHKD analysis for annual data from [188]. The grey shaded area is the CI for the network connectivity of random data without any causal coupling. Note that the lines that connect the dots are only a visual aid and no linear interpolation between the periods is assumed.

curve (a two-sided χ^2 test has a p value of 0.21 and cannot reject the null hypothesis of Gaussianity) whereas the outflux has a higher variance and, notably, a fat tail at high values. Hence, while most sectors have a similar influx of Granger causality, some sectors drive the other sectors with a much stronger outwards Granger causality than most others. It is therefore more interesting to focus on the sectors with particularly high or low outflux of Granger causality. The sectors Rubbr (Rubber and Plastic Products), BldMt (Construction Materials), Mach (Machinery), Trans (Transportation) and, perhaps surprisingly, Banks show no outflux of Granger causality during any of the periods. Gold (Precious Metals) and, on the second position, Cnstr (Construction) have a much higher sum of Granger causality outflux and a much higher rate of outwards linkage than all other sectors. An overview of the sector abbrevitions is provided at the end of this chapter in table 5.1.

5.3.2 Connected Graphs during the 2020 Covid Pandemic

To further investigate the connected network for the year 2020, the RCGCI-HHKD pipeline is used to analyse time windows of 12 months which are shifted by one month and scan over the year 2020. This process starts with the time interval January 2019 to January 2020 and ends with the period from December 2020 to December 2021. Note that the figures displaying these results use the midpoint of each time period on the x-axis,

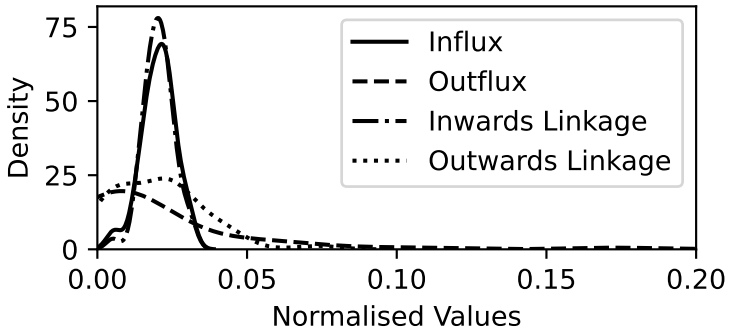

Figure 5.4: For the analysis of annual data from 2004 to 2023, KDE of the sum of all influx and outflux of Granger causality and of the total number of years with at least one inwards or outwards link in the RCGCI network. Values on the x-axis have been normalised to the same scale.

e.g. July 2019 for the period from January 2019 up to January 2020. For each period with a connected network, the parameter γ is calculated according to equation (5.10) to quantify the contribution of the gradient flow to the observed flux. Because $\lambda = 1 - \gamma$, the calculation of λ is omitted.

Figure 5.5 shows the results for the connectivity and the gradient contribution γ. Whether the network is connected or not strongly depends on whether March 2020 is included in the data. During this period, the gradient contribution is typically around $\gamma \approx 0.8$ and therefore stronger than the rotational flow λ. Though due to the quadratic L^2 norm used to calculate (5.10), a rotational component $\lambda \approx 0.2$ is nevertheless a non-negligible contribution to the total flow. Looking at the ensemble of Granger causality flux matrices for the 12 time windows with connectivity 1 reveals that every single sector has always some causality influx (i.e. it is estimated to be Granger caused by another sector) for all 12 time windows with the exception of the sector Gold which only has an influx of Granger causality for 3 of the 12 periods. The sectors PerSv (Personal Services), Other, Aero (Aircraft) and Trans (Transportation) have no outflux of Granger causality during any of the 12 time windows and for the latter two sectors, this might reflect the travel restrictions imposed during this period. In contrast to the sum of Granger causality outflux for the disjoint long-term analysis in the previous section, Gold is no longer the sector with the highest total outflux (≈ 5.00) but is far overtaken by Drugs (Pharmaceutical Products; ≈ 35.8), Hshld (Consumer Goods; ≈ 28.4) and Cnstr (≈ 10.6) with some other sectors at a slightly higher level than Gold, too.

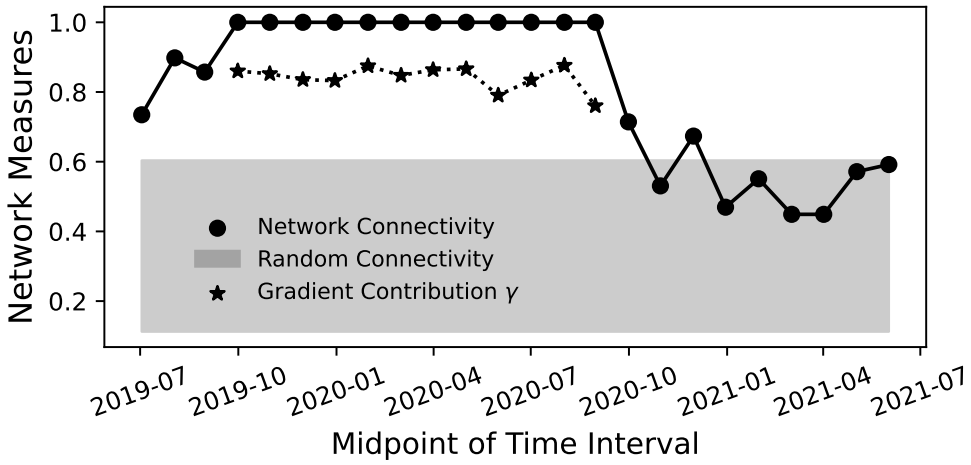

Figure 5.5: For time periods of 12 months, two network measures are depicted here: the network connectivity and the gradient contribution γ. The network connectivity is the percentage of sectors connected to the network and is displayed here against the random connectivity expected for independent time series. If the network has a connectivity of 1, the gradient contribution γ is also calculated according to equation (5.10). Note that the time on the x-axis is the midpoint of the 12 months intervals of data.

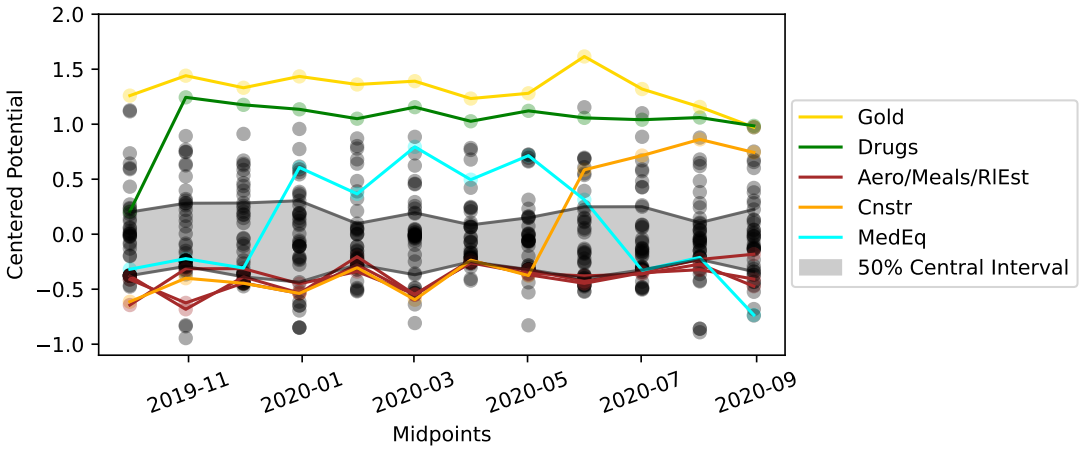

Figure 5.6: For the same time intervals as in figure 5.5,the potentials Φ_i of each sector are shown as dots. Note that for each time interval, the potentials have been centred via $\sum_i \Phi_i = 0$. Some selected sectors are shown in colour and the grey area shows the spread between the 25% and 75% quantiles for each time period.

For the periods with connected graphs, the potential Φ_i can be calculated for every single node. High Φ_i indicate a high position in the hierarchy of Granger causality and that the sector is a cause rather than an effect. Because of the large contribution γ of the gradient flux shown in figure 5.5, this hierarchy is not obstructed by strong circular fluxes in the system and indeed reflects the underlying dynamics. Figure 5.6 shows the potentials for all 12 time windows with a connected graph. These potentials can be compared to each other because they have been normalised to fulfil $\sum_i \Phi_i = 0$ for each time period. The range between minimum and maximum potential values has a mean of 2.0 with standard deviation ± 0.2 across these periods and therefore reflects a network of approximately 3 different levels with a fairly stable potential range. Additionally, for the period from October 2019 to September 2020, the full network is depicted in figure 5.7 (bottom) where the nodes' vertical positions reflect their potential values. Some selected sectors have been highlighted in these plots: The sectors Gold and, with the exception of the first interval, Drugs are consistently at the top of the hierarchy and their potentials have a low variance. Similarly, the potentials of the sectors Aero, Meals (Restaurants, Hotels, Motels) and RlEst (Real Estate) are consistently found at the bottom of the potential hierarchy.

Some other sectors have a high variance and change their position in the hierarchy rather drastically: The potential of Cnstr has the highest variance and this sector moves upwards in the hierarchy during the latter third of the periods. This might reflect the increase of construction material prices and their effect on the construction businesses and, as a cascading effect, on other business sectors during the beginning of 2021 [199]. The second highest variance is observed for the potential of MedEq (Medical Equipment). This sector rises in the hierarchy during the peak of Covid, which probably reflects the increasing demand for products such as face masks and testing equipment. The sudden decline of MedEq in the hierarchy starts in the time windows which already include the first weeks of widespread vaccinations in western countries which was interpreted as a sign of the end of the pandemic and hence of a lower importance of such equipment.

Because of the small but notable contribution of the rotational flux to the system during the Covid crisis, we also investigate which nodes have a strong rotational component Λ_i as in equation (5.8). For each time period with a connected network graph, the values Γ_i and Λ_i are calculated according to equations (5.7) and (5.8) and the rotational component is normalised in two ways: $\Lambda_i^{(N)} = \Lambda_i / N$ denotes how much the rotational flux of node i contributes to the total flux N from equation (5.9). $\lambda_i = \frac{\Lambda_i}{\Lambda_i + \Gamma_i}$ denotes whether the flux of node i is rather dominated by rotational flows or by the gradient flow. For each sector i, the mean values of $\Lambda_i^{(N)}$ and λ_i are calculated over the time periods of consideration.

While the sectors Drugs, Hshld and Cnstr have the highest, second highest and fourth highest mean value of $\Lambda_i^{(N)}$ and contribute much to the rotational flow, their own flow does not show a particularly high contribution λ_i of the rotational component. Rather their $\Lambda_i^{(N)}$ is high because they in general have strong causality links to other sectors. The third and fifth highest values of $\Lambda_i^{(N)}$ are found for the sectors Toys (Recreation) and Softw (Computer Software) and they also have the second and third highest values of λ_i with $\lambda_i \approx 0.47$ for both sectors. These sectors not only provide a strong rotational contribution to the total observed causality flow ($\Lambda_i^{(N)}$) but also experience an almost equally strong effect from the gradient and rotational flows (λ_i) on their own dynamics. This makes them interesting candidates for future research to better understand the circular dependencies in the financial network.

5.3.3 The 2007 Financial Crisis

Because the Granger network for 2007 does not have a connectivity of 100%, it will not be analysed as deeply as the 2020 network in this study. But since only a single sector (FabPr, Fabricated Products) is disconnected from the rest, it might be justified to briefly focus on the reduced network of the 48 connected sectors; not least because this period also coincides with the onset of the financial crisis in the late 2000s [200] and serves as an interesting comparison to the Covid crisis. This reduced network is depicted in figure 5.7 (top) and visual inspection shows a much more streamlined flow than for the network during 2020 and hence a more hierarchically organised potential: Whereas the 2020 network is much more entangled, the 2007 network mostly consists of links from the sector Gold to other sectors with much fewer links between the other sectors, resulting in a shape reminiscent of the depictions of Aton in ancient Egyptian artworks. This is confirmed by the estimation of the gradient flow contribution for the reduced network which yields $\gamma^{(2007)} = 0.98$ and shows an even higher gradient contribution than for any of the 2020 networks. Although the crisis starting in 2007 is generally known as the financial crisis, the financial sectors do not have a particularly important position in the hierarchy of Granger causality during this period and perhaps rather act as mediators of causality than as causal drivers. This surprising result might indicate that the causality analysis for this period does not fully represent the processes in the real economy but uncovers more subtle relationships between the time series.

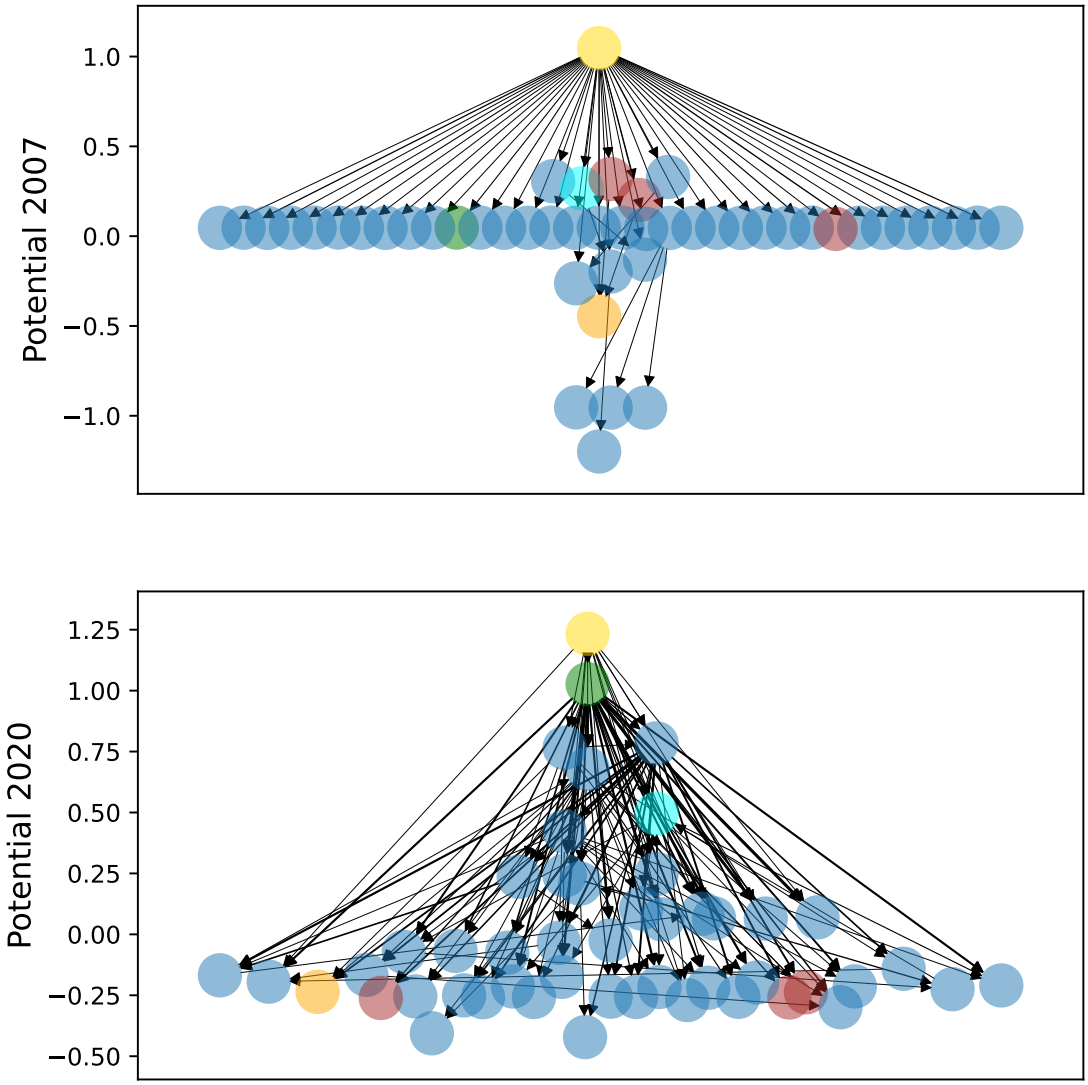

Figure 5.7: The estimated Granger causality influence network ordered by the HHKD potentials for the periods from January 2007 to December 2007 (the sector FabPr is not shown because it has no link to any other sector) and from October 2019 to September 2020. The width of the arrows reflects strength of the Granger causality and selected sectors are highlighted with the same colour coding as in fig 5.6.

5.4 Discussion

Analysis of the disjoint annual periods in figure 5.3 indicates that a highly connected Granger causality network coincides with market crises. The highest connectivity values close to 100% were observed during the 2020 Covid crisis and the beginning of the financial crisis in 2007. Other spikes occur during the year 2023, possibly reflecting the collapse of several mid-size banks and the threat of a contagion in the US banking crisis [201], and in 2016 following the market turmoil after the unexpected election of Donald Trump [202]. These results align with the causality estimation in [186] as well as with research on financial correlation matrices which also shows an increase in coupling between financial time series during times of crises [36, 116, 117]. The lower coupling during periods of a healthy market reflect that the time series are more independent and diversified which reduces the overall risk in the market.

The precious metal sector is usually found upstream at the top of the Granger causality hierarchy. This is to some extent in line with the clustering analysis in [148] which identified the precious metal mining companies as having different dynamics compared to the other assets. Deeper insights into the estimated $CGC_{X\to Y}$ values between the sectors show that the precious metal sector is, however, not typically the strongest driver of the market dynamics. Its position at the top of the hierarchy rather reflects that this sector is rarely driven by other sectors. Hence, this result should be interpreted with caution. This might be a specific feature of the precious metal sector because the pharmaceutical sector is also frequently found at the top of the hierarchy, but has a strong outflux of Granger causality and therefore acts as a driving force of the system's dynamics. Note that the returns of this sector are calculated based on the companies which trade precious metals and do not directly contain the prices of gold and other metals. Adding this might be an interesting endeavour for future work. During the Covid crisis, the high position of Drugs in the causality hierarchy as well as the rise of MedEq during the pandemic's peak reflect our intuition about the economy during the year 2020. Perhaps surprisingly, the financial sectors do not have a high position in the hierarchy of Granger causality and especially the banking sector is found rather far downstream. This might be interpreted as the financial companies being only a mediator of causal influence and providing the infrastructure to the flow of causality in the financial markets, but not actually driving the flow themselves. Our results therefore differ from the study in [203] where an analysis of the input-output network of business sectors shows that the energy and finance sectors have a high upstream position in the hierarchy. This is an important indicator that the financial market network analysed in our study does not simply resemble the

real economy, but has its own dynamical behaviour. The high signal to noise ratio of the full regression models in the RCGCI algorithm and the strong contribution $\gamma \gg \lambda$ of the gradient flow indicates that the hierarchy estimated by the HHKD is reflective of the true underlying structure of the market dynamics. Interestingly, γ was notably higher for the 2007 financial crisis than for the Covid crisis, possibly because the former was an endogenous crisis and the latter an external shock. Even though the inherent uncertainty of the estimated $CGC_{X\to Y}$ values should bias the HHKD towards a more cyclical result [194], this bias seems to have been compensated by the sparsity of the RCGCI algorithm. We hence conclude that the resulting cyclical component estimated by the RCGCI-HHKD pipeline is indeed a structural component and not merely noise.

Numerous extensions can be made in future work to this project. Adding return time series of the precious metals' prices has already been suggested, but beyond this, macroeconomic variables like the inflation rate might be used as background variables Z in equation (5.1). Without attempting to make regression models to predict Z, these variables can still be used to calculate the Granger causality conditional on the macroeconomic information provided by them. This might be interpreted as the third translational component which is usually omitted in Helmholtz-Hodge considerations but represents an influx or outflux into the whole system of interest (i.e. that it is an open sytem) as discussed in [44]. Moreover, the linear regression can be extended with interaction terms between two variables $X^{(i)}_{t-\tau} \cdot X^{(j)}_{t-\tau}$ or nonlinear functions [204] to alleviate the shortcomings of Granger causality methods [205, 206], but this might require larger amounts of data for reliable estimation and thus higher frequency data than the one available in [188]. Other methods from causal inference, such as the lead-lag relationship of complex Hilbert PCA [207, 208] or transfer entropy [179, 185] can capture nonlinear nonlinear effects, but might require more data, too. Although the heightened position of precious metals and pharmaceutical products can be related to real effects in the data, the restriction of Granger causality as a linear measure may have been the reason why the financial companies do not have a high position in the estimated hierarchy of causation. As this is rather counter-intuitive, nonlinear causality measures might be used to provide a different perspective on the data and to check the robustness of these findings. Also, one could extend the RCGCI algorithm to include a bootstrapping procedure in the estimation of (5.3) to get an uncertainty estimation of the Granger causality $CGC_{X\to Y}$. While the RCGCI and the standard formulation of Granger causality do not distinguish between positive and negative influences between variables similar to [187], a multi-layer network approach could be used to separate the causal couplings based on their sign. However, extending the HHKD to multi-layer networks is required to evaluate this, perhaps based

on the approach in [209].
Because of the high network connectivity during crises, the RCGCI-HHKD pipeline is especially useful to describe the system dynamics during such periods. In particular, understanding the flow of causality and identifying the causal drivers during a crisis might allow policymakers to more effectively intervene to stop the crisis by focusing on the sectors which are upstream in the causality hierarchy. This could open up the possibility to stabilise the market with a minimally invasive intervention.
Though beyond the scope of this work, we believe that the HHKD can help to overcome the limitations of the causality framework described by Judea Pearl [89]. Pearls approach to causality relies on directed acyclic graphs (DAG) between the variables and therefore requires an interaction network without any closed loops. While this is not always present in real systems, an adaptation of the HHKD might provide a suitable tool to extract such DAGs from real-world systems as the gradient component of the original graph.
Comparing the results from this chapter to the correlation analysis in chapter 4 we see that although both methods agree on finding high levels of interactions during crises, their sector rankings show important differences. The XAI analysis identifies Energy and IT as the most important sectors. While the latter might be attributed to the dot-com bubble as a temporal effect, the former should be expected to have a continuously high importance. Yet, the RCGCI-HHKD pipleine identifies neither of them as important causal drivers of the financial market. While this might in part be explained by the different time windows and databases used for these studies, it also highlights the differences between correlation and Granger causality. Interestingly, neither of them identifies the financial sectors as particularly important sectors even though one might intuitively expect the contrary. This might imply that these sectors either only act as a mediator within the network (and not as a driving force) or that their influence cannot be accurately analysed with the linear methods of Pearson's correlation or RCGCI.

Abbreviations

The following sector abbreviations were introduced by Ken French and are used in this chapter:

Aero	Aircraft
BldMt	Construction Materials
Cnstr	Construction
Drugs	Pharmaceutical Products
FabPr	Fabricated Products
Gold	Precious Metals
Hshld	Consumer Goods
Mach	Machinery
Meals	Restaurants, Hotels, Motels
MedEq	Medical Equipment
PerSv	Personal Services
RlEst	Real Estate
Rubbr	Rubber and Plastic Products
Softw	Computer Software
Trans	Transportation
Toys	Recreation

Table 5.1: Sector abbreviations according to [188].

Part III

Non-Ergodic Economics

Considering that this thesis is concerned with complex socio-economic systems on various time scales, one might raise the question whether the longer time scales can simply be derived from repeating the shorter ones and chaining them together. Originating from statistical physics, the concept of non-ergodicity reminds us that this is not always trivial. Only ergodic systems fulfil that the ensemble mean and the time average converge to the same values and only then can the conclusions from the dynamics of a single time step be generalised for a repetition over many time steps. Not recognising the non-ergodicity of a system can lead to wrong conclusions as will be explained later in this chapter. Transferring the considerations of ergodicity to financial and economic systems has given rise to ergodicity economics (EE) as a new branch of econophysics [210].
One major result of EE was to explain why insurance contracts can be beneficial for both participants, even though expected value theory might suggest that one party's profit is always at the other's expense [211, 212]. Chapter 6 extends the insurance model from EE by simulating it on a lattice of agents which can cooperate by agreeing on an insurance contract. This leads to the emergence of rich and poor neighbourhoods which superficially resembles kin selection, but arise endogenously from the cooperation dynamics. This chapter provides an explanation of how large-scale cooperation structures can arise over long time frames from the wealth optimisation of individual agents in the face of uncertainty. Hence, part III links the stochastic dynamics of the financial market from part II to the long-term societal dynamics from part IV.

6 Cooperation in a non-Ergodic World on a Network - Insurance and Beyond

'It is easy to mourn the lives we aren't living."

— Matt Haig, The Midnight Library

Abstract Cooperation between individuals is emergent in all parts of society, yet mechanistic reasons for this emergence are not fully understood in the literature. A specific example of this is the cooperation via insurances. But recent work has shown that assuming the risk individuals face is proportional to their wealth and optimising the time average growth rate rather than the ensemble average results in a non-zero-sum game, where both parties benefit from cooperation through insurance contracts. In a recent paper, Peters and Skjold present a simple agent-based model and show how, over time, agents that enter into such cooperatives outperform agents that do not. Here, we extend this work by restricting the possible connections between agents via a lattice network. Under these restrictions, we still find that all agents profit from cooperating through insurance. We further find that clusters of poor and rich agents emerge endogenously on the two-dimensional map and that wealth inequalities persist for a long duration, consistent with the phenomenon known as the poverty trap. By tuning the parameters which control the risk levels, we simulate both highly advantageous and extremely risky gambles and show that despite the qualitative shift in the type of risk, the findings are consistent.

This chapter and the associated appendix F are based on *Tobias Wand, Oliver Kamps and Benjamin Skjold, Chaos 34, 073137* (2024) [213]. Tobias Wand conceptualised this research project, wrote the code and carried out the research. Oliver Kamps and Benjamin Skjold supervised the project. All figures of this chapter and appendix F are taken from [213] and were created by Tobias Wand.

6.1 Introduction

Self-organisation of individuals is observed in many systems and is a cornerstone of complexity science [1, 22, 214, 215]. Mutual cooperation between agents exists in many forms, one such being insurance. The insurance industry is one of the largest industries in the modern economy and has existed for centuries. Yet, we find that mainstream economics, which is rooted in expected value theory, does not provide a satisfactory reason for its emergence: If both parties want to profit from the insurance contract in terms of expected value theory, the insurance buyer requires an insurance premium that is less expensive than the expected value of the risk. Conversely, the insurance seller demands a premium that is higher than the risk. No contract should be agreed upon under these assumptions, but of course, insurance contracts nevertheless exist in our economy. As an alternative approach, Peters and Adamou argue that the system in which insurance takes place is more reminiscent of a multiplicative system (i.e. that the risk is proportional to the wealth) and note that the increments in such a system are not ergodic and therefore expected value theory is not applicable [211]. Using time averages rather than ensemble averages, they show that there exist regimes in which insurance contracts are favourable for both buyer and seller, which calls for such contracts to be made. In a recent paper [33, 212], Peters and Skjold explore this idea using a simple agent-based model, which shows that agents who cooperate via insurance contracts over time systematically outperform agents who do not.

Inspired by the 2d Ising model and its many applications to socio-economic phenomena [3, 4, 45, 216, 217, 218, 219, 220, 221, 222, 223], we expand the model proposed in [33, 212] to large ensembles on a lattice network. Now, each agent is placed in a neighbourhood of other agents and can only cooperate with nearby agents, giving rise to the formation of spatial patterns in our model similar to the models in [216, 224] and resembling the crime hotspots found in [225] via weakly nonlinear analysis.

In this chapter, we first present a brief background of non-ergodicity in economics and illustrate the implication of broken ergodicity with a simple coin toss example before detailing the insurance paradox and its treatment in the context of ergodicity economics in section 6.2. Next, we describe the setup and our implementation of the agent-based insurance model on a network and the mathematical techniques used to analyse its results in section 6.3. Results for a specific parametric setting are presented in section 6.4 before we perform a parameter scan by varying the parameters c and r representing the relative costs or rewards of the risk in section 6.5, which reveals different clustering regimes in our model. Finally, we conclude with a summary and discussion of our work in section 6.6.

6.2 A Primer on Ergodicity Economics

In an ergodic system, an individual trajectory at different time steps $x_i(t_1), \ldots, x_i(t_n)$ has the same statistical properties (e.g. averages) as an ensemble observed at a single point in time $x_1(t), \ldots, x_n(t)$ for large n. This makes many calculations easier, and thus the ergodic hypothesis is often explicitly or implicitly assumed for many statistical methods. The core of ergodicity economics (EE) is a careful analysis of whether the ergodic hypothesis is true. EE thus makes an explicit distinction between ensemble averages $\langle x \rangle = \frac{1}{N} \sum_i x_i(t)$ and time averages $\bar{x} = \frac{1}{T} \sum_t x_i(t)$, e.g. for financial time series [210]. Assuming ergodicity implies that $\langle x \rangle = \bar{x}$ in the limit of large T and N. However, it is the exception rather than the norm that this is justified in real-world systems. A simple example of a non-ergodic system is the reinvesting coin toss gamble taken from [31].

6.2.1 Reinvesting Coin Tosses

Consider some initial wealth $x(0)$. You are now offered a gamble on which a fair coin is tossed, i.e. with probability 1/2, you win and with probability 1/2, you lose. If you win, your wealth is multiplied by a factor of $\alpha_w = 1.5$, and if you lose, your wealth is multiplied by a factor $\alpha_l = 0.6$. We will repeatedly toss the coin and sequentially update your wealth according to the outcome of the toss. Is this a favourable bet? The expected value computed via the ensemble average of identical agents with initial wealth x shows that

$$\langle x(t+1) | x(t) = x \rangle = \frac{\alpha_w x + \alpha_l x}{2} = 1.05x > x = x(t), \tag{6.1}$$

and therefore implies that the gamble is favourable. However, consider that for a single agent with initial wealth $x(0)$ and enough repetitions T, the law of large numbers guarantees that both coin toss outcomes will appear equally often. We can thus compute the time average

$$\bar{x}(T) = \alpha_w^{(T/2)} \alpha_l^{(T/2)} x = 0.9^{T/2} x \overset{T \to \infty}{\longrightarrow} 0. \tag{6.2}$$

Hence, the agent's wealth approaches 0 almost surely, implying that the gamble is not favourable. As proven via numerical simulations, this seemingly paradoxical situation can be explained by noting that the increasing ensemble average of the wealth trajectory is dominated by the asymmetry of the final wealth [210]: While almost all agents have negligible wealth, few have exponentially increasing wealth, which dominates the ensemble average.

6.2.2 Insurance Paradox

Consider a simple type of insurance, namely an agent, A, faces a risk and another agent, B, offers to take over that risk in exchange for receiving a fee, F, typically known as the insurance premium. The classical economic analysis of the insurance problem is based on the expected value model, which posits that agent A should accept to pay a fee lower than the expected value of the risk, whereas B should demand a fee higher than the expected value of the risk. This anti-symmetry results in no fee where both agents simultaneously are satisfied, and thus, no contracts should be agreed upon. The orthodox solution is usually either that the agents must have an information asymmetry or that risk aversion makes the agents deviate from the supposedly optimal decision [226, 227]. We find this unsatisfactory as the former is unjustified ad hoc reasoning and the latter simply raises a new question as to why such risk aversion exits.
Assuming that the risk an agent faces is proportional to their wealth and optimising time averages, however, Peters and Adamou show in [211] that there exist many situations in which agent A is willing to pay a higher fee than agent B demands. This framework provides a mechanistic reasoning for the existence of insurance contracts and shows that insurance is theoretically advantageous in the long run. These arguments are discussed in more detail in [33, 211] and a broader overview of cooperation in non-ergodic systems can be found in [228]. We highlight that behavioural experiments show evidence that humans are capable of heuristically optimising the time average growth rate when faced with risky decisions in laboratory settings [229, 230, 231], which gives credibility to the time solution of the insurance puzzle.

6.3 Methods and Models

6.3.1 Insurance among Agents

As we expand on the model from [33, 212, 232], it is prudent to review its results alongside its model specifications. This model consists of n agents $i = 1, \ldots, n$, which are observed at time $t \in (0, T)$. Each agent's wealth is initialised at 1, $w_i(0) = 1 \ \forall i$, and in each time step, an agent A with wealth $w_A(t)$ is randomly chosen to face a risky gamble: with probability p, A either loses a relative amount of its wealth $c \cdot w_A(t)$ or with probability $1 - p$ wins a relative amount $r \cdot w_A(t)$. Note that the simulations in [33, 212] use $c = 0.95$, $r = 0$ and $p = 0.05$ as parameters, but that the qualitative results do not depend on the exact parameter choice [232].
When faced with a risky gamble, agent A approaches another agent B for an insurance

contract. Agent B will have a minimum fee, $F_{\min}$, where they will accept to take over A's risk and agent A will have a maximum fee, $F_{\max}$, they will accept to pay. If $F_{\min} < F_{\max}$, there exist a range of fees in which signing the contract is mutually beneficial for the agents, such that the risk and the fee is transferred from agent A to agent B.[1] To ensure the results are not conflated by the number of agents used, we distinguish between a time step, t, and a time unit, t_u, with a time unit being N^2 time steps, i.e. each agent is on average chosen to face a risk ones per time unit. Note that the original model in [33, 212] is well-mixed, i.e. all other agents are equally likely to be approached as agent B without any spatial restriction. The network model presented in this manuscript lifts this assumption and restricts the agents' interaction to a lattice network.

The EE solution to the insurance problem is to focus on time average growth rates. Recognising that the underlying process of the system is mainly multiplicative[2], it follows that additive wealth increments are not ergodic, but logarithmic wealth increments are and, therefore, better indicate what happens over time and allow us to assume an underlying growth rate model $g(\cdot)$ [31, 211]. Comparing agents' growth rates with $g_{A,B}^{with}$ and without $g_{A,B}^{without}$ signing an insurance contract at some fee F, we can calculate the maximum fee $F_{\max}$ the agent facing the risk (agent A) is willing to pay, as well as the minimum fee $F_{\min}$ another agent (agent B) is willing to accept. We therefore get that

$$\underbrace{\frac{1}{\Delta t}\ln\left(\frac{w_A - F}{w_A}\right)}_{g_A^{\text{with}}} = \underbrace{\frac{1}{\Delta t}\left(p\ln\left(\frac{w_A - cw_A}{w_A}\right) + (1-p)\ln\left(\frac{w_A + rw_A}{w_A}\right)\right)}_{g_A^{\text{without}}} \tag{6.3}$$

$$\Leftrightarrow \quad F_{\max} = w_A\left[1 - \exp\left(g_A^{\text{without}}\right)\right]$$

for agent A and

$$\underbrace{(1-p)\frac{1}{\Delta t}\ln\left(\frac{w_B + F + rw_A}{w_B}\right) + p\frac{1}{\Delta t}\ln\left(\frac{w_B + F - cw_A}{w_B}\right)}_{g_B^{\text{with}}} = \underbrace{0}_{g_B^{\text{without}}} \tag{6.4}$$

[1] For stability reasons we restrict contracts to be formed by enforcing $w_B(t) \geq cw_A(t)$ to explicitly ensure agent B can pay for the cost $cw_A(t)$ without defaulting. However, we note that this is not strictly necessary but has no qualitative effect on the results.

[2] We note that the multiplicative structure of the simulations is not entirely realistic, and even within the simulation, the wealth increments are not fully multiplicative because the transferred fee F is an additive wealth change and depends on both w_A and w_B. However, these subtle distinctions are outside the scope of this analysis and part of ongoing research.

for agent B, where $g_B^{\text{without}} = 0$ because without the insurance, only agent A experiences any change during the observed time step. Equation (6.4) does not have a general closed-form solution, but can be solved for $p = 1/2$. We therefore use this parameter value for our simulation.[3] For $p = 1/2$, we get that $F_{\max}$ and $F_{\min}$ are given by

$$\begin{aligned} F_{\max} &= w_A - (w_A + r w_A)^{0.5}(w_A - c w_A)^{0.5} \\ F_{\min} &= -w_B + \frac{1}{2}\sqrt{4w_B^2 + (c+r)^2 w_A^2} + \frac{(c-r)w_A}{2}. \end{aligned} \tag{6.5}$$

For suitable w_A and w_B, $F_{\min} < F_{\max}$ is fulfilled and therefore, an interval of fees F in $[F_{\min}, F_{\max}]$ exist for which insurance contracts are mutually beneficial for both agents. For simplicity, choose then the midpoint fee $F = \frac{1}{2}(F_{\min} + F_{\max})$ to sign the contract. As shown in figure 3 of [33], the agents who sign contracts based on this criterion outperform agents who do not (for example, expected value optimisers). Interestingly, the simulations show how "large" agents, who act as the insurance company for all other agents, emerge. However, if such an agent becomes "too large", they lose the ability for other agents to insure its large risk and thus, its growth rate declines until it is no longer larger than the other agents.

6.3.2 Simulation on a Network

The model in [33] is deliberately simple but showcases two interesting findings: First, all agents participating in the insurance contracts profit compared to the uninsured agents. Second, even though the cooperation reduces the overall wealth inequality compared to a model without insurance, it still endogenously creates large wealth differences, leading to the emergence of a large insurance-company-like agent.
However, the well-mixed setup in which all agents can approach each other is only realistic for small systems. Therefore, we propose to perform the simulation on a lattice and allow each agent to approach only its nearest neighbours to negotiate an insurance contract. The lattice simulation allows us to investigate the spatial clustering of agents based on relative wealth levels. We use a lattice coordination number of four and periodic boundary conditions such that the first and last agents of each row and column are treated as neighbours. Throughout the simulations, we use a square lattice with length $N = 64$, i.e. $N \times N = 2^{12}$ agents and our code is available at [233].

[3] This is especially useful because numerical optimisation of equation (6.4) breaks down for sufficiently small wealth levels.

6.3.3 Clustering Evaluation Methods

The main difference between our model and the one presented in [33] is the spatial constraint, which gives rise to the analysis of spatial clustering. We, however, first investigate the temporal clustering and compare that to the temporal clustering from the model in [33].

Temporal Clustering

To evaluate the temporal clustering of agents, i.e. whether the rich stay rich and the poor stay poor, we use Spearman's rank correlation. All agents are ordered in an enumeration $i = 1, \ldots, N^2$, and we record their ranks $\rho_i(t_u)$ as the k^{th} richest agent for each time unit t_u. Spearman's rank correlation is now the regular Pearson's correlation between $\rho(t_u)$ and $\rho(t'_u)$. For each time lag Δt_u, multiple pairs $t_u - t'_u = \Delta t_u$ exist, and therefore, we calculate Spearman's correlation for all pairs with a given time lag and compute the mean and standard deviations for all of them. Interestingly, Spearman's correlation and Pearson's correlation of the logarithmic wealth values show almost the same results.

Spatial Clustering

A naive approach to evaluate the neighbourhood formation of the richest or poorest quantile is formulated by defining a clustering coefficient κ_{clust} for each group as the relative amount of agents in this group which have at least one neighbour from their own group. For the definition of the groups via deciles and for a sufficiently large ensemble $N \gg 1$,[4] the expected clustering coefficient for a purely random ensemble can be calculated as

$$\kappa_{clust}^{random} = 1 - (1 - 0.1)^4 \approx 0.34, \tag{6.6}$$

and is the chance that none of the i.i.d. four neighbourhood sites is occupied by another member from the in-group. Approximate confidence intervals of this quantity can be calculated empirically by simulating a large number of random matrices and evaluating their κ_{clust}.

An alternative way to evaluate the spatial clustering closer to the framework of Ising-like models is to compute the spatial autocorrelation. To this end, the ensemble mat-

[4] With sufficiently large, we mean that when checking whether any of the four sites is occupied by another member of the same decile, it can be modelled as sampling without replacement from a group with a relative size of $\frac{1}{10}$. This holds if $\frac{0.1 \cdot N^2 - 1}{N^2} \approx \frac{1}{10}$.

rix is transformed into a binary matrix $\mathbf{B}$ with $B_{ij} = 1$ if a member of the group under consideration (i.e. the richest or poorest decile) is in the i^{th} row of the j^{th} column and 0 otherwise. The autocorrelation matrix is then computed via the function *scipy.signal.correlate2d* [234] by correlating $\mathbf{B}$ with itself and using the settings *mode ="full", boundary = "wrap", fillvalue=0.* This method computes the full discrete linear cross-correlation of $\mathbf{B}$ with itself as a 2-dimensional array. For each offset (k, l), it overlays the matrix $\mathbf{B}$ over itself with an offset of (k, l). The value of the $(k, l)^{\mathrm{th}}$ entry of the 2-dimensional correlation matrix is then the sum of the element-wise products which can be positive or negative and can be normalised to lie between -1 and 1. Note that the boundary condition *boundary = "wrap"* means that if the matrix shifted by (k, l) does not fully lie on the original matrix, the original matrix is periodically extended with its own entries.

6.4 Results: Typical Observations

In an initial run, we set the simulation time $T_u = 200$ and the risk parameters $c = 0.6$ and $r = 1.4$ with equal probability $p = 0.5$. These parameters are deliberately chosen, such that for an agent without any insurance, this yields the interesting regime of

$$\langle w(t+1)\rangle = \frac{(1-c)+(1+r)}{2} w(t) = 1.4w(t) > w(t) \tag{6.7}$$

$$\bar{w}(t+1) = \sqrt{(1-0.6)(1+1.4)}w(t) = \sqrt{0.96}w(t) < w(t), \tag{6.8}$$

i.e. from the ensemble perspective, it is advantageous to keep the risk, but from the time perspective, it is not. This also ensures that the trajectories do not decay too quickly to zero because $\sqrt{0.96} \lesssim 1$, thereby guaranteeing numerical stability of the simulation. As a general remark, the wealth distribution quickly develops a fat right tail which at least superficially is reminiscent of a power law distribution similar to what is found in many empirical wealth distributions.

6.4.1 Temporal Clustering

In figure 6.1, we show Spearman's rank correlation, which reveals that even for large time differences, there is a strong ranking correlation showcasing that there is a long memory in the agents' wealth ranking. This means that the general wealth ranking is preserved for extended periods of time, i.e. that, for example rich agents stay rich for long time intervals. This is similar to what is observed in [33] as shown for the mixed

network in figure 6.1 and can be related to the well-documented poverty trap observed in real societies [235, 236].

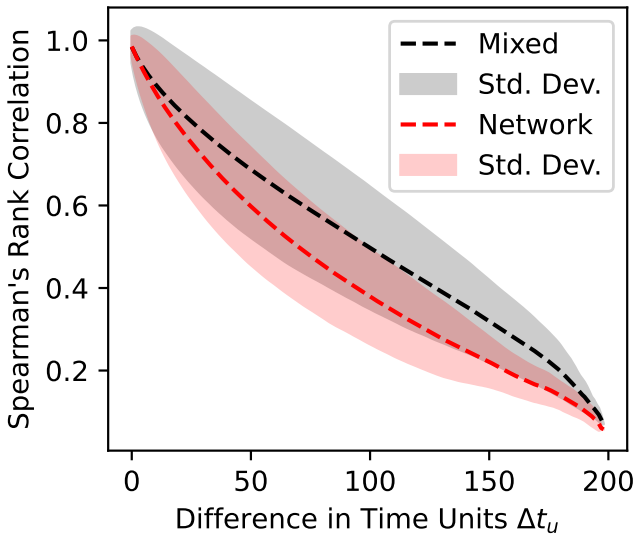

Figure 6.1: Spearman's rank correlation shows a highly significant correlation in the ranking order of the agents' wealth levels even after long time lags for both our network model and the mixed model in [33]. The parameters used for this simulation are $T_u = 200$, $N = 64$, $c = 0.6$ and $r = 1.4$.

6.4.2 Spatial Clustering

After $T_u = 200$ time units, we observe a notable pattern when plotting the spatial distribution of the richest or poorest 10% of the agents in figure 6.2. Both deciles form small clusters of neighbourhoods where they have mostly neighbours from within their own group. Using the clustering coefficient defined in section 6.3, we find that in the given parameter setting, both quantiles converge to the same level of clustering and far exceed the expectation for random matrices, indicating that a significant level of clustering emerges from the system's dynamics (figure 6.3 left panel). Note that even within a neighbourhood, the wealth levels can still be heterogeneous and vary wildly, even though they all fall in the same decile of the wealth distribution.

We further find that (after normalising the autocorrelation matrix with its maximum) there are no noteworthy differences between the row-wise and column-wise autocorrelation (corresponding to the x and y directions). Both directions for both the top and bottom decile show a notably higher autocorrelation for the first nontrivial lag than the

Figure 6.2: On the $N \times N$ lattice grid, we highlight the richest 10% of agents (left panel) and the poorest 10% of agents (middle panel) in our model. As a comparison, we also show a purely random ensemble of agents (right panel) and highlight its 10% richest agents. Our model shows a marked pattern of clustering into neighbourhoods when compared to the random matrix. The parameters used for this simulation were $T_u = 200$, $N = 64$, $c = 0.6$ and $r = 1.4$.

random matrix, thereby confirming the analysis of the clustering coefficients that rich and poor neighbourhoods emerge in the system (figure 6.3 right panel).

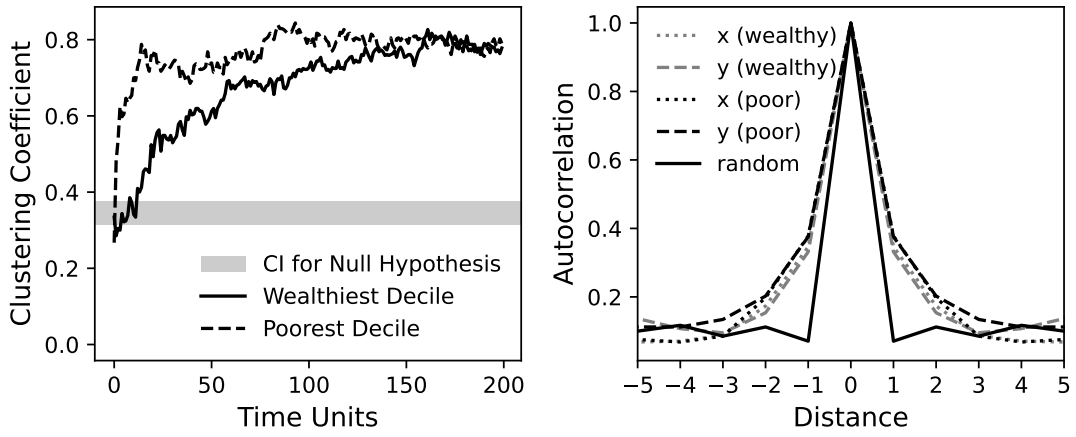

Figure 6.3: Left: Clustering coefficients of the richest and poorest deciles are far above the expectation for random ensembles. While the clustering emerges faster for the poorest than for the wealthiest decile, both seem to converge to similar levels. The confidence interval for the null hypothesis that the geographical distribution is purely random is calculated via eq. (6.6) and sampling random matrices. Right: Autocorrelation function (ACF) for the wealthiest and poorest deciles in x and y direction. Distance is taken in units of the agents' positions on the lattice network. All show essentially the same behaviour and a notably higher value at the first nontrivial distance than for the random matrix. The parameters used for this simulation are $T_u = 200$, $N = 64$, $c = 0.6$ and $r = 1.4$ for both graphs.

6.5 Parameter Scan

To check if the clustering observed in the previous section is a general feature of the model or an artefact of the specific parameter configuration, we perform a parameter scan across the space of (r, c). For constant $N = 64$ and $T_u = 200$, we vary c across $0, 0.05, \ldots, 0.95$ and r across $0, 0.05, \ldots, 2$ and calculate the clustering coefficients for each pairing. We show the results in the (r, c) phase space in figure 6.4. By repeating the analysis with different values for T_u, we show that the qualitative behaviour is robust, though with the clustering in both regimes slowly increasing with larger T_u.

For both the top and bottom decile, we find a gradual increase in clustering with c and r (figure 6.4). However, for the bottom decile, we find a much more pronounced behaviour with a well-defined region of high clustering in the centre of the phase space. We explore this further in section 6.5.1, where we show that its borders (black lines on right panel of figure 6.4) can be derived analytically and reflect the onset of degenerate values for the fees $F_{\min\,/\,\max}$.

To visualise how suddenly the clustering emerges, we scan through a slice of the phase space at $r = 0.5$ by varying c and plot the autocorrelation function for the top and bottom decile. Normalising the first nontrivial value of the ACF by its maximum AC_1/AC_0 reveals a well-defined maximum for the clustering of the bottom decile as shown in figure 6.5. On the contrary, the top decile (see figure F.1 in the appendix) only has a linear increase of AC_1/AC_0 with c without any indications of a discontinuity or maximum. An exploration of the highly volatile regime in the top right corner of the phase space is given in appendix F.2.

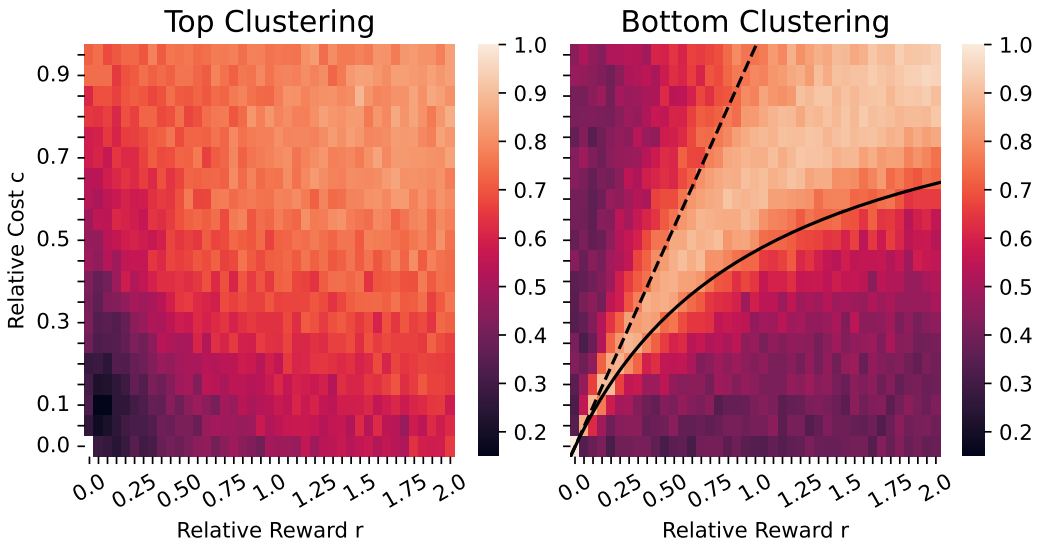

Figure 6.4: For $T_u = 200$ and $N = 64$, c and r are varied and the clustering coefficients for the top and bottom decile are recorded after the last iteration. Note that $c = r = 0.0$ means the absence of gambling dynamics and, therefore, shows no dynamical behaviour at all. The regime for high clustering of the bottom decile has contours given by the black lines. These lines correspond to the onset of degenerate regimes for $F_{\min\,/\,\max}$ and are derived analytically as equations (6.9) and (6.10) in section 6.5.1.

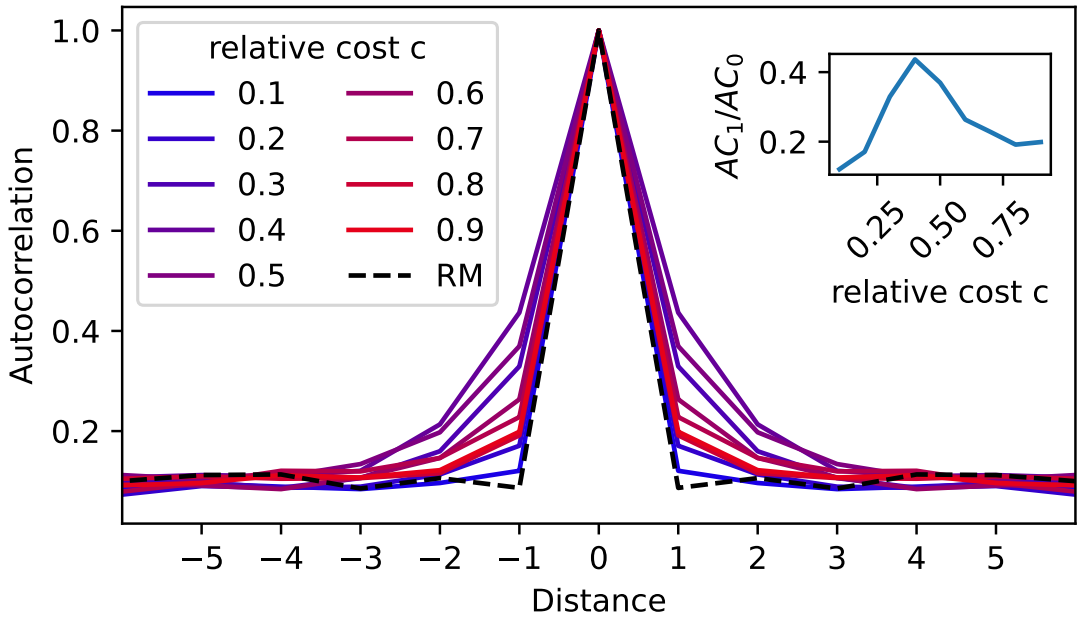

Figure 6.5: For constant $r = 0.5$, c is varied and the autocorrelation of the poorest decile is plotted. The displayed ACFs are the mean values across the slices in the x and y directions shown in the right part of figure 6.3. The inset depicts the strength of the first nontrivial autocorrelation value AC_1 normalised by the maximum autocorrelation AC_0 which reaches a well-defined maximum at $c = 0.4$.

6.5.1 Degenerate Regimes

Insurance can be thought of as a special form of cooperation in the face of uncertainty. The model described in [33, 211, 212] includes a stochastic risk which is entirely negative ($c > 0$, $r = 0$), but the parameter scan in this chapter explores a more general setup, where both $c \geq 0$ and $r \geq 0$ are varied. Hence, we also simulate regimes where the gamble is actually favourable to agent A ($r \gg c \gtrsim 0$). Though even in such a positive regime, growth rate optimisation might incentivise an agent A to rather accept a deterministic fixed payment instead of a fluctuating reward. Because this regime is reached by simple parameter variation, it is not a qualitatively new model and can still be simulated and analysed with the same methods as above. As these regimes describe a different kind of cooperation in an uncertain environment than what we traditionally think of as "insurance", we decided to call this a degenerate regime to stress that the model has adopted a new behaviour. The question arises; how does this situation with beneficial risk parameters $r \gg c \gtrsim 0$ affect the fees?

The first degenerate regime is defined as $F_{\max} < 0$, i.e. the risk is so favourable for agent A that they demand to receive money in exchange for giving away the risk. The border of this regime can be derived from equation (6.5) and is given by

$$c = 1 - \frac{1}{1+r}. \tag{6.9}$$

Setting the time average growth rate $g_A^{\text{without}} \stackrel{!}{=} 0$, i.e. having no time average growth or decline for the individual, leads to the same condition for c. This degeneracy line is shown as a solid black curve in the right panel of figure 6.4. Everything below this curve results in $F_{\max} < 0$.

Similarly, from agent B's perspective, the question becomes, when is $F_{\min} < 0$, i.e. when is agent B willing to pay the agent A to take over the risk? While it may appear strange that B should pay A to take over the risk, consider a situation in which A holds a risky asset which might either default or dramatically increase its value. If agent B can stomach the loss in case of a default without any problems, the potential upside might be alluring enough to pay A in order to purchase this risky asset. Because both agents' wealth levels w_A and w_B are necessary to calculate $F_{\min}$ in (6.5) and their relative difference w_A/w_B can vary wildly, there exists no closed analytical solution to this. However, in the case of $w_B \gg w_A$ we can rewrite (6.5) as

$$\begin{aligned} 0 \stackrel{!}{=} F_{\min} &= -w_B + \frac{1}{2}\sqrt{4w_B^2 + (c+r)^2 w_A^2} + \frac{(c-r)w_A}{2} \\ &= -w_B + \frac{1}{2} w_B \sqrt{4+\epsilon^2} + \frac{(c-r)w_A}{2} \\ &= -w_B + \frac{1}{2} w_B \left(2 + \mathcal{O}\left(\epsilon^2\right)\right) + \frac{(c-r)w_A}{2} \\ &\approx \frac{(c-r)w_A}{2} \\ \Rightarrow c &= r, \end{aligned} \tag{6.10}$$

by setting $\epsilon = \frac{(c+r)w_A}{w_B} \ll 1$ and using first-order Taylor expansion.[5] The condition $c = r$ is shown as a dashed line in the right panel of figure 6.4.
As seen in figure 6.4 (left panel), there is no immediate connection between the degeneracy regimes and the clustering in the top decile. However, for the poorest decile (figure 6.4,

[5] Note that as discussed in [211], when $w_A/w_B \to 0$, the time average of the risk approaches the expected value from agent B's perspective. This expression is therefore identical to using the expected value operator.

right panel), the degeneracy lines approximate the contours of the regime of high clustering. Clustering into poor neighbourhoods, thus, happens if and only if $c \geq 1 - \frac{1}{1+r}$ ((6.9)) and simultaneously $r > c$ ((6.10) under the condition that $w_B \gg w_A$).

6.6 Discussion

6.6.1 Summary

The model presented in this chapter is an extension of the agent-based model presented in [33, 212]. We find that despite restricting the agents' ability to make contracts to their nearest neighbours on a spatially restricted grid, the overall conclusion from [33], i.e. that insurance is advantageous in the long run for all agents, still holds. However, we find that the general wealth ranking is preserved for shorter periods of time, indicating that not having spatial constraints leads to what can be thought of as monopoly-like structures, consistent with tendencies in society with increasingly global access through the internet. We, though, highlight that the memory in the system with spatial constraints is still high, which is consistent with the well-documented poverty trap observed in real societies [235, 236]. We further find that this restriction leads to a rich phenomenology of spatial cluster formations, again similar to what we see in the real world. While the clustering of the wealthiest agents increases continuously for larger parameters r and c, we find that the poorest agents only exhibit high clustering for particular combinations of (r, c) and that this regime can be approximated by the degeneracy lines.

6.6.2 Model Relevance

While including spatial constraints to the insurance model presented in [33] is an attempt to explore a more realistic system, we recognise that the system is still a very simplified representation of the world. Although this would be closer to reality, we have decided not to include specific insurance companies rather than having the insurer simply be another agent in the system. However, this is an unnecessary restriction to the model: Even if a large insurance company exist, one still has to rely either on the non-satisfactory solutions discussed earlier or on the time solution provided here to solve the insurance paradox. As the argument often is that such an insurance company can offer contracts as expected value (either through having such high wealth that the time average of the risk approaches the expected value or through having access to the actual ensemble), cooperation through insurance would only be stronger than what we present. We, therefore, see our model as more general.

Within the broader scope of this thesis, it is interesting to think about which time scale this model's dynamics correspond to. Considering that this model describes the emergence of cooperation clusters between agents in fixed positions from isolated origins, it might make sense to interpret each agent as an early settlement at the onset of the agricultural revolution. Then, the passing of one time unit as a period of uncertain reward or risk can be interpreted as a harvest cycle which means that its time scale is on the order of one year.

6.6.3 Relation to Classical Cooperation

This type of insurance can be considered a restricted form of cooperation. Although the setup on the lattice network restricts the agents in our model, the cooperation is still beneficial for them, which is in line with previous findings. Studies on evolutionary game theory have shown that cooperative strategies can become dominant over the defectors' strategy purely due to the effect of fluctuations on the reward [237, 238]. This has also been studied within the time average framework, where cooperation is similarly found to be stable, i.e. defecting is not advantageous [239], which has been backed by empirical research [240, 241, 242]. An important note, however, on the pure form of cooperation in these studies is that trust does play a role. Despite the fact that full cooperation is mathematically stable and defecting can be shown not to be advantageous, this is only true in the long run; if you give a large part of your wealth to someone in an attempt to cooperate and they decide to defect, it is, of course, not advantageous for you.
This is not a consideration in our model, as both agents improve their growth rate in every time step when they sign a contract and therefore need not worry about what the other agent does in the future: taking the deal is good regardless. Whether there is a path from non-cooperation through insurance-like cooperation to full cooperation is the subject of ongoing research.

6.6.4 Spatial Clustering

The main feature of this model compared to [33] is the study of spatial cluster formations. Whenever the volatility (i.e. the variation in the gamble's payoff) is high enough, the top decile of agents tends to cluster into neighbourhoods while the bottom decile displays a similar behaviour in a specific parameter regime. This regime is bound by the degeneracy lines which will be further discussed in section 6.6.6 and is characterised by $F_{\max} > 0 > F_{\min}$ for $w_A \ll w_B$. B is willing to transfer wealth to A in order to obtain the fluctuating payoff and A gladly accepts this as A would actually be willing to pay

money in order to get rid of the varying risk. However, this only holds for $w_A \ll w_B$ and if A is in a poor neighbourhood without such rich neighbours, it has no chance to profit from the beneficial cash transfer that a rich agent would offer and A will remain as poor as its neighbourhood. Further discussion on the validity of this regime is in section 6.6.6. The emergence of rich and poor neighbourhoods might superficially resemble kin selection [243]. However, such a mechanism is not present in our model's algorithm. Instead, the clustering emerges endogenously: Rich agents do not choose to connect with other rich agents but rather become and stay rich because there are other rich agents in their neighbourhoods. Indeed, the original model already shows that a rich agent eventually collapses if no other agent is rich enough to insure their risk [33, 212, 244]. Hence, with our spatial constraint and whenever the volatility is high enough, only agents with other rich neighbours have a chance to stay rich in the long run. This phenomenon might be interpreted as a special case of network reciprocity described by Nowak [243] and is not restricted to a narrow regime in the phase space, but a general feature that emerges whenever the volatility (quantified by c and r) is high.

6.6.5 Persistent Inequality

Real economies have a strong memory effect, i.e. rich people tend to stay rich and poor people tend to stay poor and remain in the so-called poverty trap. The well-mixed and network model both have such memory effects as shown in figure 6.1. We find that the inequality memory quantified via the Spearman correlation is higher in the well-mixed model from [33] than in the spatially constrained network model. This is an interesting finding when we compare it to the two theories for the emergence of poverty traps discussed in [235]: scarcity-driven and friction-driven. The scarcity-driven poverty traps describe a different behaviour of agents under the pressure of extreme scarcity, whereas the friction-driven poverty traps describe situations in which market inefficiencies or different initial conditions lead agents with identical decision criteria to different wealth trajectories.
Interestingly, our findings seem to contradict the friction-driven theory of poverty traps: Including spatial constraints clearly a market inefficiency, but we find that the inequality persistence of the system is greater without it. We suspect that this can be explained by monopoly-like structures which emerge in the well-mixed system but are not possible with the spatial constraint. Note that in both our and the well-mixed system, there are no inefficiencies in forming contracts (such as administration costs), nor any agents acting in bad faith (such as agents demanding higher prices than based on their own assessment), which might change these results.

6.6.6 Generalising the Original Model Gives Rise to Negative Fees

Another extension of [33] is to include a positive outcome of the gamble with reward $r > 0$ in addition to the harmful cost c. We show that this can lead to regimes where it might be mutually beneficial to have $F \leq 0$, i.e. the "insurer" pays an amount to take over the risk. The phase space spanned by the two parameters (r, c) in figure 6.4 shows that the bottom quantile's clustering corresponds to the two degeneracy lines (6.9) and (6.10) derived in section 6.5.1. Equation (6.9) marks the regime at which $F_{\max} < 0$ and the time average growth rate of the gamble becomes neutral, while (6.10) indicates the regime where the ensemble average becomes neutral and, under the condition $w_A \ll w_B$, $F_{\min}$ can become negative. It might be tempting to dismiss the regimes with negative fees as beyond the scope of an insurance model. However, insurance is just a special case of cooperation in the face of uncertainty and, as will be discussed at the end of this section, this degenerate regime can be interpreted as a relationship between employee A and employer B. Our analysis shows that there are no qualitative differences or irregularities in the model behaviour in the different regimes:

First, the condition $F_{\min} < F_{\max}$ is still a meaningful model of reality. Consider $F_{\min} < 0 < F_{\max}$ as discussed above. The negative $F_{\min} < 0$ of agent B means that they want to acquire the risk, while the positive $F_{\max} > 0$ of agent A means that they want to get rid of it. In this scenario, e.g. a transfer at $F = 0$ (fulfilling $F_{\min} < F = 0 < F_{\max}$) is beneficial for both of them. Alternatively, if both $F_{\min} < F_{\max} < 0$, then one can more easily understand the setup by reversing the signs and considering the price $P_B = -F_{\min}$ that agent B is willing to pay and $P_A = -F_{\max}$ that agent A demands. Now, $0 < P_A < P_B$ and agent B is willing to match agent A's demanded price. The interested reader can play around with the parameters for $F_{\min} < 0 < F_{\max}$ and $F_{\min} < F_{\max} < 0$ to verify that the condition $F_{\min} < F_{\max}$ still yields sensible results.

Second, the regime of high clustering is characterised by $F_{\max} > 0$ and the possibility that $F_{\min} < 0$ if $w_A \ll w_B$, i.e. if B is much wealthier than A, it will consider the gamble (whose reward and cost are relative to A's wealth) so advantageous that B offers A money ($F_{\min} < 0$) to reap the potential rewards. At the same time, the risk c is so large relative to A's wealth that A is willing to pay money in order to get rid of the risk. This is just a more extreme version of the risk assessment $0 < F_{\min} < F_{\max}$ in the non-degenerate regime and illustrates that purely by considering time averages, the agents can come to mutually beneficial deals without assuming the existence of any subjective utility functions or subjective risk assessment [211]. Both agents operate under the same rules, but

the time average considerations lead them to different risk assessments based on their individual wealth $w_{A/B}$ whereas expected value theory would fail to explain this scenario.

Third, if $F_{\min} < F_{\max} < 0$, one can rethink this setup as modelling employment rather than insurance contracts. While insurances are used to replace A's overall detrimental risk with a fixed payment from A to B, one can think about employment in similar terms. Now, A can either work as a self-sustained freelancer and, while being mostly profitable, will have varying success. Or A can be hired by another agent B and give B the (varying) rewards of its work while, in return, getting a fixed income (a negative fee $F < 0$ corresponds to A getting $-F$ from B). This regime is considered in more detail in [245] but arises naturally from our model's parameter scan as another regime of cooperation in the face of uncertainty.

6.6.7 Further Work

We consider this a first attempt to generalise the model presented in [33]. First, we find qualitatively similar results on the benefits in the long run of evaluating the risk using time average growth rates, which leaves the question of whether any non-trivial network structures may lead to qualitatively different results. Second, our spatial constraint reveals a decreased memory in wealth ranking. The question now is: How much does the temporal autocorrelation change if the neighbourhood of the agents is expanded or if the network structure is changed? And are there any configurations that lead to qualitatively different results? Finally, we recognise that consumption is an important feature in real systems [246]. This leads to the possibility of bankruptcy, which in this model is equal to being erased from the system, but we are not aware of any solutions for dealing with such a problem.

Part IV

Data-Driven Insights into Pre-Modern Societies

Cliodynamics is the study of the emergence, long-term development and dissolution of large groups of humans via quantitative methods [47]. The notion that statistical patterns might predict the behaviour of such systems had been discussed by historians [247] and in science fiction literature [248] alike, but only recent access to large databases and modern statistical methods has allowed this research field to flourish. The demographic fiscal model was one of the first major results of this research and it derives coupled differential equations for the population density N and resources S of a state with carrying capacity $k(S)$ to describe the emergence and collapse of historical empires and to replicate historical population data [48, 249].

Particularly the *Seshat* databank has sparked numerous new directions of cliodynamics research [250]. Seshat contains historical records of more than 500 historical polities spread across over 30 geographical areas and covering several millennia. Archaeologists, historians and other scholars collaborate to quantify continuous variables like the population of the largest settlement or the degree of literacy and ordinal variables like the development of a monetary system (e.g. trading goods, using a natural resource like cowrie shells or gold as currency or even adopting paper money) or the diversification and specialisation of social roles (e.g. one person is simultaneously king, high priest and military commander or these rules are given to different people). Due to the scarcity of accurate data, proxy measurements often have to be derived from academic sources and statistical techniques are often necessary, e.g. to aggregate disagreeing expert assessments or to impute missing data [251]. The databank is constantly updated to account for new research and to expand its geographic scope [252]. More details on the databank will be given in the following chapters and in appendix G.3. Seshat allows researchers to quantitatively analyse the emergence of complex social systems [50, 253] or to test causal hypotheses about the drivers of socio-economic complexity [52, 254] and the field of cliodynamics still benefits from an interplay between domain expertise, data science and insights from complex systems research [53, 54]. Recently, new efforts have been undertaken to represent the spatio-temporal data from Seshat as structured geographic information [255], but new analyses and results based on this data are still pending.

The Seshat project seeks to record the state-of-the-art knowledge about past societies, but several possible challenges arise in the collection and assessment of historical data. One of them is that short-lived polities with a small geographic scope of action might simply not appear in the data records. This constitutes a survivorship bias and will be analysed in chapter 7. Afterwards, a characteristic time scale for the growth of socio-economic complexity is derived in chapter 8.

7 A Bayesian Approach to Survivorship Bias in Historical Data Analysis

"As we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tends to be the difficult ones."

— Donald Rumsfeld, US DoD news briefing

Abstract Datasets such as *Seshat* have allowed researchers to quantitatively test hypotheses about premodern societies and states with great success. Nevertheless, one has to consider potential sources of bias in the data such as a survivorship bias favouring the inclusion of long-lived over short-lived states. If we are not aware that some instances are missing in our sample because of a systematic bias, we might draw wrong conclusions. Bayesian methods can be used to complement standard modelling procedures to take this issue into account as is demonstrated by analysing the longevity distribution of premodern states.

This chapter is based on *Tobias Wand, Cliodynamics 15, pp. 99-106* (2024) [256]. Figure 7.1 is taken from [256] and was created by Tobias Wand.

7.1 Introduction

Cliodynamics has paved the way for the scientific analysis of human societies on long time scales via modelling and comparison to empirical data [48]. Complemented by the efforts of the Seshat databank [49], this methodology allows for detailed quantitative analyses of how the sampled states and societies developed [50, 253]. In a recent article, Scheffer et al. have used the Seshat data in combination with their own Moros database to analyse the longevity of historical states by modelling the states' vulnerability or resilience against existential threats [257]. They conclude that the resilience of states decreases during their first 200 years of existence until it reaches a plateau. Efforts are currently underway to find a microscopic model that can reproduce such behaviour [54]. However, the data gathering might suggest an alternative explanation: What if the states that collapsed within their first 200 years simply did not leave enough traces behind to be included in the dataset, i.e. what if a survivorship bias makes us believe that young states were more resilient than older states? As an example, the present chapter explores this possibility for the resilience analysis of Scheffer et al., but its main idea should be kept in mind whenever historical data is analysed.

7.2 Modelling the Data

7.2.1 Resilience Analysis by Scheffer et al.

The resilience of a system can be defined as its tendency to return to its original state after a shock whereas vulnerability describes conversely the inability to restore the original state [257]. These concepts are strongly connected to the stability of fixed points in section 2.1 and chapter 3. Observing historical societies as the systems of interest, one can use their longevity distribution as a measure of how resilient they have been because a non-resilient state would collapse after it faces the first major crisis. Results from complexity science show that if the resilience of a state is constant throughout its lifetime, the resulting ensemble of recorded life times will resemble an exponential survival distribution

$$f_s(t) = \exp\left(-\lambda_s t\right). \tag{7.1}$$

Comparing this to empirical data, it is shown that the naive f_s overpredicts the number of states that should have collapsed at a fairly young age whereas the Seshat and Moros data have a well-defined maximum in their historgram distribution at around 200 years.

Hence, Scheffer et al. develop the Saturating Risk model which assumes that the resilience is high for young states (i.e. just very few of them collapse) and decreases towards a constant plateau after about 200 years. Compared to other alternative mechanisms that lead to different survival functions, Scheffer et al. show that the Saturating Risk model has the best fit to the data by using the AIC model selection criterion from equation (2.13) where a lower AIC indicates a better fit [66].

7.2.2 Bayesian Modelling of Survivorship Bias

What if the amount of states which died young is not *over*estimated by the naive model but rather *under*sampled by the data gathering? If a premodern state collapsed within the first few years after its foundation, it may well have left so few artefacts behind that archeologists have not been able to recover any traces of its existence, thereby causing this state to be lost to time. Additionally, short-lived sociopolitical structures may have been classified as revolutionary periods or uprisings rather than full-fledged states (e.g. the eleven years of the republican Commonwealth of England). And because the Seshat data was mostly sampled at a frequency of 100 years, short-lived states which did not overlap with the turn of a century either must be deliberately included by researchers or simply fall by the wayside. These mechanisms could be responsible for an undersampling of such states and therefore reflect a survivorship bias in databanks like Seshat or Moros.

Such concerns can be taken into account by using Bayesian statistics and incorporating these beliefs into a prior distribution [55]. As a simple example, the chance of a political formation with age t to be considered a state by researchers and included in the data may be modelled naively by an exponential discovery rate via the prior distribution

$$f_p(t) = 1 - \exp(-\lambda_p t). \tag{7.2}$$

Hence, young states with age close to zero have a very low chance to be included in the datasets whereas long-lived states converge to a probability 1 to almost surely be included. Assuming that the states now follow the naive longevity distribution for constant resilience (irrespective of whether they have been discovered by researchers or not), the observed longevity distribution is proportional to the product of these equations

$$f_B(t) \sim f_p(t) f_s(t) = \exp(-\lambda_s t) - \exp(-(\lambda_p + \lambda_s)t) \tag{7.3}$$

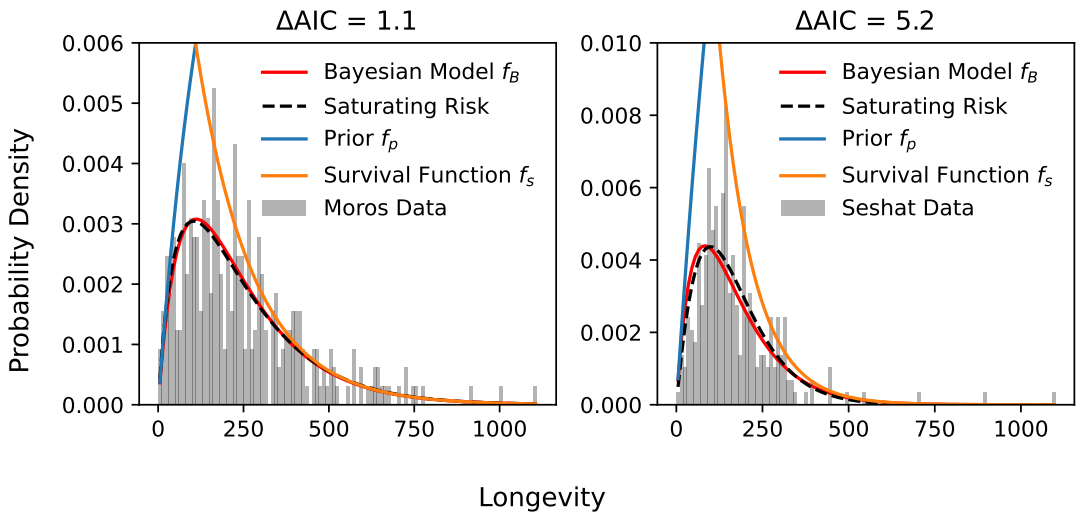

Figure 7.1: Scheffer et al.'s Saturating Risk model is compared to the Bayesian model f_B for the Moros (left) and Seshat (right) data. The ΔAIC indicates how much better Scheffer et al.'s model is compared to the two-stage Bayesian model. The Bayesian prior f_p and the naive survivorship function f_s are also depicted.

according to the Bayesian methodology. Note that because the prior distribution for the discovery rate converges to 1, f_p cannot be normalised, but the posterior distribution f_B can be normalised as a probability density. One could formally solve the problem of $\int_0^\infty f_p(t)dt = \infty$ by replacing the infinite upper limit of the integral with the highest possible age of states τ, e.g. as the time since human settlements and agriculture were first established. Nevertheless, the normalisation constant is an irrelevant scaling factor for the problem presented here.

By using the likelihood optimisation scheme from [257] and implementing it via the function *optimize.minimize* from scipy [234], the Seshat and Moros data is fitted to the Bayesian model in equation (7.3). The results in figure 7.1 show that this model captures the general trend of the data extremely well. The AIC is used to evaluate its goodness of fit as in [257] by comparing the differences in AIC

$$\Delta\text{AIC} = \text{AIC}_B - \text{AIC}_{SR} \tag{7.4}$$

between the Bayesian and the Saturating Risk model. More information about the interpretation of ΔAIC and characteristic threshold values can be found in [67].

For the Moros data, the Bayesian model is essentially as accurate as the Saturating Risk model because the ΔAIC ≈ 1.1 is negligibly small. For the Seshat sample, ΔAIC is slightly larger, but the model is still reasonably accurate as discussed in [67]. Hence, even though both datasets are fitted slightly better by the Saturating Risk model, the Bayesian model is essentially an equally good description within the fluctuation expected by statistical noise. Because both compared models have the same number of parameters, likelihood-ratio tests and ΔAIC evaluation imply the same results. Note that for short and high longevity, $f_p(t)$ and $f_s(t)$ approximate their respective tail of the data.

7.3 Discussion

While this brief study is by no means exhaustive, it indicates that the Bayesian model is at least viable for the Moros data and perhaps, if coupled with a more refined survival function from [257], could also become viable for the Seshat data and even surpass the accuracy of the Saturating Risk model. The exact form of the prior distribution could certainly be varied, too. Considering that a long-lasting state did not only have more time to produce its own set of archeological records, but also to be mentioned in contemporary sources from other geographic areas, it might be reasonable to use a super-exponential discovery rate instead. Fine-tuning the Bayesian estimation might be an interesting endeavour, but is ultimately beyond the scope of this chapter as our intention is to merely illustrate the potential effects of survivorship bias with an example from recent research. This chapter is not supposed to falsify the findings in [257], but rather to highlight issues that need to be considered during the analysis of historical data and offer the first attempt at a solution. Note that the problem of survivorship bias not only appears in historical data like the Sehsat databank, but also in more modern data as discussed in chapter 4 for the survivorship bias in financial correlation matrix analyses which tend to disregard companies that went bankrupt during the time period of interest.

The product in the Bayesian model equation (7.3) might superficially resemble the product of hazard function h and survival function S in [257], but there is an important difference in the methodologies behind these approaches. In [257], both factors of the product stem from the same underlying process and are therefore not independent of each other. Moreover, the analysis assumes that the maximum of the longevity distribution is caused by a decrease in resilience and therefore originates from the underlying dynamics. On the contrary, the Bayesian approach of this chapter highlights potential problems with the data gathering as the source of this feature in the observed distribution.

Of course, Scheffer et al. were aware of the imperfections found in historical datasets such as Seshat and Moros. They described some of these issues in their Supplementary Information to [257] and listed some problematic sources and polities. Moreover, they explicitly stated that their polity selection "does pose a bias against areas with less well documented states", thereby motivating the present study. Hence, it seems natural to posit an alternative model that takes into account the probability of a given polity to actually be considered as a state and to be included in the datasets via a Bayesian prior distribution in order to account for this potential source of bias. Additionally, because historical polities often had less direct control over their territories than modern states and often originated as a subpolity of a larger empire, it is challenging to assign a clear beginning and end to their lifetime, which is reflected by the efforts to construct the CrisisDB database within the Seshat project [258, 259, 260]. "Crafting better databases" is included in the research agenda in [257] to alleviate the issues presented here, but if the accuracy of such datasets is nevertheless limited by the sparsity of archaeological records, Bayesian methods might provide a suitable alternative to incorporate the underlying uncertainty.

8 The Characteristic Time Scale of Cultural Evolution

"History doesn't repeat itself, but it often rhymes."

— attributed to Mark Twain

Abstract Numerous researchers from various disciplines have explored commonalities and divergences in the evolution of complex social formations. Here, we explore whether there is a characteristic time-course for the evolution of social complexity in a handful of different geographic areas. Data from the *Seshat: Global History Databank* is shifted so that the overlapping time series can be fitted to a single logistic regression model for all 23 geographic areas under consideration. The resulting regression shows convincing out-of-sample predictions and its period of extensive growth in social complexity can be identified via bootstrapping as a time interval of roughly 2500 years. To analyse the endogenous growth of social complexity, each time series is restricted to a central time interval without major disruptions in cultural or institutional continuity and both approaches result in a similar logistic regression curve. Our results suggest that these different areas have indeed experienced a similar course in the their evolution of social complexity, but that this is a lengthy process involving both internal developments and external influences.

This chapter and the associated appendix G are based on *Tobias Wand and Daniel Hoyer, PNAS Nexus, Volume 3, Issue 2, pgae009* (2024) [174]. Tobias Wand and Daniel Hoyer conceptualised the project. Tobias Wand wrote the code to analyse the data and carried out the research. Daniel Hoyer curated the data and supervised the project. All figures in this chapter and appendix G are taken from [174] and were created by Tobias Wand.

8.1 Introduction

8.1.1 Motivation to Find Characteristic Time Scales

Researchers from various disciplines have analysed commonalities and divergences in the evolution of complex social systems [48, 50, 253, 254, 261, 262, 263]. The recent emergence of cliodynamics as a discipline has started the analysis of the dynamics of human societies and states with data-driven scrutiny and modelling approaches from natural sciences [47, 264]. Previous work established that a common set of factors associated with complex social formations typically moved in tandem across a wide variety of regions and time-periods; factors such as social scale, the use of informational media, administrative hierarchies, monetary instruments, and others [50]. These were interpreted as comprising the primary dimension of what could be called "social complexity" across cultures, though other dimensions can be adduced as well [253].

Various studies have already discussed or tried to identify the causal drivers of cultural evolution and evaluated the evidence for different theories of *why* cultures become more complex [254, 265, 266, 267, 268]. Beyond the *causal* similarities behind cultural evolution across cultures, researchers have also found evidence for *temporal* similarities and seemingly parallel time scales in the dynamics of various social structures. For example, models for societal collapse have been derived from demographic and fiscal data that show characteristic oscillation periods of a few centuries and a fine structure with a faster periodicity of approximately two human generations [48]. Other theories suggest that cultural evolution leads to the emergence of similar political institutions and schools of thought at roughly identical time intervals across different geographic regions [247, 269, 270]. Another recent study has evaluated the connection between the first emergence of complex societies in different world regions and the age of widespread reliance on agriculture in these areas [271], supporting the theory that agriculture is a necessary condition for the evolution of complex societies. While the time lag between the primary reliance on agriculture and the emergence of states was found to decrease over time, an average time lag of roughly 3,400 years for pristine states suggested the existence of a characteristic time scale, though this was not the explicit focus of that study. Similarly, the study on causal drivers in [254] also found that the time since the adoption of agriculture had a statistically significant effect as a linear predictor variable (called AgriLag) for socio-political complexity, providing additional evidence for temporal regularities in the growth of complexity across different cultures and civilisations. Finally, using the same data as this chapter (cf. section 8.1.2), it was possible to identify

periods of cultural macroevolution with either slow or rapid change in social complexity [263].

Nevertheless, as yet there is no consensus on whether there is a "typical" time scale for socio-political development cross-culturally, let alone what that time-course might be. Such characteristic time scales of dynamic systems are, however, well documented in different areas of the natural sciences such as physics and chemistry [272, 273]. Differentiating between fast and slow time scales in a dynamical system can lead to useful insights and can inform modelling assumptions for data analysis [122, 274]. In particular, Haken's theory of the "enslaving principle" [21], according to which the dynamics of fast-relaxing modes are dominated (enslaved) by the behaviour of slowly relaxing modes in a dynamical system, pioneered the research on how dynamics on different time scales influence each other in the same observed system. The existence of temporal regularities among societal dynamics would suggest that cultural evolution not only occurs in similar developmental stages across geographic regions and time periods, but also in similar time intervals. This would add an important dimension to our understanding of *how* complex social formations evolve, and raise a number of critical questions about what drives these cross-cultural patterns.

Here, we adapt some of the methods employed in the natural sciences in an attempt to identify characteristic time scales in the evolution of complex societies. We utilise data collected by the *Seshat: Global History Databank* [49, 250, 251], a large repository of information about the dynamics of social complexity across world regions from the Neolithic to the early modern period [275]. We find that, despite significant differences in the timing and intensity of major increases in social complexity reached by polities across the Seshat sample, there is a typical, quantitatively identifiable time course recognisable in the data. This result is robust to a variety of checks and covers polities from all major world regions and across thousands of years of history. Our findings offer a novel contribution to the study of cultural evolution, indicating the existence of a general, cross-cultural pattern in both the scale as well as the pace of social complexity development.

8.1.2 Seshat Databank

The Seeshat databank includes systematically coded information on over 35 natrural geographic areas (NGAs) and over 200 variables across up to 10,000 years in time steps of 100 years ([49, 275]. During the time interval captured by the Seshat databank, these NGAs are occupied by over 370 different identifiable polities, defined as an "independent political unit". This sample is constructed by identifying all known polities that occu-

pied part or all of each NGA over time (see [49, 250, 251] for details). The recorded variables are aggregated into nine complexity characteristics (CCs) and a principal component analysis shows that 77% of the variation in the data can be explained by the first principal component (SPC1), which has almost equal contributions from all nine CCs [50]. In the case of missing data or expert disagreement in [50], multiple imputation [276] was used to create several datasets with the differently imputed values which were aggregated into the principle component analysis. The NGAs in the Seshat data cover a wide geographical range and different levels of social complexity, though it is important to note that the Seshat sample is focused largely (but not exclusively) on relatively complex, sedentary societies. Data on the CCs is sampled at century intervals, giving a time series of each polity's estimated social complexity measure throughout its duration.

Seshat data has allowed researchers to quantitatively test hypotheses on cultural evolution such as identifying drivers of social complexity and predictors of change in military technology, for example gauging the effect of moralising religions on cultural evolution or predicting historical grain yields [51, 254, 277, 278]. Further analysis of the Seshat data includes a discussion of ideas from biological evolutionary theory with respect to the *tempo* of cultural macroevolution, defined as "rates of change, including their acceleration and deceleration", concluding that "cultural macroevolution is characterized by periods of apparent stasis interspersed by rapid change" [263]. These results strongly relate to the question of the present chapter, whether there is some generality in the time scale of cultural evolution in the Seshat data.

8.1.3 Data on Culture/Polity Boundaries and Duration

Each NGA's time series can contain data about very different polities that succeeded each other. Sometimes, a gradual and continuous change between the polities justifies treating predecessor and successor polities as closely related; for instance, in the Latium NGA (modern-day central Italy), Seshat records three separate polities for the Roman Republic, indicating the Early, Middle, and Late phases. These phases are culturally and (to a signfiicant degree) institutionally continuous and can therefore be treated as a single polity-sequence. In other cases, there may have been an invasion or mass migration as a clear breakpoint between the two polity's continuity; for instance, between the Ptolemaic Kingdom and Roman Principate polities in the Upper Egypt NGA. Data from [275] and other information recorded in the Seshat sample, notably information on the relationship between polities, is here used to establish a list of continuous polities. The continuity

is evaluated either as cultural continuity or as political-institutional continuity and our cutout data for both approaches is published on [279].

8.1.4 Organisation of this Chapter

The mathematical methods and technical details of the logistic regression analysis are discussed in section 8.2. Section 8.3 explains how we transformed the time series data on each NGA in the Seshat sample to establish a common reference point to investigate the time course of changes in social complexity across NGAs. In short, we shift each NGA's time series with respect to a single anchor time such that the transformed time variable *RelTime* shows major overlap between the RelTime-vs-SPC1-curves of all NGAs. Exploratory data analysis for the whole dataset reveals that there is a logistic relationship between RelTime and the SPC1 response variable. Section 8.4 identifies the time scale of growth from the lower to the upper plateau of the logistic curve via bootstrapping. The logistic curve is compared to a regression using only either the culturally or institutionally continuous time series and moreover, the duration of the continuous time series is compared to the estimated characteristic time scale. Finally, the results of the analyses are summarised and discussed in section 8.5. Appendix G gives more details on the data used for this analysis.

8.2 Methods and Technical Details

8.2.1 Logistic Regression Curve

Logistic regression is used to model time series data which is mostly distributed at two plateaus with a transitory area between them [280]. It is based on the characteristic sigmoid curve of the logistic growth model described in [281], which models an exponential growth process constrained by a carrying capacity. The logistic curve has the functional form f with an asymptotic behaviour

$$f(x) = \frac{a}{1+\exp(-c(x-d))} + b, \quad f(-\infty) = b \quad \text{and} \quad f(\infty) = a + b. \tag{8.1}$$

Often, data is scaled such that $b = 0$ and $a = 1$, i.e. an asymptotic behaviour between two binary plateaus at height 0 and 1. To analyse the accuracy of the logistic regression curve, the methods from chapter 2.2.2 are used, namely the KDE, the *RMSE*, the ρ^2 coefficient and bootstrapping. For the KDE, the Gaussian density is used as the kernel via *scipy.stats.gaussian_kde* [234].

8.2.2 Direction Reversal

Estimating the coefficients (a, b, c, d) can lead to numerical instabilities because it is possible to transform a logistic curve with $c > 0$ to an equivalent equation $\hat{f}$ with $c < 0$. For example, consider $a = 1, b = 0, c = 1$ and $d = 0$, then

$$f(x) = \frac{1}{1 + \exp(-x)} = \frac{\exp(x)}{\exp(x) + 1} = \frac{\exp(x) + 1 - 1}{\exp(x) + 1} = 1 + \frac{-1}{1 + \exp(x)}. \tag{8.2}$$

The last reformulation of f can now be parametrised via $\hat{a} = -1, \hat{b} = 1, \hat{c} = -1$ and $\hat{d} = 0$. This ambiguity can lead to the regression algorithm yielding positive and negative results for c during multiple runs. It can be prevented by setting an initial parameter guess with $c > 0$, which locks the algorithm into positive values for c.

8.3 Data Transformation and Exploratory Analyses

First, all raw SPC1 time series are rescaled via a min-max scaling, i.e.

$$SPC1 = \frac{SPC1_{raw} - \min(SPC1_{raw})}{\max(SPC1_{raw}) - \min(SPC1_{raw})}. \tag{8.3}$$

This has the advantage of making the interpretation of high and low SPC1 values much easier as high/low correspond to close to 1 or close to 0, respectively. It also makes the parametrisation of a logistic curve easier by restricting the observed data to a range between 0 and 1.

8.3.1 Anchor Time

Considering that most NGAs have an SPC1 time series that starts at a low value barely above 0 and ends at a high value close to 1, a logistic regression model seems like a reasonable suggestion for the data. Although all NGAs experience a growth in SPC1 over time, they start at very different calendar years. Therefore, it is necessary to shift the time series via an anchor time so that in the new "relative" time, the growth phase in each NGA's time series coincide. Then, one logistic regression from equation (8.1) can be used for all shifted time series. Hence, each NGA i needs an anchor time $T_a^{(i)}$ so that if all time series are shifted by $-T_a^{(i)}$, they roughly overlap. The shifted time series of the NGAs and the logistic fit are shown in the main part of figure 8.1.

The anchor time can be chosen as the year during which the NGA i's SPC1 value crosses a threshold value. It has already been shown that there is a clear threshold SPC1_0 between high and low values of SPC1 in the data, which was used to define the

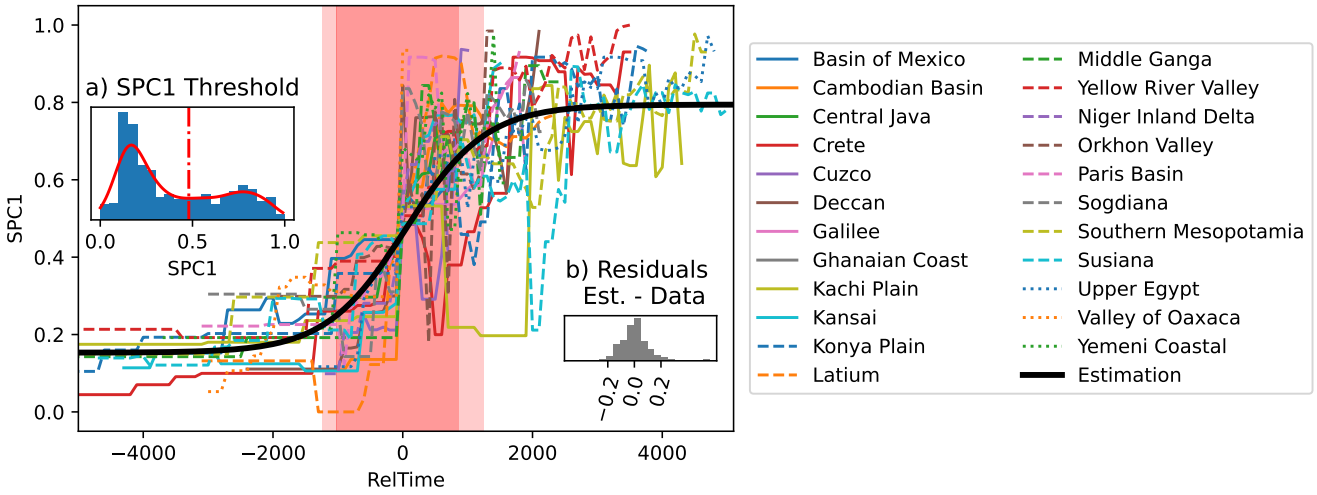

Figure 8.1: Main figure: Time series of RelTime vs. SPC1 for all 23 NGAs that cross $SPC1_0$ and the logistic regression. Marked in red is the area of growth between the two plateaus of the curve as identified in section 8.4. The various time series are shown in appendix G.2 in multiple plots to make the identification easier. Insets: a) distribution of SPC1 for all 35 NGAs, the associated KDE (red) and the threshold $SPC1_0$ (vertical); b) residuals of the logistic regression.

RelTime variable in [277]. A similar methodology was also used in [52], but there, the authors used the emergence of a moralising religious belief as the "year zero" to shift each NGA's time series. Copying the procedure from [277] to get the RelTime variable, $SPC1_0$ is chosen as the minimum between the two maxima in the kernel density estimation (KDE; explained in 2.2.2) of the SPC1 values (figure 8.1, inset a). The anchor time $T_a^{(i)}$ is then selected as the first recorded data point when the NGA i exceeds $SPC1_0$. An illustration of the anchor time shift is provided in the appendix G.1. Thus, the 12 NGAs that never exceed $SPC1_0$ are discarded from this analysis. On the one hand, this is not too problematic because their limited growth in SPC1 means that they would have only contributed little information to the estimation of SPC1's characteristic growth time, but on the other hand, chapter 7 has highlighted the problems that can arise from this. Moreover, this discards all NGAs from the world region Oceania-Australia in the Seshat sample, meaning that it might introduce a geographical bias. We discuss this and other possible limitations of the approach further in section 8.5 below.

8.3.2 Logistic Regression

The RelTime-vs-SPC1 data is fitted to a logistic regression curve like equation (8.1) via the optimisation algorithm *scipy.optimize.curve_fit* from [234]. The quality of the

regression curve is evaluated with the methods from 2.2.2. With the exception of a few outlier observations occurring in several of the NGAs, all time series qualitatively agree with the regression curve fairly well. Also, the majority of the residuals shown in figure 8.1 (inset b) are distributed roughly symmetrically in a neighbourhood of zero. The distribution of the residuals and the rather low value of the root-mean-square-error $RMSE \approx 0.11$ both indicate that the logistic regression is a suitable model for the shifted SPC1 data.

To further increase our trust in the quality of the regression, it is evaluated via the coefficient of determination ρ^2 in an out-of-sample prediction. The data is split randomly into equally sized training and testing data and a logistic regression curve f_i is estimated by only using the training data. Then, f_i is used to predict the values for the test data and the prediction is evaluated via the ρ^2 metric in 2.2.2. The random training-test-split is repeated $i = 1, \ldots, 100$ times, each time using the estimated parameters from the full time series as initial values, and the resulting ρ^2 values have an average of $\rho^2 = 0.81 \pm 0.01$ far above 0 and therefore further strengthens our trust in the logistic model.

8.4 Analysis of Time Scales

8.4.1 Finding a Characteristic Time Scale

Having established that the data can be accurately captured by a logistic curve, we can investigate our research question; namely, how many years did it typically take in these different regions to transition from a polity with low SPC1 to one with high SPC1? Or to reformulate the question: when does the curve leave the low plateau and when does it reach the high plateau? We attempt to answer these questions by estimating the heights of the plateaus and their respective uncertainties and by checking when the regression curve crosses these thresholds.

1000 steps of bootstrapping are performed by sampling from the list of NGAs and by estimating the regression parameters $(a_i, b_i, c_i, d_i)_{i=1,\ldots,1000}$ for each sample (see 2.2.2). According to the asymptotic behaviour in 8.2.1, the plateaus are given by b_i and $a_i + b_i$. In order to make conservative estimates instead of being influenced by noise, an upper boundary for the lower plateau's value Th_1 and a lower boundary for the upper plateau's value Th_2 are used as the thresholds. Th_1 is chosen as $Th_1 = \mu(b) + 3\sigma(b)$ of the bootstrapped distribution of b, Th_2 as $Th_2 = \mu(a+b) - 3\sigma(a+b)$. For each bootstrapped logistic curve $f_i(t)$, it is then determined at which RelTime values $t_1^{(i)}$ and $t_2^{(i)}$ it crosses

the lower and upper thresholds Th_1 and Th_2. We can then understand the mean value

$$\mu\left(t_2^{(i)} - t_1^{(i)}\right) = \mu\left(t_2^{(i)}\right) - \mu\left(t_1^{(i)}\right) \approx 2500 \text{ years} \tag{8.4}$$

as the characteristic time scale for the period of rapid cultural evolution between low and high plateaus of socio-political complexity, across geography and not in reference to any specific time period. Note that one can also choose less restrictive thresholds via $Th_1 = \mu(b) + \sigma(b)$ and $Th_2 = \mu(a+b) - \sigma(a+b)$. With these thresholds, the regression curve leaves the vicinity of the lower plateau rather quickly but needs much longer until it is close enough to the upper plateau to be considered as having reached the upper plateau. These 1σ thresholds would result in a longer time scale of roughly

$$\mu\left(t_2^{(i)} - t_1^{(i)}\middle|1\sigma\right) \approx 4000 \text{ years.} \tag{8.5}$$

We can check the general validity of these results by explicitly identifying for each NGA i the first time $\tau_1^{(i)}$ that their SPC1 value exceeds Th_1 and the first time $\tau_2^{(i)}$ they exceed Th_2. With the exception of the Ghanaian Coast, all NGAs cross Th_2 and therefore, this procedure yields 22 time durations $d^{(i)} = \tau_2^{(i)} - \tau_1^{(i)}$. Only for the two NGAs Kachi Plain and Middle Yellow River Valley (two 'pristine' states, cf. the discussion in section 8.5.2) does the duration $d^{(i)}$ exceed the 4000 years estimated as an upper boundary in (8.5). Both the mean (approximately 2200 years) and median (2100 years) are in line with the main estimation in (8.4).

8.4.2 Continuous Polities

There are two reasons why it makes sense to restrict the logistic regression only to a central part of each NGA's time series, during which the polities in that NGA are not disrupted by external influence or major dislocations in socio-political structures. First, the logistic regression starts at a plateau of low values of SPC1 close to 0 and ends at a plateau of high values close to 1. Therefore, even a bad interpolation for the central part can achieve a good *RMSE*, if the plateaus of the high and low tails are sufficiently accurate. However, this would not be a reliable estimation to make an inference on the growth phase in the centre of the curve. Second, if the NGA's polity is annexed by another, more developed polity, then it inherits the invading polity's high SPC1 value and may make a sudden jump in the SPC1 curve. However, the logistic regression here is intended to model steady, uninterrupted growth like in [281] and not major transitions driven by developments experienced elsewhere, as through annexations by an external invader. Therefore, it makes sense to divide each NGA's time series into intervals which

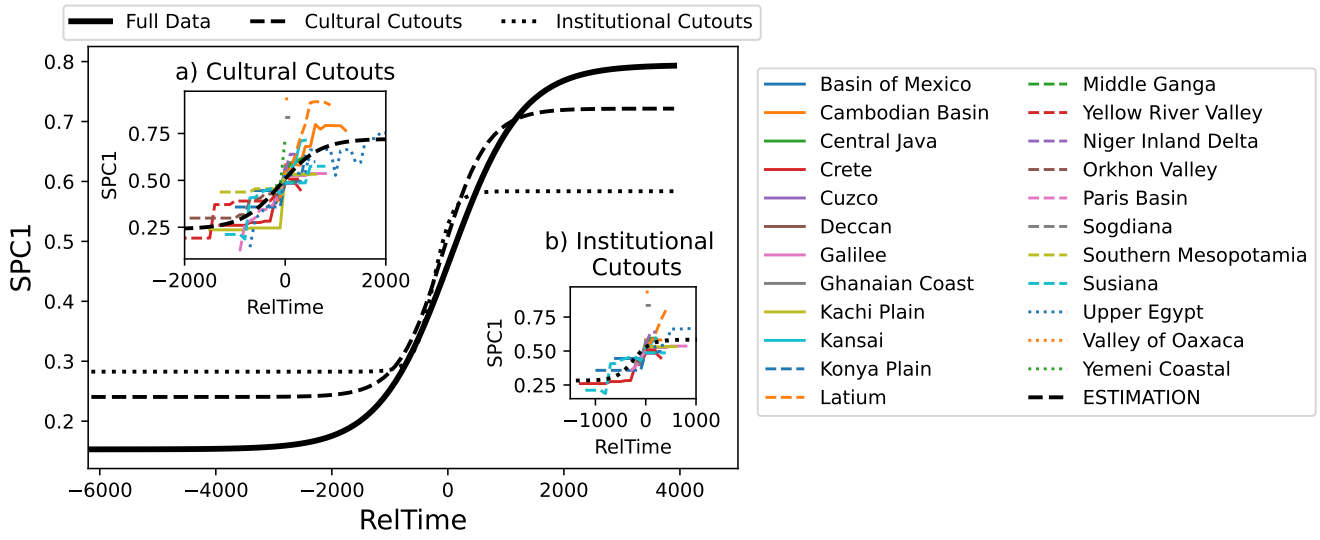

Figure 8.2: Main figure: estimated logistic curves for the full data and the two cutout methods. Insets: a) each NGA's central time series and resulting logistic curve for the culturally continuous time series; b) the same as subfigure a) for the institutionally continuous time series.

are separated by sharp, discontinuous changes within each NGA and to restrict the analysis of the NGA to its central interval, i.e. to the time series from the polities that cross the $SPC1_0$ threshold.

As mentioned earlier, there are two ways of identifying such discontinuous changes: either via cultural changes of via major institutional changes of the polity's governance. Both approaches are analysed separately. The central time series for both methods and their resulting logistic regressions are shown in figure 8.2.

Cultural Continuity

One set of sequences was determined by the absence of a major cultural dislocation; namely, the introduction of a new ideological and linguistic system, major population displacement or major technological advance (e.g. the adoption of iron metallurgy). This is a very broad and lenient definition of continuity, as it allows for very different social formations to be part of a single sequence and can include significant developments. In Egypt, for instance, we treat nearly the entire Pharaonic period (from the Naqada period to the Achaemenid conquest) of over 3000 years, including the so-called Intermediate periods when central rule was fragmented (though many cultural and social features were retained), as a culturally continuous time period.

For the 23 NGAs under consideration, the mean value of data points for the culturally

continuous central interval is approximately 11.7, i.e. there is on average a bit more than one millennium of data. While this is much shorter than the characteristic time scale of roughly 2500 years, the longest continuous time series of the NGAs show a similar length to that of the characteristic time scale (cf. the left half of table 8.1). Hence, the logistic regression for these cutouts is rather close to the regression of the full data (cf. main part of figure 8.2) and in particular, the regression curves' steepness (i.e. their time of growth) is quite similar.

Institutional Continuity

For institutionally continuous time periods, we follow a similar procedure as above, though with different criteria for continuity leading to shorter sequences. Namely, we break each sequence at any significant political/institutional change, even if there was much continuity in socio-cultural forms. In Egypt, for instance, the institutional sequence starts at the 1st Dynasty period and ends a the end of the Old Kingdom period and the First Intermediate Period, which we call the "Period of the Regions". The mean value of data points for institutionally continuous central time series is only 5.9 and represents approximately 500 years of data. Even the longest continuous sequences now do not last as long as the characteristic growth time of 2500 years (cf. right half of table 8.1). Moreover, the logistic regression has only very little data for the parameter estimation (cf. inset *b* of figure 8.2) and hence, the logistic regression has a much lower SPC1 level for the upper plateau than the regression to the full data (main part of figure 8.2) because the cutout time series are too short to reach the high-SPC1 levels.

NGA	Cultural Continuity Length	NGA	Institutional Continuity Length
Yellow River	38	Susiana	17
Upper Egypt	33	Crete	17
Kachi Plain	22	Konya Plain	15
Susiana	21	Upper Egypt	10

Table 8.1: For both methods of identifying continuous time sequences, the four longest continuous central time series are shown and the amount of data points they contain (given as their length). The data points are sampled at intervals of one century. The culturally continuous time series are much longer than the longest institutionally continuous sequences.

8.5 Summary and Discussion

8.5.1 Summary

Exploratory data analysis shows in figure 8.1 that the logistic regression is a suitable model for the RelTime-vs-SPC1 time series. Bootstrapping allows us to narrow down the time interval of rapid SPC1 growth to approximately 2500 years, as highlighted in figure 8.1. Together, these results illustrate that there is a uniform behaviour in growth of social complexity represented by the time evolution of SPC1.
If the data is restricted to the central part of each NGA's time series without any discontinuous cultural or institutional transitions, the logistic regression is still a reasonable model and shows a similar shape to the full data as depicted in figure 8.2. In particular, the regression based on the culturally continuous time series show a very similar steepness (i.e. growth period) to the full regression curve.

8.5.2 Discussion of the Time Scale and Continuous Sequences

Figure 8.2 shows that the culturally continuous and institutionally continuous time series result in a similar logistic regression to the full data. Notably, table 8.1 shows that the culturally continuous time series have a much longer duration than the institutionally continuous ones. In particular, in the Yellow River Valley, Upper Egypt, Kachi Plain and Susiana, the culturally continuous time series is approximately as long (or even longer) as the characteristic time scale of SPC1 growth. This is expected for regions that saw the emergence of large, complex states relatively early in history and without any precedent from neighbouring societies – the so-called *pristine* or *primary* states [282, 283] – which these regions all experienced. However, this is not the case for most other NGAs, indicating that in these NGAs, the growth from the lower to the higher SPC1 plateau did not take place over the course of just one culturally continuous era, but rather included developments across cultural spheres and, in most cases, including developments brought in from the outside in the form of direct conquest or more indirect influence. The institutionally continuous time series are all significantly shorter than the characteristic growth time, as expected from the criteria used to generate these sequences. This is notable, as it suggests that in order to transition from low to high social complexity, major shifts in the NGA's governing institutions are necessary to facilitate the increase in social complexity. In other words, our findings suggest that major transitions in social complexity are not feasible for a single polity to accomplish, but require multiple social formations to build successively (but not monotonically, as the above figures illustrate) on prior developments. Nevertheless, the general similarity of the three regression curves

in figure 8.2 shows that our analysis is stable with respect to the exact selection of time periods and different cutout criteria used.

It is interesting to compare the NGAs that crossed the threshold $SPC1_0$ to those that failed to do so and stayed at lower complexity values. The latter group had a mean of only 6.4 recorded data points, i.e. there were only complex social formations coded as part of the Seshat sample for a period of roughly six centuries. On the other hand, the NGAs that did reach a high complexity and exceeded the threshold $SPC1_0$ had a mean of 57.3 recorded data points, corresponding to almost 6 millennia of observed data. Partly this is explained by different availability of historical and archaeological evidence in different regions, but it suggests also that cultural developments in the low complexity NGAs could have followed the same trajectory of logistic growth, if they had been given enough time. Unfortunately, the necessity to identify an anchor time for this analysis means that all NGAs from the Seshat world region Oceania-Australia had to be discarded for this research. The bias introduced by this has to be kept in mind while interpreting our results.

8.5.3 Interpretation and Comparison to Previous Work

With the shifted time index RelTime, the logistic regression model of the SPC1 time series achieves a high accuracy in capturing the evolution of socio-political complexity measured by SPC1. Previous work has already demonstrated a significant amount of cross-cultural generality in the factors contributing to the evolution of socio-political complexity ([50], supplemented by findings in [253, 284]). Notably, a previous study has already identified a characteristic growth pattern of SPC1 and the second principal component SPC2 and found that a rapid period of scale is first followed by a growth of information processing and economic complexity and then by further growth in scale [253].

Here, we expand on this prior work by identifying that the time scales involved in these developments also exhibit a general, characteristic shape. Nevertheless, the evolution of social complexity is a lengthy and non-monotonous process; this emerges clearly from our analyses distinguishing the full regional time-series involved in the transition from low to high thresholds of SPC1 from sequences of cultural or institutional continuity. We see no examples of this evolution accomplished during a single institutionally-continuous sequence. Further, in all NGAs there are noisy periods during which SPC1 grows but also crises during which socio-political complexity sharply declines, only to recover later and continue increasing. These findings highlight both that different parts of the world experienced similar processes of social complexity growth, involving multiple phases of

cultural and socio-political structures building on (and occasionally recovering from) prior developments in each region.

While the sample of past societies explored in this chapter is certainly not exhaustive, they comprise a fairly representative sample of regions from different parts of the world and include societies from different periods, cultures and different developmental experiences. These results thus lend novel empirical support to the idea (e.g. from [247, 263]) that socio-cultural evolution does indeed occur in similar time scales across different cultures and geographies. Future research can expand these insights by including additional societies and exploring alternate thresholds of complexity to identify anchor times to include more NGAs from the original sample, because the current thresholding procedure in particular excluded some NGAs from modern-day Oceania-Australia from our analysis.

In terms of the underlying approach, our study tries to single out the autocatalytic effect of social complexity growth. To this end, we have focused on only one NGA at a time and compared our regression results to the culturally and institutionally continuous periods for the respective NGA. Thus, we uncover an empirical pattern in the temporal evolution of SPC1 that has not yet been fully discussed by previous work, e.g in causal analyses of the drivers of social complexity like in [254]. Our methodology differs from the regression model in [254] by deliberately choosing a very simple model to single out the temporal evolution whilst disregarding possible drivers of the observed dynamics. We believe this approach can be utilised to answer other questions about long-run cultural evolution, for instance if the processes by which key technologies (e.g. metallurgy, military technology, communications media etc.) are invented in certain locations and then adopted in others.

While the autocatalytic growth model provides an elegant interpretation of our findings (the current level of complexity facilitates further growth until the presence of an upper boundary of complexity is approached), it has to be regarded with caution: We sought as far as possible to disentangle culturally and institutionally endogenous developments from those driven by interactions with other polities, though even the internal developments are not free from external influence. Previous work, for instance, shows the strong effect of military conflicts with other states on the growth of socio-political complexity [51, 254]. Hence, the autocatalytic model might be a useful low-dimensional description of the data, but not an exhaustive explanation. In short, our findings exposes a cross-cultural temporal pattern whose causes need to be fleshed out in future work.

Finally, the findings of the present chapter can be used as a benchmark for future additions to the Seshat data: if a new NGA is added to the databank and shows a clear divergence from the logistic curve, it may be prudent to either check if there are any

mistakes in the data generation and interpolation or if the divergences can be explained by historical developments. Such a benchmark may thus be useful for further expansion of the Seshat databank.

Part V

Conclusion

This chapter first summarises the main results presented in this dissertation and suggests possible future research directions. Second, the challenges in getting access to higher-quality data are explained along with research opportunities that better databases might offer. Finally, more general conclusions that go beyond the immediate scope of this thesis are discussed.

Chapter 3 has used high frequency and daily financial price time series to estimate polynomial stochastic differential equations. A second order polynomial (i.e. a geometric Brownian motion with an additional quadratic term) has been found to be the most reliable description of the data and shows the existence of stable fixed points which help to qualitatively describe the time series dynamics and provide a different perspective than the double well potential in [104]. The great success of the geometric Brownian motion in applied finance is due to its simplicity allowing for closed-form equations, e.g. in the Black Scholes model [16]. Whether the second order extension proposed in this thesis might also lead to closed expressions like this could be a valuable endeavour for future research.
Chapter 4 has analysed the correlation of daily return time series to understand collective effects of the financial market and expands on the work in [36, 116, 117]. Analysis via explainable artificial intelligence of the market states revealed by k-means clustering of the correlation matrices has shown that the IT sector (probably due to the dot-com bubble) and the energy sector have a heightened importance for the market states. The mean correlation has been modelled with a generalised Langevin equation (GLE) and has been found to exhibit memory effects on the order of several weeks. In both cases, a survivorship bias might be present in the data because only companies with return time series for the full period have been considered for the analysis. Incorporating knowledge from time series that terminate during the period (e.g. because of a company's bankruptcy) might be possible via a smoothened transition between the mean correlation and sector returns for n and $n-1$ time series and could further strengthen these results.
Going beyond correlations, chapter 5 has analysed the Granger causality between daily returns of sector portfolios. The hierarchy of this network of causality has been extracted via the Helmholtz-Hodge-Kodaira decomposition for bidirectional networks (HHKD). For the year 2020, this analysis has revealed the dominant contribution of the hierarchical over the rotational component in the causality network and has shown the heightened importance of the pharmaceutical sector (probably due to the Covid crisis) and the precious metal sector as causal drivers. This is in contrast to the findings for the correlation matrix analysis in chapter 4 which underlines the importance of a clear distinction

between correlation and causality. Further expansions of this work might include more complicated causality measures (e.g. transfer entropy or convergent cross mapping) to capture nonlinear causal effects as well as the inclusion of other time series (such as gold prices, currency exchange rates or unemployment rates) as background information for the conditional Granger causality.
The agent-based model in chapter 6 has simulated a cooperation scheme between agents and their nearest neighbours via insurances against an unknown multiplicative effect on their wealth based on the model in [33]. Growth rate optimisation has shown that cooperation between agents is favourable for both participants. Over time, regions emerge on the simulated lattice which are either occupied by poor or rich agents. Hence, the immobile agents form clusters which superficially resemble kin selection but actually emerge purely from the cooperation scheme.
Chapter 7 has expanded on a recent resilience analysis of the longevity of historical countries in [257]. It has been shown that the conclusion from [257] that a state's resilience decreases after 200 years might have been mislead by a survivorship bias in the data. If Bayesian statistics are used to account for such bias, a constant resilience function can explain the observed distribution with almost equally high accuracy. This short study has underlined a general problem of historical data analysis in the field of cliodynamics as data records tend to be sparse and unreliable during the earliest periods of pre-modern states.
Finally, chapter 8 has compared the development of socio-economic complexity across different regions in the Seshat databank quantified by the first principle component SPC1 from [50]. The SPC1 time series has been shifted according to an anchor time so that geographic areas in which cultural complexity started to evolve at vastly different times have overlapping SPC1 curves which has been described with a logistic growth curve. Out-of-sample predictions have shown the validity of this regression curve whose characteristic growth time can be estimated as approximately 2500 years.
Overall, this thesis has shown that complex socio-economic systems can be analysed with methods from physics at all time scales of human interactions. The interplay between statistical analysis, modelling approaches and large datasets for social and economic processes has achieved major contribution to the understanding of complex social systems and will continue to do so with the increasing availability of such data.

All of the presented data-driven projects could, of course, have benefited from access to data with higher quality. For the financial market data, this is in principle possible as high frequency data down to the level of individual trading decisions (ticks) is avail-

able. The GLE estimation in chapter 4 has exhibited instabilities because of the short time series length after data pre-processing and could have been improved with access to longer time series. Also, this may have enabled more elaborate causality methods than Granger causality to be used for the HHKD network analysis in chapter 5. However, such data is not available free of charge, but is unfortunately associated with high access costs. One has to wonder how many phenomena have already been understood by researchers at financial companies whose results are kept secret as the company's intellectual property. Additionally, even though the US stock market data is frequently used for research, one should also test the universality of these findings on data from other markets such as the EURO STOXX 50 or the DAX. For the cliodynamics research projects, having higher quality data is a huge challenge, but efforts are underway to construct more refined and fine-grained datasets through the CrisisDB project [259]. However, due to the multidisciplinary nature of this research, even improvements for a single geographic area require the work of domain experts from several fields (archaeology, history, anthropology) for various historical eras. Hence, while more reliable historical data would be helpful for analyses such as the regression in chapter 8, developing tools such as the Bayesian prior in chapter 7 to account for potential errors and bias in the data might be the more realistic route for progress.

A more fundamental issue with respect to the quantitative study of socio-economic systems should not be omitted from this discussion: Suppose researchers manage to uncover equations, rules and natural laws which describe how human economies and societies develop. If this allows us to predict the future of socio-economic systems – what does that mean for the individual's ability to influence the course of history? What is the role of the individual in society, if our collective future is set in stone and – no matter how hard we try to change it – our actions only lead to one predetermined outcome, predicted by mathematics and statistics? There are two arguments against this quite concerning and worrisome argument. First, human behaviour adapts to changing circumstances and hence, whatever rules can be found in the data, they might only hold true for a finite duration of time. Second, and more importantly, such socio-economic laws always concern the macroscopic behaviour of the system. Synegetics has shown us that the microscopic constituents have some degrees of freedom even despite the influence of the macroscopic behaviour on their dynamics. In socio-economic systems, the individual human therefore retains a certain scope of action even against the overarching developments of the macroscopic level. Uncovering the governing equations of macroscopic social and economic development can provide us with guardrails which help us to see the directions where we

can actually achieve a meaningful change if we put in the effort to do so. They help us to channel our efforts and resources so that they are not wasted on trying to achieve the impossible. Instead, we will be able to see exactly which degrees of freedom are truly available to us and make an informed decision on how to use them to, perhaps not fulfil, but come close to our ideal goals by navigating with and within the flow of time and not trying to swim against it. Fate leads the willing and drags the unwilling along with it.

"Ducunt volentem fata, nolentem trahunt" — Seneca, Epistulae morales

Acknowledgements

First and foremost, I would like to thank my wife Lisa for her emotional support over the past three and a half year as well as my parents Andrea and Theo, my sister Alexandra, my grandparents Karin, Winfried, Anni and Theo and my uncle Christoph for accompanying me since the beginning of my education and supporting me to pursue my interests. I would further like to thank my supervisors Oliver Kamps, Uwe Thiele and Svetlana Gurevich and the research group for their academic support and guidance and my former teachers and professors Wolfgang Mensing, Karol Kovarik, Matthias Löwe, Valentin Popov, Michael Rohlfing, Rodolfo Smiljanic and John Joseph Valetta for fostering my interests. I would further like to thank my research collaborators Jan Harren, Martin Hessler, Daniel Hoyer, Hiroshi Iyetomi, Benjamin Skjold, Timo Wiedemann and, again, Oliver Kamps for their support, Pia Petrak for fresh air, my extended family and the Alemannia for emotional support and Henrik Bette, Colm Connaughton, Yoshi Fujiwara, Lennart Gevers, Thomas Guhr, Anton Heckens, Andreas Heuer, Philip Hövel, Mateusz Iskrzyński, Kiyoshi Kanazawa, Tim Kroll, Oliver Mai, Helena Monke, Alexander Midlen, Yu Ohki, Ole Peters, Christian Philipp, Benjamin Risse, Martin Salinga, Yuki Sato, Katrin Schmietendorf, Wataru Souma, Peter Turchin, Michael te Vrugt, Christoph Wand, Clemens Willers, Jonathan Wolf and Fabian Zelesinski for valuable discussions. I would like to thank the Japan Society for the Promotion of Science, the G-Research grant programme, the Wilhelm und Else Heraeus-Stiftung and especially the Studienstiftung des deutschen Volkes for financially supporting my research. Finally, I would like to thank Franz von Fürstenberg for founding and Wilhelm II for re-establishing the university in Münster as without it, none of this work would have been possible.

Part VI

Appendix

A List of Publications

The content of this thesis is mainly taken from peer-reviewed research publications which were written during the author's time as a PhD student by him as the first author. Content from the following articles was used in this thesis:

- **Tobias Wand**, Martin Heßler and Oliver Kamps, *Identifying dominant industrial sectors in market states of the S&P 500 financial data*, Journal of Statistical Mechanics: Theory and Experiment 043402 (2023) [123]
- **Tobias Wand**[1], Martin Heßler[1] and Oliver Kamps, *Memory Effects, Multiple Time Scales and Local Stability in Langevin Models of the S&P500 Market Correlation*, Entropy 2023, 25(9), 1257 (2023) [124]
- **Tobias Wand** and Daniel Hoyer, *The characteristic time scale of cultural evolution*, PNAS Nexus, Volume 3, Issue 2, pgae009 (2024) [174]
- **Tobias Wand**, Timo Wiedemann, Jan Harren, and Oliver Kamps, *Estimating stable fixed points and Langevin potentials for financial dynamics*, Physical Review E 109, 024226 (2024) [95]
- **Tobias Wand**, Oliver Kamps and Benjamin Skjold, *Cooperation in a non-ergodic world on a network - insurance and beyond*, Chaos 34, 073137 (2024) [213]
- **Tobias Wand**, *A Bayesian Approach to Survivorship Bias in Historical Data Analysis*, Cliodynamics 15 (2024) [256]
- **Tobias Wand**, Oliver Kamps and Hiroshi Iyetomi, *Causal Hierarchy in the Financial Market Network – Uncovered by the Helmholtz-Hodge-Kodaira Decomposition*, Entropy 2024, 26(10), 858 (2024) [175]
- **Tobias Wand**, *Introduction to Artificial Intelligence*, in Michael te Vrugt (Ed.), *Artificial Intelligence and Intelligent Matter*, Springer, Cham, forthcoming in 2025 [56]

[1] Joint first authorship between Tobias Wand and Martin Heßler; both authors agreed on splitting the content of this article for their PhD theses so that only each person's original contribution is contained in each thesis.

Additionally, the following peer-reviewed publications were written during the author's time as a PhD student, but were not incorporated into this thesis:

- **Tobias Wand**, *Analysis of the Football Transfer Market Network*, Journal of Statistical Physics 187, 23 (2022) [285]
- **Tobias Wand**, *Entropy scaling of high resolution systems — Example from football*, Physica A 600, 127536 (2022) [286]
- Martin Heßler, **Tobias Wand** and Oliver Kamps, *Efficient Multi-Change Point Analysis to Decode Economic Crisis Information from the S&P500 Mean Market Correlation*, Entropy 2023, 25(9) (2023) [200]

B List of Abbreviations

The sector abbreviations used in chapters 4 and 5 are found in tables 4.1 and 5.1.

ACF	autocorrelation function
AI	artificial intelligence
AIC	Akaike information criterion
BIC	Bayesian information criterion
BM	Brownian motion
CC	complexity characteristic
CGC	conditional Granger causality
CI	confidence interval or credible interval (both are in use)
DAG	directed acyclic graphs
$\mathbb{E}$	expectation value of a random variable
EE	ergodicity economics
EMH	efficient market hypothesis
GBM	geometric Brownian motion
GICS	global industry classification standard
GLE	generalised Langevin equation
HHKD	Helmholtz-Hodge-Kodaira decomposition
KDE	Kernel Density Estimation
LRP	layer-wise relevance propagation
NGA	natural geographic area
MAP	maximum a-posteriori estimation
MLE	maximum likelihood estimation
ML	machine learning
NN	(artificial neural network
ODE	ordinary differential equation
PCA	Principle Component Analysis
RCGCI	Restricted Conditional Granger Causality Index
RelTime	relative time
RMSE	root mean sqared error

S&P500	Standard & Poor's 500
SDE	stochastic differential equation
SPC	principle components of the Seshat databank
$\mathbb{V}$	variance of a random variable
XAI	explainable artificial intelligence

C Software and Data Availability

Several chapters of this thesis use Python code to analyse datasets. This section lists the available software or data and provides references to them.
Code for the estimation of the Langevin equation in chapter 3 with an example for synthetic data is given in [287]. The correlation time series analysed in chapter 4 is available at [142]. Code and data used for the Helmholtz-Hodge-Kodaira decomposition in chapter 5 is available at [195]. The simulation code for the insurance model in chapter 6 is available at [233]. The data used for the resilience analysis in chapter 7 is taken from the supplementary material of [257]. The pre-processed data used for the logistic regression in chapter 8 is available at [279].

D Details on Explainable AI

D.1 Details on Layer-Wise Relevance Propagation for K-Means

This appendix gives a more detailed explanation of how to reformulate the k-means clustering algorithm and how to use the layer-wise relevance propagation as described in [143].

D.1.1 Neuralisation of K-Means

It has been proven in the supplementary materials of [143] that the k-means decision criterion (4.4) of cluster j vs. the other clusters $l \neq j$ can be rewritten in two layers and a third decision layer. This procedure is described by algorithm 1. Note that while the k-means classifier tests all clusters against each other, the neuralised classifier only tests if an instance belongs to a specific cluster j versus all other clusters $l \neq j$. If the instance does not belong to the j^{th} cluster, algorithm 1 does not say which of the other clusters is correct. Therefore, one needs as many neuralised classifiers as there are clusters to fully neuralise equation (4.4).

Algorithm 1 Neuralisation of a K-Means Classifier

Goal: Test assignment of $\mathbf{C}$ to cluster j against all alternative clusters $l \neq j$.
Input: Instance $\mathbf{C}$ and cluster centroids $(\mathbf{C}_l)_l$
Procedure in three layers:
1: Calculate $\forall l \neq j$: $h_l = \mathbf{w}_l \cdot \mathbf{C} + b_l$ with $\mathbf{w}_l = 2(\mathbf{C}_j - \mathbf{C}_l)$ and $b_l = ||\mathbf{C}_l||^2 - ||\mathbf{C}_j||^2$
2: Calculate $f_j = \min_{l \neq j} h_l$
3: assign $\mathbf{C}$ to cluster j, if $f_j > 0$

D.1.2 Layer-Wise Relevance Propagation

Layer-wise relevance propagation is an XAI method that runs backwards through a neural network. At each step, it calculates the relevance that the current neuron has for the deeper layer and distributes it accordingly to the next higher layer's neurons.

A conservation property ensures that the combined relevance is preserved while being propagated to the higher layers and is illustrated in figure D.1. According to [151], this reasoning is analogous to Kirchhoff's laws in electrical circuits and used similarly in other XAI techniques such as [288, 289, 290].
For the neuralised k-means algorithm 1, [143] suggests the following propagation rule of relevance ρ_l from f_j backwards to h_l:

$$\rho_l = \frac{\exp(-\beta h_l)}{\sum_{l \neq j} \exp(-\beta h_l)} f_j \tag{D.1}$$

with the inverse mean $\beta = \mathbb{E}[f_j]^{-1}$ of f_j over the entire data. The second rule for the propagation of relevance ρ_l of each node h_l to the original input features c_i of $\mathbf{C}$ is given by

$$\rho_i = \sum_{l \neq j} \frac{(c_i - m_{i,l}) w_{i,l}}{\sum_i (x_i - m_{i,l}) w_{i,l}} \rho_l \tag{D.2}$$

with $\mathbf{m}_l = (\mathbf{C}_j + \mathbf{C}_l)/2$ as the point directly between the centroids and $w_{i,l}$ as the i^{th} component of the difference vector $\mathbf{w}_l$ between the cluster centroids j and l (cf. the pseudocode in 1). Note that in our study, each c_i is a correlation like in equation (4.3).

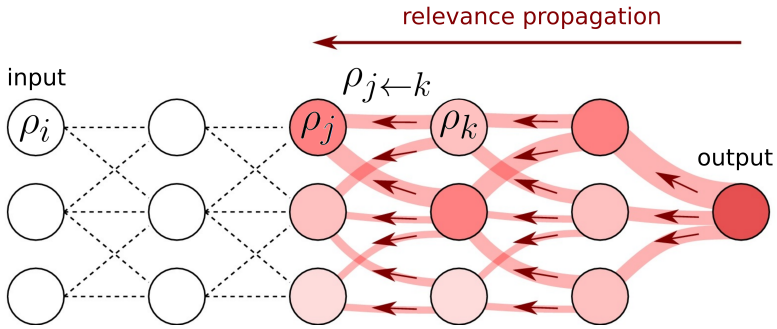

Figure D.1: Illustration of the layer-wise relevance propagation, taken from [84] and adapted under a CC BY 4.0 license (`https://creativecommons.org/licenses/by/4.0/`) to stay consistent with this thesis's notation.

D.2 Elbow Plots

The same plots as in figure 4.5 are depicted in this section for the other seven clusters. While a pronounced elbow is visible for all clusters in the mode-mode aggregation, the median method often shows a continuum.

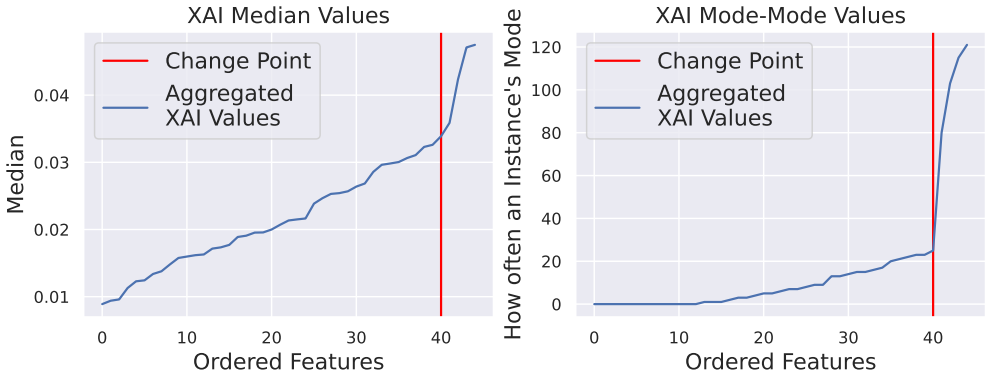

Figure D.2: The same figures as in figure 4.5, but now for cluster 0.

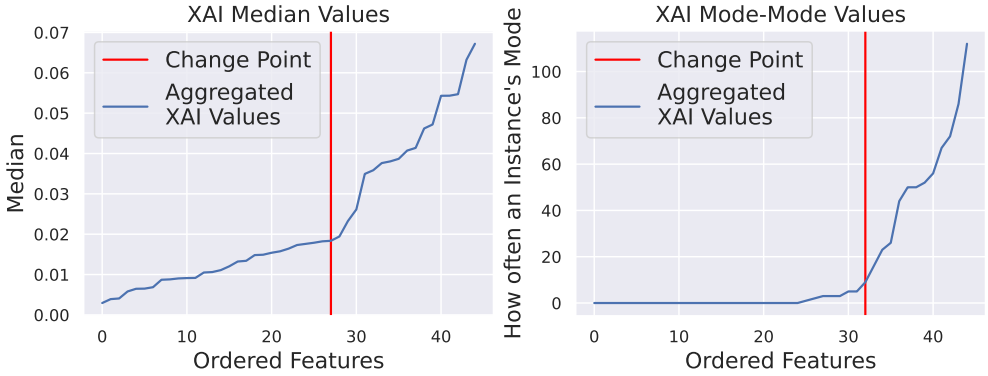

Figure D.3: The same figures as in figure 4.5, but now for cluster 1.

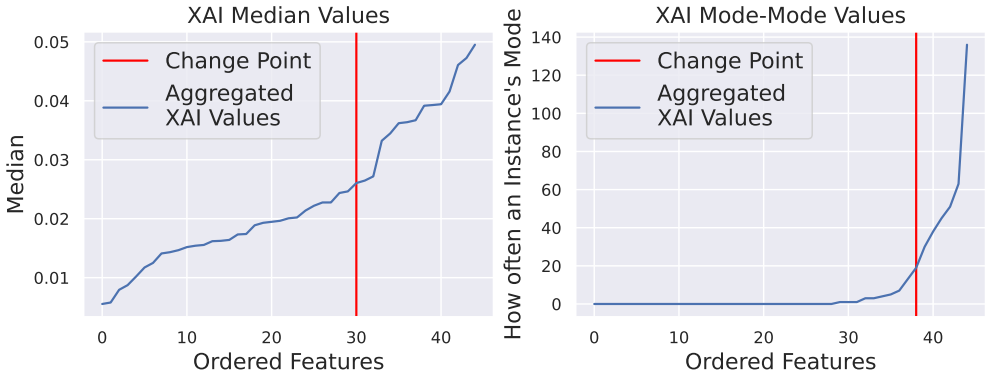

Figure D.4: The same figures as in figure 4.5, but now for cluster 3.

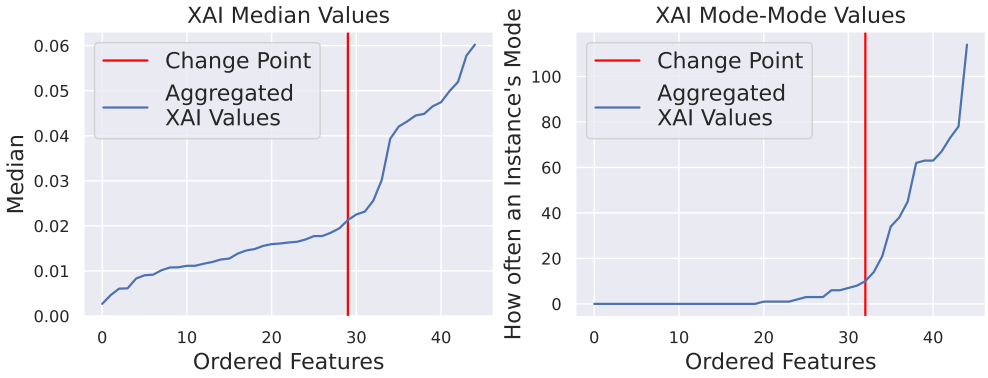

Figure D.5: The same figures as in figure 4.5, but now for cluster 4.

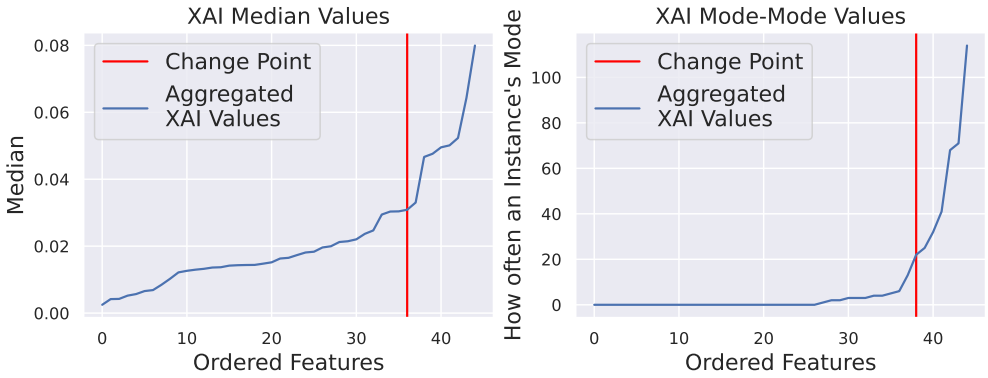

Figure D.6: The same figures as in figure 4.5, but now for cluster 5.

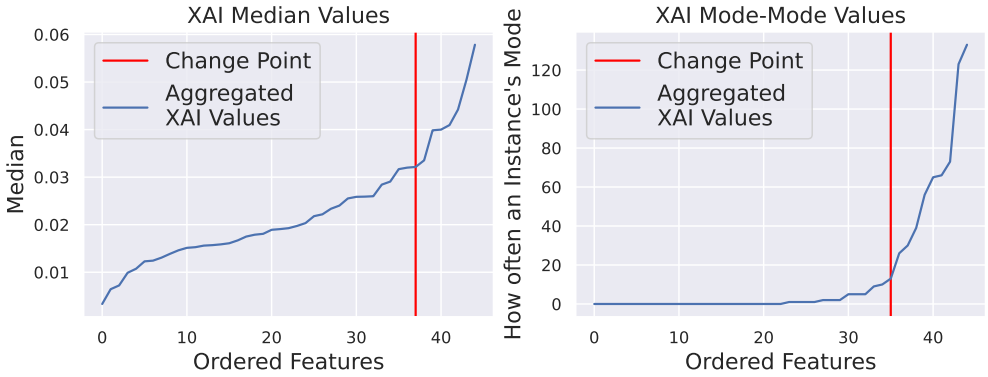

Figure D.7: The same figures as in figure 4.5, but now for cluster 6.

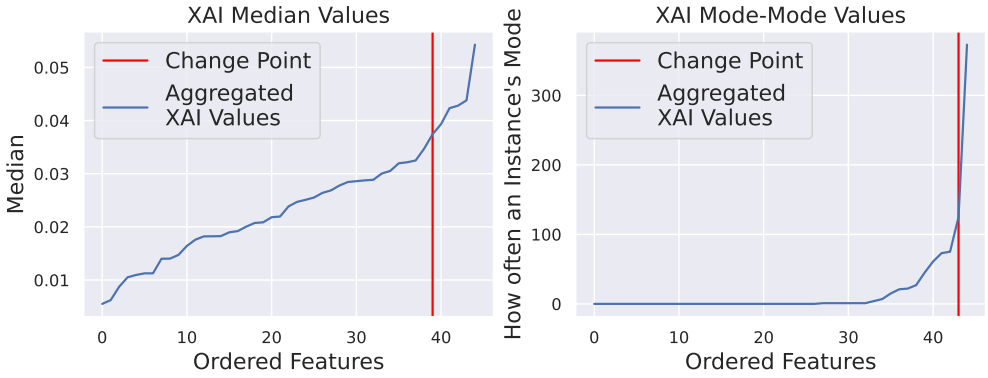

Figure D.8: The same figures as in figure 4.5, but now for cluster 7.

D.3 Neural Network Architecture

The neural network used as a surrogate model in this thesis is a deep multi-layer perceptron network for a classification task to predict which of the eight clusters the instance belongs to. The architecture of such a network is given in table D.1. Note that two different types of layers are used here, *dense* and *dropout* layers. If the connections are dense, then all N_n previous outputs are combined in linear combinations to contribute to the next layer's inputs. The weights of the linear combinations are then optimised to achieve the best accuracy. If the connections are in a dropout layer of rate β, then at every optimisation step, β of the connections are forcibly set to 0 to prevent an overfitting of the network (i.e. that the network is fit so closely to the known data that it does not generalise well on unknown data). The very first layer (input layer) receives the features (observables) of the data (i.e. the correlations between sectors) and the output layer is

Layer	Dense	Dense	Dense	Dense	Dense	Dropout	Dense
Units	256	128	128	1024	128	-	8
Activation	selu	relu	relu	LeakyReLu	LeakyReLu	-	softmax
Parameter	-	-	-	0.05	0.01	0.3	-

Table D.1: Architecture of the neural network with 100 training epochs for the surrogate model in section 4.2.2 from left to right. There are eight input units for the eight features of the surrogate model and eight output units for the eight clusters.

giving a prediction about which of the 8 clusters the data belongs to.
The softmax function is often used to get from the continuous values of the neurons' outputs to an actual prediction. If the network has to choose between k possible predictions, then it will end with k linear combinations $l_1, \ldots, l_k$ of the neurons' outputs in its penultimate layer. For the k possible classes it transforms these linear combinations into probabilities $p_1, \ldots, p_k$ via the softmax function following

$$\text{softmax}(l_i) = \frac{\exp\{(l_i)\}}{\sum_i \exp\{(l_i)\}}.$$

Note that $\sum_i p_i = 1$. Finally, it uses the highest probability as its prediction. Further details on the construction of neural networks can be found e.g. in [72] and technical details on the implementation of the neural network are given in [76, 291]. The network architecture used in this thesis is detailed in table D.1.

E Details on the Generalised Langevin Equation

E.1 GLE: Overlapping Windows for the Mean Correlation

If the mean correlation is calculated via a window size of $\tau = 42$ and a shift of $s = 1$, as in [117], then the GLE fit indicates some problems: The estimated memory kernel shows very strong absolute values at $\mathcal{K}_{42}$ and the doubled temporal distance $\mathcal{K}_{84}$. Although the memory kernel suggests that there is a strong memory effect, this is most probably only attributable to the overlap of the window ranges used to calculate $\bar{C}$.

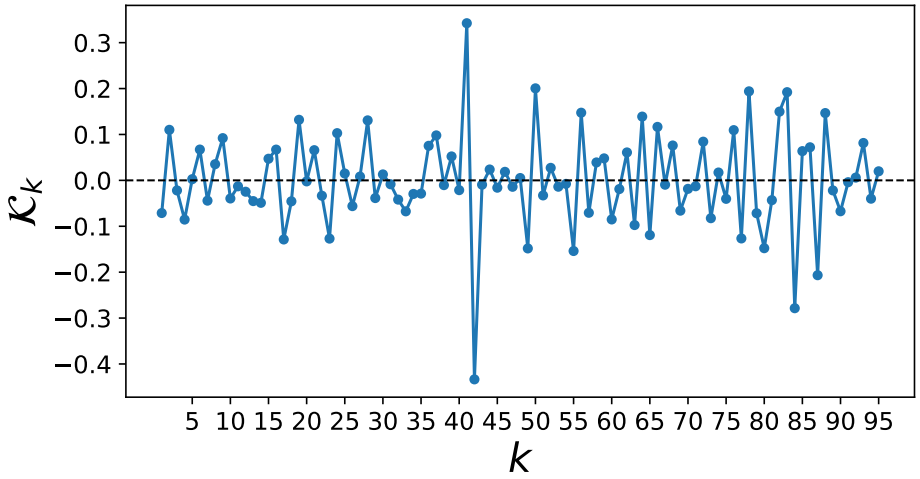

Figure E.1: Example for a model with a long kernel memory for correlation matrix window size $\tau = 42$ and shift $s = 1$ with the lag k in days on the x-axis. The kernel has large negative values for $\mathcal{K}_\tau$ and $\mathcal{K}_{2\tau}$. This probably indicates that the overlap of the artificially created windows ends at a distance of τ instead of reflecting a real-world relationship.

E.2 LE: Increment Distribution

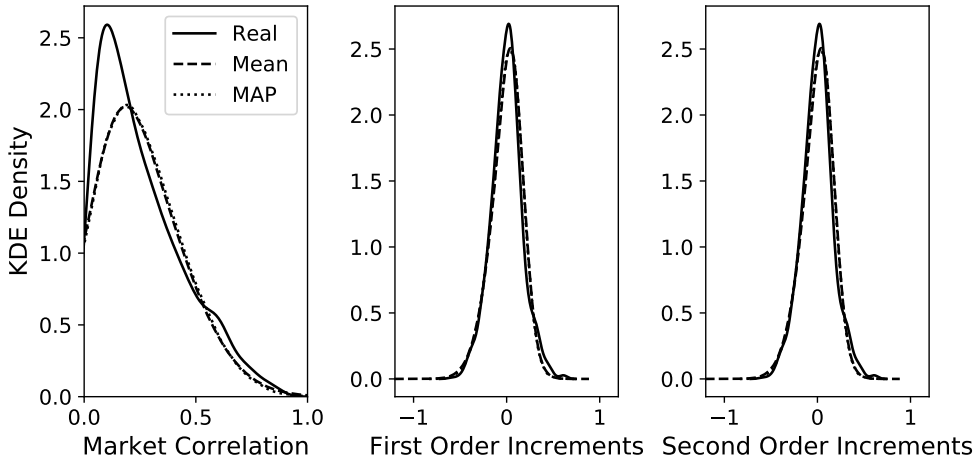

Figure E.2: The same plots as in Figure 4.11 for the Langevin Equation without any memory kernel. The distributions look very similar to the ones in Figure 4.11.

E.3 GLE: Prediction

The results of the naive forecasting method are compared to the GLE in this figure, showcasing the superior accuracy of the GLE, which can also be seen by comparing the ρ^2 values in Table 4.2.

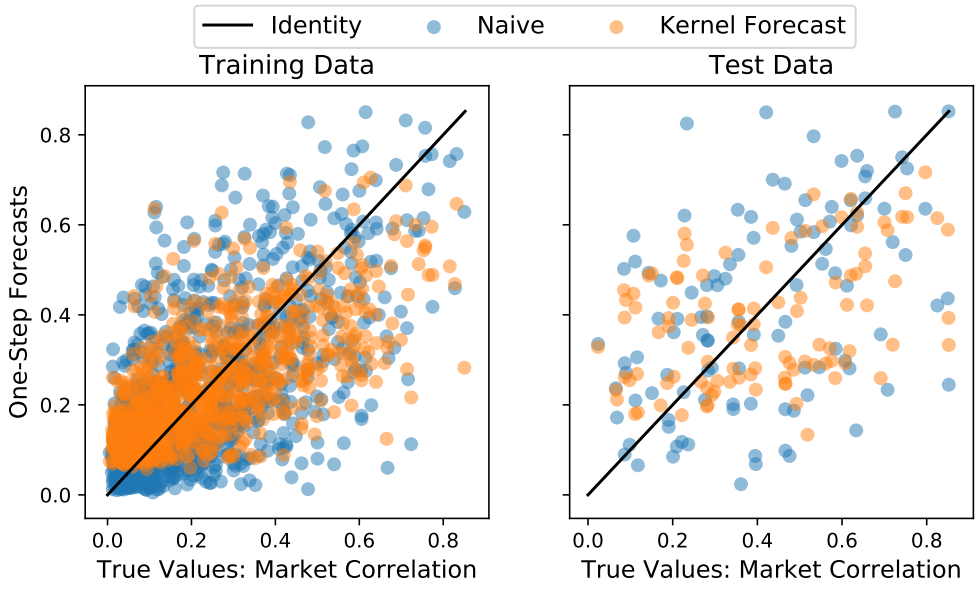

Figure E.3: The same comparison between the prediction results as in Figure 4.13, but now for the naive forecasting method $\hat{y}_{t+1} = y_t$ vs. the GLE. Again, $\alpha = 90\%$.

F Additional Plots for the non-Ergodic Insurance Model

F.1 ACF for the Richest Decile

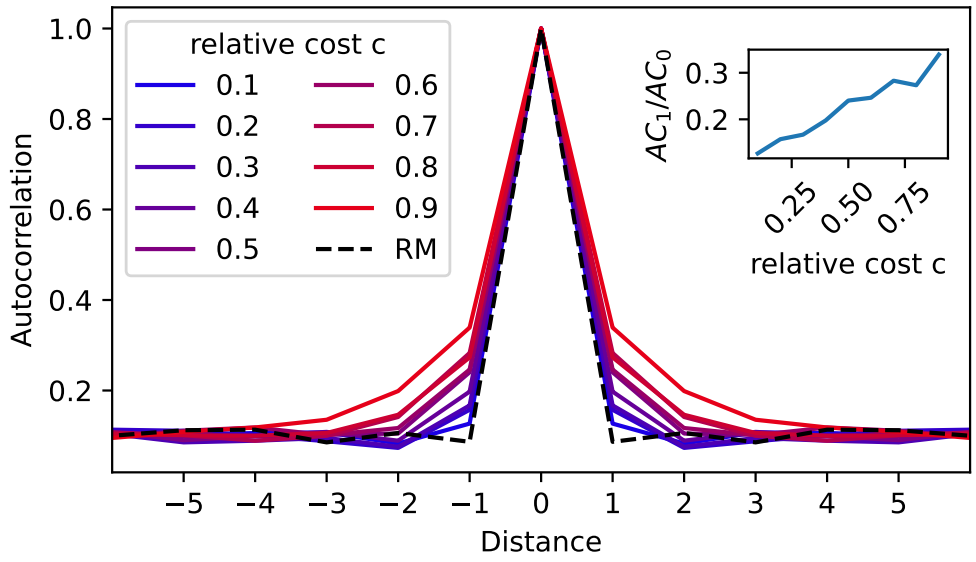

Figure F.1: The same plot as fig. 6.5 but for the richest decile. The ACF seems to continuously increase with rising c.

F.2 Highly Volatile Regime

The phase space in fig. 6.4 reveals that a particularly high clustering for both the top and bottom quantile can be found in a region of high c and high r, i.e. where the volatility of the gamble is very large in both directions. While the low-volatile regime transitions into the deterministic case, it is not obvious how the model behaves in the high-volatile regime. To investigate this, we explore a highly volatile system with $c = 0.95$ and $r = 2$.

In figure F.2 (left), we show the richest (white), middle (grey) and poorest (black) third of the agents and the ACF (right panel) of the richest and poorest thirds. We see that the ACF for both have fat tails compared to the ACF of a random ensemble, reflecting the high degree of clustering seen in the left panel. The formation of large, macrocsopic neighbourhood areas for both the richest and the poorest third is clearly visible in this plot, whereas the middle third mostly forms a thin boundary layer between the two extremes instead of manifesting into large-scale neighbourhoods. Hence, the neighbourhood map shows a high polarization, reminiscent of the magnetic domains in ferromagnets. This superficial similarity should not be taken too seriously, though. The magnetic domains form because of energetic considerations of the macroscopic magnetic field, whereas our model and the Ising model are microscopic descriptions.

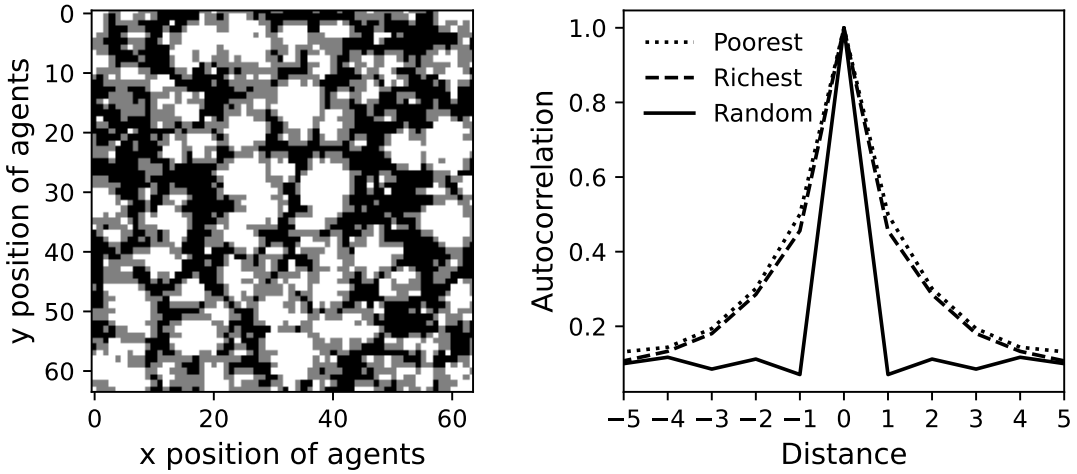

Figure F.2: Results after $T = 200$ time units for $N = 64$ in the highly volatile regime of $c = 0.95$ and $r = 2$. Left: The colour coding now shows the agents divided into three categories as the richest (white), central (grey) and poorest thirds (black). Right: The ACF of the richest and poorest decile now shows much wider tails than in the previous parameter setting in fig. 6.3.

G Seshat Databank

G.1 Example of the Data Pre-Processing

We illustrate the pre-processing of the raw SPC1 time series using the NGA "Latium" (modern day central Italy) as an example. Figure G.1 shows the shift between the original time series to RelTime, i.e. to the time frame relative to the anchor time for Latium $T_a^{(Latium)} = 500$ BC. Additionally, this figure depicts how the SPC1 time series of Latium is dissected into culturally or institutionally continuous time series intervals of SPC values for the NGA. Note that the time index is given in $RelTime$ for the shifted data and real-world time for the original data.

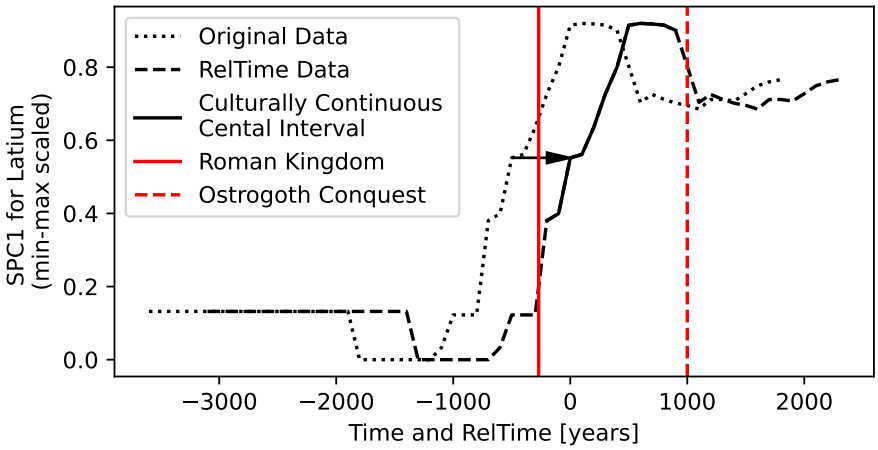

Figure G.1: Illustration of how the SPC1 time series of the Latium NGA is shifted by its anchor time to the new relative time frame $RelTime$. Hence, at $RelTime = 0$, the shifted time series is equal to the threshold value SPC1_0. The two recorded discontinuities for the Latium NGA (red) are used to divide its SPC1 time series into different intervals. The interval containing the threshold value SPC1_0 (the "relevant snippet" with the solid line) is used for further analysis whereas the rest (dotted) is discarded. Note that the times of the discontinuities do not line up with the 100-year time intervals of the SPC1 measurement which is why the transition between the black dashed/solid lines does not perfectly align with the red vertical lines.

G.2 Detailed Data

The SPC1 data from figure 8.1 is shown in figure G.2, but spread out onto several subplots to help the reader identify different NGAs.

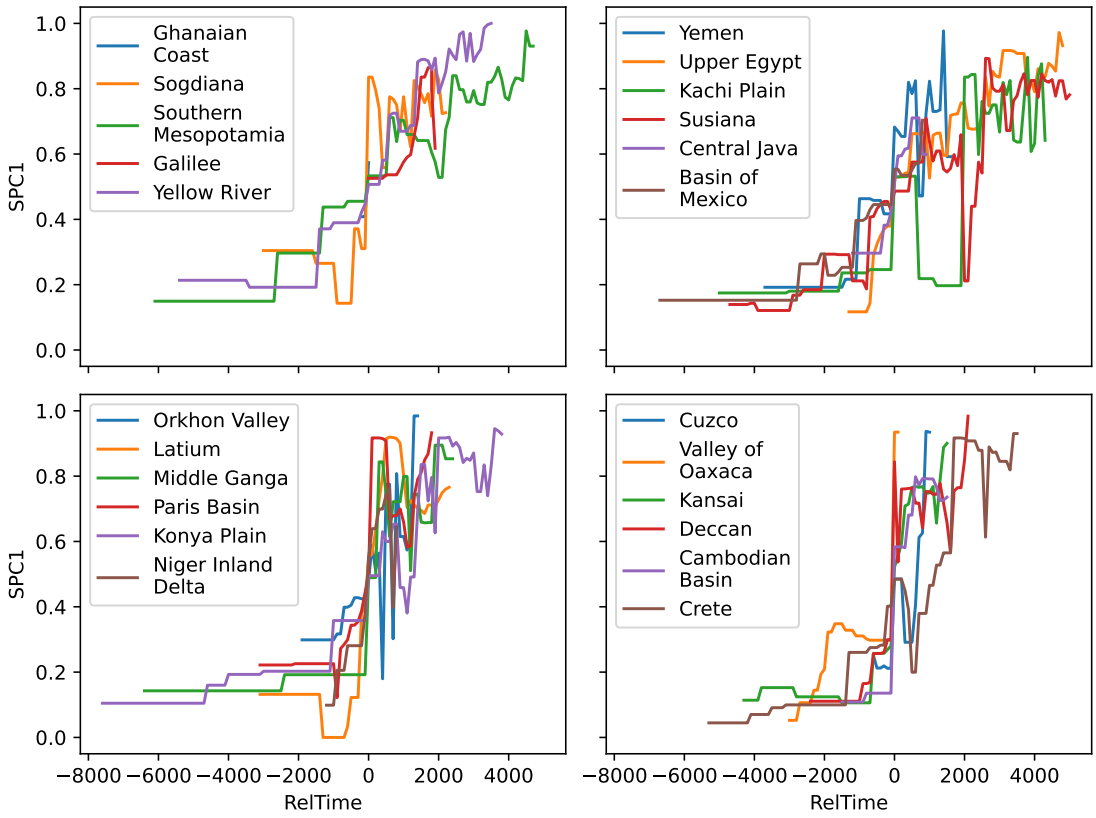

Figure G.2: Spreading the various time series from figure 8.1 onto multiple plots to make it easier to distinguish the NGAs from each other.

G.3 Data Collection and Availability

As described in the main text, the data used in chapter 8 of this thesis is derived from [275], supplemented with information provided by the authors. The original data is described in [252]. It was gathered, cleaned, reviewed, and managed by members of the Seshat project following standard project methods, as described in the references cited in the main text. For more details on the social complexity data utilised here and development of the SPC1 values, see especially [50] and [254].
The full dataset used in the analyses presented here is available on [279]. This shows:

- **NGA**: Name of the NGA
- **PolID**: Unique identifier for each polity in the sample.
- **AbsTime**: The "absolute" or calendar time-point for each SPC1 value. Note that we sample at 100-year intervals, for as many intervals as there are polities in the sample occupying each NGA.
- **RelTime**: The shifted time series, such that RelTime = 0 in the century interval in each NGA during which SPC1 values crossed the calculated threshold, as described in the main text and illustrated in figure G.1. Other rows in the NGA express their relation before or after this threshold, in century intervals. Note that rows that fall outside of the central cultural or institutional sequence are not expressed.
- **SPC1**: Raw SPC1 values for each polity at each century interval, calculated as described in the main text (but not yet min-max scaled).
- **Culture.Sequence**: Label indicating the central time-series identified as part of the culturally-continuous sequence that surrounds each NGA's RelTime=0 threshold (labelled "cultural.continuity"); all other century intervals in the NGA are labelled "outside.central" to indicate they fall outside of this central interval sequence.
- **Institutions.Sequence**: Same as above, but indicating the institutionally continuous sequence.

Bibliography

[1] S.H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Studies in nonlinearity. Westview, 2000.

[2] Manlio De Domenico et al. *Complexity Explained*. 2022. DOI: 10.17605/OSF.IO/TQGNW.

[3] Ernst Ising. 'Beitrag zur Theorie des Ferromagnetismus'. In: *Zeitschrift für Physik* 31.1 (Feb. 1925), pp. 253–258. DOI: 10.1007/bf02980577.

[4] Lars Onsager. 'Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition'. In: *Phys. Rev.* 65 (3-4 Feb. 1944), pp. 117–149. DOI: 10.1103/PhysRev.65.117.

[5] Kevin P. O'Keeffe, Hyunsuk Hong and Steven H. Strogatz. 'Oscillators that sync and swarm'. In: *Nature Communications* 8.1 (Nov. 2017). DOI: 10.1038/s41467-017-01190-3.

[6] Michael Tinkham. *Introduction to Superconductivity*. 2nd ed. Dover Books on Physics. Mineola, NY: Dover Publications, June 2004.

[7] Per Bak, Chao Tang and Kurt Wiesenfeld. 'Self-organized criticality: An explanation of the 1/f noise'. In: *Physical Review Letters* 59.4 (July 1987), pp. 381–384. DOI: 10.1103/physrevlett.59.381.

[8] Didier Sornette. 'Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth'. In: *Proceedings of the National Academy of Sciences* 99.suppl$_1$ (Feb. 2002), pp. 2522–2529. DOI: 10.1073/pnas.022581999.

[9] A. Einstein. 'Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen'. In: *Annalen der Physik* 322.8 (Jan. 1905), pp. 549–560. DOI: 10.1002/andp.19053220806.

[10] William D Callister and David G Rethwisch. *Materials Science and Engineering*. Nashville, TN: John Wiley & Sons, Nov. 2013.

[11] K. Hasselmann. 'Stochastic climate models Part I. Theory'. In: *Tellus* 28.6 (Dec. 1976), pp. 473–485. DOI: 10.3402/tellusa.v28i6.11316.

[12] Martin Steinhauser. *Computational multiscale modeling of fluids and solids*. en. Berlin, Germany: Springer, Oct. 2010.

[13] Auguste Comte and Jean-Paul Enthoven. *Physique sociale*. fr. Paris: Hermann, 1975. ISBN: 9782705657314.

[14] F. A. Hayek. 'The Use of Knowledge in Society'. In: *The American Economic Review* 35.4 (1945), pp. 519–530. ISSN: 00028282. URL: http://www.jstor.org/stable/1809376 (visited on 30/04/2024).

[15] Friedrich August von Hayek. *Der Weg zur Knechtschaft*. Originally published as 'The Road to Serfdom' in 1944. Reinbek, Germany: Olzog, Feb. 2017.

[16] Carlo Requião da Cunha. *Introduction to Econophysics: Contemporary Approaches with Python Simulations*. CRC Press, Aug. 2021. DOI: 10.1201/9781003127956.

[17] Neil F Johnson, Paul Jefferies and Pak Ming Hui. *Financial Market Complexity*. en. Oxford Finance Series. London, England: Oxford University Press, Aug. 2003.

[18] Rosario Mantegna and H. Stanley. *An Introduction to Econophysics*. Correlations and Complexity in Finance. Cambridge University Press, 2000.

[19] Didier Sornette. *Why Stock Markets Crash*. Princeton, NJ: Princeton University Press, Feb. 2004.

[20] Stanislao Gualdi et al. 'Tipping points in macroeconomic agent-based models'. In: *Journal of Economic Dynamics and Control* 50 (Jan. 2015), pp. 29–61. DOI: 10.1016/j.jedc.2014.08.003.

[21] Hermann Haken. *Synergetik*. 3rd ed. New York, NY: Springer, Jan. 1990.

[22] Wei-Bin Zhang. *Synergetic Economics*. Springer Berlin Heidelberg, 1991. ISBN: 9783642759093. DOI: 10.1007/978-3-642-75909-3.

[23] Eugene F. Fama. 'Efficient Capital Markets: A Review of Theory and Empirical Work'. In: *The Journal of Finance* 25.2 (May 1970), p. 383. DOI: 10.2307/2325486.

[24] Fabrizio Lillo, Szabolcs Mike and J. Doyne Farmer. 'Theory for long memory in supply and demand'. In: *Physical Review E* 71.6 (June 2005). DOI: 10.1103/physreve.71.066122.

[25] Yuki Sato and Kiyoshi Kanazawa. 'Inferring Microscopic Financial Information from the Long Memory in Market-Order Flow: A Quantitative Test of the Lillo-Mike-Farmer Model'. In: *Phys. Rev. Lett.* 131 (19 Nov. 2023), p. 197401. DOI: 10.1103/PhysRevLett.131.197401.

[26] Louis Bachelier. 'Théorie de la spéculation'. In: *Annales scientifiques de l'École Normale Supérieure* 3e série, 17 (1900), pp. 21–86.

[27] Fischer Black and Myron Scholes. 'The Pricing of Options and Corporate Liabilities'. In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. URL: http://www.jstor.org/stable/1831029 (visited on 17/04/2023).

[28] Robert C. Merton. 'Theory of Rational Option Pricing'. In: *The Bell Journal of Economics and Management Science* 4.1 (1973), pp. 141–183. URL: http://www.jstor.org/stable/3003143 (visited on 17/04/2023).

[29] Rosario N. Mantegna and H. Eugene Stanley. 'Scaling behaviour in the dynamics of an economic index'. In: *Nature* 376 (July 1995), pp. 46–49.

[30] Benoit B. Mandelbrot. *Fractals and Scaling in Finance*. Springer New York, 1997. ISBN: 9781475727630. DOI: 10.1007/978-1-4757-2763-0.

[31] O. Peters and M. Gell-Mann. 'Evaluating gambles using dynamics'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26.2 (Feb. 2016), p. 023103. DOI: 10.1063/1.4940236.

[32] Ole Peters. 'The time resolution of the St Petersburg paradox'. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369.1956 (Dec. 2011), pp. 4913–4931. DOI: 10.1098/rsta.2011.0065.

[33] Ole Peters. 'Insurance as an ergodicity problem'. In: *Annals of Actuarial Science* 17.2 (2023), pp. 215–218. DOI: 10.1017/S1748499523000131.

[34] Laurent Laloux et al. 'Noise Dressing of Financial Correlation Matrices'. In: *Phys. Rev. Lett.* 83 (7 Aug. 1999), pp. 1467–1470. DOI: 10.1103/PhysRevLett.83.1467.

[35] Vasiliki Plerou et al. 'Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series'. In: *Phys. Rev. Lett.* 83 (7 Aug. 1999), pp. 1471–1474.

[36] Michael C. Münnix et al. 'Identifying States of a Financial Market'. In: *Scientific Reports* 2.1 (Sept. 2012). DOI: 10.1038/srep00644.

[37] S. Bornholdt and H. G. Schuster, eds. *Handbook of Graphs and Networks*. Weinheim: Wiley-VCH, 2005.

[38] G. Caldarelli and A. Vespignani, eds. *Large Scale Structure and Dynamics of Complex Networks*. Singapore: World Scientific Publishing, 2007.

[39] Ma Ángeles Serrano and Marián Boguñá. 'Topology of the world trade web'. In: *Phys. Rev. E* 68 (1 July 2003), p. 015101. DOI: 10.1103/PhysRevE.68.015101.

[40] M. E. J. Newman, S. H. Strogatz and D. J. Watts. 'Random graphs with arbitrary degree distributions and their applications'. In: *Physical Review E* 64.2 (July 2001). DOI: 10.1103/physreve.64.026118.

[41] Joseph B. Kruskal. 'On the shortest spanning subtree of a graph and the traveling salesman problem'. In: *Proceedings of the American Mathematical Society* 7.1 (1956), pp. 48–50. DOI: 10.1090/s0002-9939-1956-0078686-7.

[42] Faheem Aslam et al. 'Network analysis of global stock markets at the beginning of the coronavirus disease (Covid-19) outbreak'. In: *Borsa Istanbul Review* 20 (2020), S49–S61. DOI: 10.1016/j.bir.2020.09.003.

[43] James L. Johnson and Tom Goldring. 'Discrete Hodge Theory on Graphs: A Tutorial'. In: *Computing in Science & Engineering* 15.5 (Sept. 2013), pp. 42–55. DOI: 10.1109/mcse.2012.91.

[44] Alexander Strang. 'Applications of the Helmholtz-Hodge Decomposition to Networks and Random Processes'. Available at https://case.edu/math/thomas/Strang-Alexander-2020-PhD-thesis-final.pdf (visited on 04/11/2024). PhD thesis. Cleveland, Ohio: Case Western Reserve University, Aug. 2020.

[45] D. Stauffer. 'Social applications of two-dimensional Ising models'. In: *American Journal of Physics* 76.4 (Apr. 2008), pp. 470–473. DOI: 10.1119/1.2779882.

[46] A. Grabowski and R.A. Kosiński. 'Ising-based model of opinion formation in a complex network of interpersonal interactions'. In: *Physica A: Statistical Mechanics and its Applications* 361.2 (Mar. 2006), pp. 651–664. DOI: 10.1016/j.physa.2005.06.102.

[47] Peter Turchin. 'Arise 'cliodynamics''. In: *Nature* 454.7200 (2008), pp. 34–35. DOI: 10.1038/454034a.

[48] Peter Turchin. *Historical Dynamics: Why States Rise and Fall*. Princeton University Press, 2003.

[49] Peter Turchin et al. 'Seshat: The Global History Databank'. In: *Cliodynamics* 6.1 (July 2015). DOI: 10.21237/c7clio6127917.

[50] Peter Turchin et al. ‘Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization’. In: *PNAS* 115.2 (2018), E144–E151. DOI: 10.1073/pnas.1708800115.

[51] Peter Turchin et al. ‘Rise of the war machines: Charting the evolution of military technologies from the Neolithic to the Industrial Revolution’. In: *PLOS ONE* 16.10 (Oct. 2021), pp. 1–23. DOI: 10.1371/journal.pone.0258161.

[52] Harvey Whitehouse et al. ‘Testing the Big Gods hypothesis with global historical data: a review and “retake”’. In: *Religion, Brain & Behavior* (June 2022), pp. 1–43. DOI: 10.1080/2153599X.2022.2074085.

[53] S. Vakulenko et al. ‘Rise of nations: Why do empires expand and fall?’ In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30.9 (Sept. 2020). DOI: 10.1063/5.0004795.

[54] Florian Schunck et al. ‘A Dynamic Network Model of Societal Complexity and Resilience Inspired by Tainter’s Theory of Collapse’. In: *Entropy* 26.2 (Jan. 2024), p. 98. DOI: 10.3390/e26020098.

[55] Wolfgang von der Linden, Volker Dose and Udo von Toussaint. *Bayesian Probability Theory*. Applications in the Physical Sciences. Cambridge University Press, 2014.

[56] Tobias Wand. ‘Introduction to Artificial Intelligence’. In: *Artificial Intelligence and Intelligent Matter*. Ed. by Michael te Vrugt. In Production. Springer, 2025.

[57] Sauro Succi and Peter V. Coveney. ‘Big data: the end of the scientific method?’ In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 377.2142 (2019), p. 20180145. DOI: 10.1098/rsta.2018.0145.

[58] Rudolf Friedrich et al. ‘Approaching complexity by stochastic methods: From biological systems to turbulence’. In: *Physics Reports* 506.5 (2011), pp. 87–162. DOI: 10.1016/j.physrep.2011.05.003.

[59] Alistair I Mees, ed. *Nonlinear dynamics and statistics*. 2001st ed. Secaucus, NJ: Birkhauser Boston, Jan. 2001.

[60] M Reza Rahimi Tabar. *Analysis and data-based reconstruction of complex nonlinear dynamical systems*. 1st ed. Understanding Complex Systems. Cham, Switzerland: Springer Nature, Aug. 2020.

[61] Andy Lawrence. *Probability in physics*. 1st ed. Undergraduate Lecture Notes in Physics. Cham, Switzerland: Springer Nature, Sept. 2019.

[62] Emanuel Parzen. 'On Estimation of a Probability Density Function and Mode'. In: *The Annals of Mathematical Statistics* 33.3 (Sept. 1962), pp. 1065–1076. DOI: 10.1214/aoms/1177704472.

[63] Murray Rosenblatt. 'Remarks on Some Nonparametric Estimates of a Density Function'. In: *The Annals of Mathematical Statistics* 27.3 (Sept. 1956), pp. 832–837. DOI: 10.1214/aoms/1177728190.

[64] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Springer US, 1994. DOI: 10.1201/9780429246593.

[65] Ian T. Jolliffe and Jorge Cadima. 'Principal component analysis: a review and recent developments'. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (Apr. 2016), p. 20150202. DOI: 10.1098/rsta.2015.0202.

[66] Hirotogu Akaike. 'Information Theory and an Extension of the Maximum Likelihood Principle'. In: *Selected Papers of Hirotugu Akaike*. Ed. by Emanuel Parzen, Kunio Tanabe and Genshiro Kitagawa. New York, NY: Springer New York, 1998, pp. 199–213. DOI: 10.1007/978-1-4612-1694-0_15.

[67] Kenneth P. Burnham, David R. Anderson and Kathryn P. Huyvaert. 'AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons'. In: *Behavioral Ecology and Sociobiology* 65.1 (Aug. 2010), pp. 23–35. DOI: 10.1007/s00265-010-1029-6.

[68] Devinderjit Sivia and John Skilling. *Data analysis*. 2nd ed. London, England: Oxford University Press, June 2006.

[69] W. K. Hastings. 'Monte Carlo sampling methods using Markov chains and their applications'. In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97.

[70] Coryn A. L. Bailer-Jones. *Practical Bayesian Inference: A Primer for Physical Scientists*. Cambridge University Press, 2017.

[71] Daniel Foreman-Mackey et al. 'emcee: The MCMC Hammer'. In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pp. 306–312. DOI: 10.1086/670067.

[72] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol: O'Reilly Media, Inc., 2022. ISBN: 978-1-098-12247-8.

[73] Collin J. Delker. *Schemdraw*. https://schemdraw.readthedocs.io/en/stable/index.html (accessed 04/11/2024).

[74] Warren S. McCulloch and Walter Pitts. 'A logical calculus of the ideas immanent in nervous activity'. In: *The Bulletin of Mathematical Biophysics* 5.4 (Dec. 1943), pp. 115–133. DOI: 10.1007/bf02478259.

[75] F. Pedregosa et al. 'Scikit-learn: Machine Learning in Python'. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[76] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available on https://tensorflow.org (accessed 04/11/2024). 2015.

[77] Eli Stevens, Luca Antiga and Thomas Viehmann. *Deep Learning with PyTorch*. New York: Simon and Schuster, 2020. ISBN: 978-1-617-29526-3.

[78] BaFin - Federal Agency for Financial Services Supervision. *Big data meets artificial intelligence*. Challenges and implications for the supervision and regulation of financial services. https://www.bafin.de/SharedDocs/Downloads/EN/dl_bdai_studie_en.html (accessed on 04/11/2024). 2018.

[79] Sebastian Lapuschkin et al. 'Unmasking Clever Hans Predictors and Assessing What Machines Really Learn'. In: *Nature Communications* 10 (Mar. 2019). DOI: 10.1038/s41467-019-08987-4.

[80] Christoph Molnar. *Interpretable Machine Learning*. A Guide for Making Black Box Models Explainable. 2nd ed. https://christophm.github.io/interpretable-ml-book/ (accessed on 04/11/2024). 2022.

[81] Wojciech Samek et al., eds. *Explainable AI*. Interpreting, Explaining and Visualizing Deep Learning. Springer International Publishing, 2019. DOI: 10.1007/978-3-030-28954-6.

[82] Lloyd S. Shapley. 'A value for n-person games'. In: *The Shapley Value*. Cambridge University Press, Oct. 1988, pp. 31–40. DOI: 10.1017/cbo9780511528446.003.

[83] Scott M Lundberg and Su-In Lee. 'A Unified Approach to Interpreting Model Predictions'. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Available at https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf (accessed on 04/11/2024). Curran Associates, Inc., 2017.

[84] Grégoire Montavon, Wojciech Samek and Klaus-Robert Müller. 'Methods for interpreting and understanding deep neural networks'. In: *Digital Signal Processing* 73 (2018), pp. 1–15. DOI: 10.1016/j.dsp.2017.10.011.

[85] Cynthia Rudin. 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead'. In: *Nature Machine Intelligence* 1.5 (May 2019), pp. 206–215. DOI: 10.1038/s42256-019-0048-x.

[86] M. Raissi, P. Perdikaris and G.E. Karniadakis. 'Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations'. In: *Journal of Computational Physics* 378 (2019), pp. 686–707. DOI: https://doi.org/10.1016/j.jcp.2018.10.045.

[87] Erica Thompson and Leonard Smith. 'Escape from model-land'. In: *Economics: The Open-Access, Open-Assessment E-Journal* 13 (Oct. 2019). DOI: 10.5018/economics-ejournal.ja.2019-40.

[88] Erica Thompson. *Escape from model land*. London, England: John Murray, Nov. 2022.

[89] Judea Pearl, Madelyn Glymour and Nicholas P. Jewell. *Causal Inference in Statistics: A Primer*. Wiley, 2016. ISBN: 978-1-119-18684-7.

[90] C. W. J. Granger. 'Investigating Causal Relations by Econometric Models and Cross-spectral Methods'. In: *Econometrica* 37.3 (1969), pp. 424–438.

[91] Ethan R. Deyle and George Sugihara. 'Generalized Theorems for Nonlinear State Space Reconstruction'. In: *PLOS ONE* 6.3 (Mar. 2011), pp. 1–8. DOI: 10.1371/journal.pone.0018295.

[92] George Sugihara et al. 'Detecting Causality in Complex Ecosystems'. In: *Science* 338.6106 (2012), pp. 496–500. DOI: 10.1126/science.1227079.

[93] Dmitry A. Smirnov. 'Transient and equilibrium causal effects in coupled oscillators'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.7 (2018), p. 075303. DOI: 10.1063/1.5017821.

[94] Joris Mooij, Dominik Janzing and Bernhard Schölkopf. 'From Ordinary Differential Equations to Structural Causal Models: the deterministic case'. In: *Uncertainty in Artificial Intelligence - Proceedings of the 29th Conference, UAI 2013* (Apr. 2013). DOI: 10.48550/arXiv.1304.7920.

[95] Tobias Wand et al. 'Estimating stable fixed points and Langevin potentials for financial dynamics'. In: *Phys. Rev. E* 109 (2 Feb. 2024), p. 024226. DOI: 10.1103/PhysRevE.109.024226.

[96] Christoph Renner, Joachim Peinke and Rudolf Friedrich. 'Evidence of Markov properties of high frequency exchange rate data'. In: *Physica A: Statistical Mechanics and its Applications* 298.3 (2001), pp. 499–520. DOI: 10.1016/S0378-4371(01)00269-2.

[97] Julian Tice and Nick Webber. 'A Nonlinear Model of the Term Structure of Interest Rates'. In: *Mathematical Finance* 7.2 (1997), pp. 177–209. DOI: 10.1111/1467-9965.00030.

[98] Jean-Philippe Bouchaud and Rama Cont. 'A Langevin Approach to Stock Market Fluctuations and Crashes'. In: *Physics of Condensed Matter* 6 (Dec. 1998), pp. 543–550. DOI: 10.1007/s100510050582.

[99] Kiyoshi Kanazawa et al. 'Kinetic theory for financial Brownian motion from microscopic dynamics'. In: *Phys. Rev. E* 98 (5 Nov. 2018), p. 052317. DOI: 10.1103/PhysRevE.98.052317.

[100] Min Lee, Akihiko Oba and Hideki Takayasu. 'Parameter Estimation of a Generalized Langevin Equation of Market Price'. In: Springer, Jan. 2002, pp. 260–270. ISBN: 978-4-431-66995-1. DOI: 10.1007/978-4-431-66993-7_28.

[101] M.M. Garcia et al. 'Forecast model for financial time series: An approach based on harmonic oscillators'. In: *Physica A: Statistical Mechanics and its Applications* 549 (2020), p. 124365. DOI: 10.1016/j.physa.2020.124365.

[102] Igor Halperin. *The Inverted Parabola World of Classical Quantitative Finance: Non-Equilibrium and Non-Perturbative Finance Perspective*. 2020. arXiv: 2008.03623 [q-fin.GN].

[103] Matthew F. Dixon, Igor Halperin and Paul Bilokon. *Machine Learning in Finance*. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-41068-1.

[104] Igor Halperin and Matthew Dixon. '"Quantum Equilibrium-Disequilibrium": Asset price dynamics, symmetry breaking, and defaults as dissipative instantons'. In: *Physica A: Statistical Mechanics and its Applications* 537 (Aug. 2019), p. 122187. DOI: 10.1016/j.physa.2019.122187.

[105] Silke Siegert, Rudolf Friedrich and Joachim Peinke. 'Analysis of data sets of stochastic systems'. In: *Physics Letters A* 243.5 (1998), pp. 275–280. DOI: 10.1016/S0375-9601(98)00283-7.

[106] Rudolf Friedrich et al. 'Extracting model equations from experimental data'. In: *Physics Letters A* 271.3 (2000), pp. 217–222. DOI: 10.1016/S0375-9601(00)00334-0.

[107] Rudolf Friedrich, Joachim Peinke and Christoph Renner. 'How to Quantify Deterministic and Random Influences on the Statistics of the Foreign Exchange Market'. In: *Phys. Rev. Lett.* 84 (22 May 2000), pp. 5224–5227. DOI: 10.1103/PhysRevLett.84.5224.

[108] David Kleinhans and Rudolf Friedrich. 'Maximum likelihood estimation of drift and diffusion functions'. In: *Physics Letters A* 368.3 (2007), pp. 194–198.

[109] Steven L. Brunton, Joshua L. Proctor and J. Nathan Kutz. 'Discovering governing equations from data by sparse identification of nonlinear dynamical systems'. In: *Proceedings of the National Academy of Sciences* 113.15 (2016), pp. 3932–3937.

[110] U. Fasel et al. 'Ensemble-SINDy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control'. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 478.2260 (Apr. 2022).

[111] Martin Heßler and Oliver Kamps. 'Bayesian on-line anticipation of critical transitions'. In: *New Journal of Physics* (2021). DOI: 10.1088/1367-2630/ac46d4.

[112] John H. Cochrane. 'Presidential Address: Discount Rates'. In: *The Journal of Finance* 66.4 (July 2011), pp. 1047–1108. DOI: 10.1111/j.1540-6261.2011.01671.x.

[113] Craig W Holden and Stacey Jacobsen. 'Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions'. In: *The Journal of Finance* 69.4 (2014), pp. 1747–1785.

[114] Timo Wiedemann. 'Quantifying the Beta Estimation Bias and its Implications for Empirical Asset Pricing'. In: *SSRN* (2022). DOI: 10.2139/ssrn.4179365.

[115] Peter E Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. Stochastic Modelling and Applied Probability. Berlin, Germany: Springer, Apr. 2013.

[116] Philip Rinn et al. 'Dynamics of quasi-stationary systems: Finance as an example'. In: *EPL (Europhysics Letters)* 110.6 (June 2015), p. 68003. DOI: 10.1209/0295-5075/110/68003.

[117] Yuriy Stepanov et al. 'Stability and hierarchy of quasi-stationary states: financial markets as an example'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2015.8 (Aug. 2015), P08011. DOI: 10.1088/1742-5468/2015/08/p08011.

[118] Martin Heßler and Oliver Kamps. 'Quantifying resilience and the risk of regime shifts under strong correlated noise'. In: *PNAS Nexus* 2.2 (Dec. 2022), pgac296. DOI: 10.1093/pnasnexus/pgac296.

[119] Steven L Heston. 'A closed-form solution for options with stochastic volatility with applications to bond and currency options'. In: *The review of financial studies* 6.2 (1993), pp. 327–343.

[120] Bruno Dupire et al. 'Pricing with a smile'. In: *Risk* 7.1 (1994), pp. 18–20.

[121] Clemens Willers and Oliver Kamps. 'Efficient Bayesian estimation of the generalized Langevin equation from data'. In: *Journal of Computational Physics* 497 (2024), p. 112626. DOI: 10.1016/j.jcp.2023.112626.

[122] Clemens Willers and Oliver Kamps. 'Non-parametric estimation of a Langevin model driven by correlated noise'. In: *The European Physical Journal B* 94.7 (July 2021).

[123] Tobias Wand, Martin Heßler and Oliver Kamps. 'Identifying dominant industrial sectors in market states of the S&P 500 financial data'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2023.4 (Apr. 2023), p. 043402. DOI: 10.1088/1742-5468/accce0.

[124] Tobias Wand, Martin Heßler and Oliver Kamps. 'Memory Effects, Multiple Time Scales and Local Stability in Langevin Models of the SP500 Market Correlation'. In: *Entropy* 25.9 (2023). DOI: 10.3390/e25091257.

[125] Harry Markowitz. 'Portfolio Selection'. In: *The Journal of Finance* 7.1 (Mar. 1952), p. 77. DOI: 10.2307/2975974.

[126] Jarosław Kwapień and Stanisław Drożdż. 'Physical approach to complex systems'. In: *Physics Reports* 515.3 (2012). Physical approach to complex systems, pp. 115–226. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.01.007.

[127] Rosario Nunzio Mantegna. 'Degree of correlation inside a financial market'. In: *AIP Conference Proceedings* 411.1 (1997), pp. 197–202. DOI: 10.1063/1.54189.

[128] Stephen J. Brown. 'The Number of Factors in Security Returns'. In: *The Journal of Finance* 44.5 (Dec. 1989), pp. 1247–1262. DOI: 10.1111/j.1540-6261.1989.tb02652.x.

[129] Gautier Marti et al. 'A Review of Two Decades of Correlations, Hierarchies, Networks and Clustering in Financial Markets'. In: *Progress in Information Geometry: Theory and Applications*. Ed. by Frank Nielsen. Springer International Publishing, 2021, pp. 245–274. ISBN: 978-3-030-65459-7. DOI: 10.1007/978-3-030-65459-7_10.

[130] Peter Ashwin and Marc Timme. 'When instability makes sense'. In: *Nature* 436.7047 (2005), pp. 36–37. DOI: 10.1038/436036b.

[131] Daniel Fernex, Bernd R. Noack and Richard Semaan. 'Cluster-based network modeling—From snapshots to complex dynamical systems'. In: *Science Advances* 7.25 (June 2021). DOI: 10.1126/sciadv.abf5006.

[132] Axel Hutt et al. 'Detection of fixed points in spatiotemporal signals by clustering method'. In: *Physical Review. E* 61 (June 2000), R4691–3. DOI: 10.1103/PhysRevE.61.R4691.

[133] Peter beim Graben and Axel Hutt. 'Detecting Recurrence Domains of Dynamical Systems by Symbolic Dynamics'. In: *Phys. Rev. Lett.* 110 (15 Apr. 2013), p. 154101. DOI: 10.1103/PhysRevLett.110.154101.

[134] Anton J Heckens, Sebastian M Krause and Thomas Guhr. 'Uncovering the dynamics of correlation structures relative to the collective market motion'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2020.10 (Oct. 2020), p. 103402. DOI: 10.1088/1742-5468/abb6e2.

[135] Anton J Heckens and Thomas Guhr. 'A new attempt to identify long-term precursors for endogenous financial crises in the market correlation structures'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2022.4 (Apr. 2022), p. 043401. DOI: 10.1088/1742-5468/ac59ab.

[136] Anton J. Heckens and Thomas Guhr. 'New collectivity measures for financial covariances and correlations'. In: *Physica A: Statistical Mechanics and its Applications* 604 (2022), p. 127704.

[137] Ran Aroussi. *yfinance 0.1.70*. https://pypi.org/project/yfinance/ (accessed on 04/11/2024). Jan. 2022.

[138] The pandas development team. *pandas-dev/pandas: Pandas*. Feb. 2020. DOI: 10.5281/zenodo.3509134.

[139] Wes McKinney. 'Data Structures for Statistical Computing in Python'. In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[140] Rudi Schäfer and Thomas Guhr. 'Local normalization. Uncovering correlations in non-stationary financial time series'. In: *Physica A* 389.18 (2010), pp. 3856–3865. DOI: 10.1016/j.physa.2010.05.030.

[141] S&P Dow Jones Indices. *Sector Classification*. https://www.msci.com/documents/10199/4547797/Effective+Until+February+28,+2014.xls. (accessed 03/03/2022). 2018.

[142] Tobias Wand. *S&P500 Mean Correlation Time Series (1992-2012)*. Published via Zenodo at https://doi.org/10.5281/zenodo.8167592. 2023.

[143] Jacob Kauffmann et al. 'From Clustering to Cluster Explanations via Neural Networks'. In: *IEEE Transactions on Neural Networks and Learning Systems* (2022), pp. 1–15. DOI: 10.1109/TNNLS.2022.3185901.

[144] Arnd Koeppe et al. 'Explainable Artificial Intelligence for Mechanics: Physics-Explaining Neural Networks for Constitutive Models'. In: *Frontiers in Materials* 8 (2022).

[145] Mark S. Neubauer and Avik Roy. *Explainable AI for High Energy Physics*. 2022. DOI: 10.48550/arxiv.2206.06632.

[146] Christophe Grojean et al. 'Lessons on interpretable machine learning from particle physics'. In: *Nature Reviews Physics* 4.5 (Apr. 2022), pp. 284–286. DOI: 10.1038/s42254-022-00456-0.

[147] Martin van den Berg and Ouren Kuiper. *XAI in the Financial Sector*. A Conceptual Framework for Explainable AI (XAI). https://www.internationalhu.com/research/projects/explainable-ai-in-the-financial-sector (accessed 04/11/2024). Sept. 2020.

[148] Matteo Marsili. 'Dissecting financial markets: sectors and states'. In: *Quantitative Finance* 2.4 (2002), pp. 297–302. DOI: 10.1088/1469-7688/2/4/305.

[149] Walter Böhm and Kurt Hornik. 'Generating random correlation matrices by the simple rejection method: Why it does not work'. In: *Statistics & Probability Letters* 87 (2014), pp. 27–30. DOI: 10.1016/j.spl.2013.12.012.

[150] J. MacQueen. 'Some methods for classification and analysis of multivariate observations'. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* 5.1 (1967), pp. 281–297.

[151] Grégoire Montavon et al. 'Layer-Wise Relevance Propagation: An Overview'. In: *Explainable AI. Interpreting, Explaining and Visualizing Deep Learning*. Ed. by Wojciech Samek et al. Springer International Publishing, 2019, pp. 193–209. ISBN: 978-3-030-28954-6. DOI: 10.1007/978-3-030-28954-6_10.

[152] Volker Dose and Annette Menzel. 'Bayesian analysis of climate change impacts in phenology'. In: *Global Change Biology* 10.2 (Feb. 2004), pp. 259–272. DOI: 10.1111/j.1529-8817.2003.00731.x.

[153] Giray Gozgor, Chi Keung Marco Lau and Zhou Lu. 'Energy consumption and economic growth: New evidence from the OECD countries'. In: *Energy* 153 (2018), pp. 27–34. ISSN: 0360-5442. DOI: 10.1016/j.energy.2018.03.158.

[154] Rafał Kasperowicz. 'Electricity Consumption and Economic Growth: Evidence from Poland'. In: *Journal of International Studies* 1 (May 2014), pp. 46–57.

[155] Wen-Cheng Lu. 'Electricity Consumption and Economic Growth: Evidence from 17 Taiwanese Industries'. In: *Sustainability* 9.1 (2017).

[156] Lorenzo Giada and Matteo Marsili. 'Algorithms of maximum likelihood data clustering with applications'. In: *Physica A: Statistical Mechanics and its Applications* 315.3 (2002), pp. 650–664. DOI: 10.1016/S0378-4371(02)00974-3.

[157] Leonidas Sandoval. 'A map of the Brazilian stock market'. In: *Advances in Complex Systems* 15.05 (2012), p. 1250042. DOI: 10.1142/S0219525912500427.

[158] Leonidas Sandoval and Italo De Paula Franca. 'Correlation of financial markets in times of crisis'. In: *Physica A* 391.1 (2012), pp. 187–208. DOI: 10.1016/j.physa.2011.07.023.

[159] Martin Heßler. *PhD Thesis on the Anticipation of Critical Transitions in Complex Systems*. Forthcoming, University of Münster.

[160] Mario Ragwitz and Holger Kantz. 'Indispensable Finite Time Corrections for Fokker-Planck Equations from Time Series Data'. In: *Phys. Rev. Lett.* 87 (25 Dec. 2001), p. 254501.

[161] Rudolf Friedrich et al. 'Comment on "Indispensable Finite Time Corrections for Fokker-Planck Equations from Time Series Data"'. In: *Phys. Rev. Lett.* 89 (14 Sept. 2002), p. 149401.

[162] Mario Ragwitz and Holger Kantz. 'Ragwitz and Kantz Reply:' in: *Phys. Rev. Lett.* 89 (14 Sept. 2002), p. 149402.

[163] Moritz Sieber, C. Oliver Paschereit and Kilian Oberleithner. 'Stochastic modelling of a noise-driven global instability in a turbulent swirling jet'. In: *Journal of Fluid Mechanics* 916 (2021), A7. DOI: 10.1017/jfm.2021.133.

[164] Viktor Klippenstein et al. 'Introducing Memory in Coarse-Grained Molecular Simulations'. In: *The Journal of Physical Chemistry B* 125.19 (2021), pp. 4931–4954. DOI: 10.1021/acs.jpcb.1c01120.

[165] Zbigniew Czechowski. 'Reconstruction of the modified discrete Langevin equation from persistent time series'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26.5 (2016), p. 053109. DOI: 10.1063/1.4951683.

[166] Hazime Mori. 'Transport, Collective Motion, and Brownian Motion'. In: *Progress of Theoretical Physics* 33.3 (Mar. 1965), pp. 423–455. DOI: 10.1143/PTP.33.423.

[167] Skipper Seabold and Josef Perktold. 'statsmodels: Econometric and statistical modeling with Python'. In: *9th Python in Science Conference*. 2010.

[168] Andrea Gaunersdorfer, Cars H. Hommes and Florian O.O. Wagener. 'Bifurcation Routes to Volatility Clustering under Evolutionary Learning'. In: *Journal of Economic Behavior & Organization* 67.1 (2008), pp. 27–47. DOI: 10.1016/j.jebo.2007.07.004.

[169] Xue-Zhong He, Kai Li and Chuncheng Wang. 'Volatility clustering: A nonlinear theoretical approach'. In: *Journal of Economic Behavior & Organization* 130 (2016), pp. 274–297. DOI: 10.1016/j.jebo.2016.07.020.

[170] S. Ghashghaie et al. 'Turbulent cascades in foreign exchange markets'. In: *Nature* 381.6585 (June 1996), pp. 767–770. DOI: 10.1038/381767a0.

[171] Ulrich A. Müller et al. 'Volatilities of different time resolutions — Analyzing the dynamics of market components'. In: *Journal of Empirical Finance* 4.2 (1997). High Frequency Data in Finance, Part 1, pp. 213–239. DOI: 10.1016/S0927-5398(97)00007-8.

[172] Joseph Schumpeter. *Konjunkturzyklen. Eine theoretische, historische und statistische Analyse des kapitalistischen Prozesses*. Göttingen: Vandenhoeck & Ruprecht, 1961.

[173] N. D. Kondratieff and W. F. Stolper. 'The Long Waves in Economic Life'. In: *The Review of Economics and Statistics* 17.6 (1935), pp. 105–115. ISSN: 00346535, 15309142. URL: http://www.jstor.org/stable/1928486 (visited on 04/07/2023).

[174] Tobias Wand and Dan Hoyer. 'The Characteristic Time Scale of Cultural Evolution'. In: *PNAS Nexus* 3 (2) (2024), pgae009. DOI: 10.1093/pnasnexus/pgae009.

[175] Tobias Wand, Oliver Kamps and Hiroshi Iyetomi. 'Causal Hierarchy in the Financial Market Network—Uncovered by the Helmholtz–Hodge–Kodaira Decomposition'. In: *Entropy* 26.10 (2024). DOI: 10.3390/e26100858.

[176] John Aldrich. 'Correlations Genuine and Spurious in Pearson and Yule'. In: *Statistical Science* 10.4 (1995). DOI: 10.1214/ss/1177009870.

[177] R. Quian Quiroga, J. Arnhold and P. Grassberger. 'Learning driver-response relationships from synchronization patterns'. In: *Phys. Rev. E* 61 (5 May 2000), pp. 5142–5148. DOI: 10.1103/PhysRevE.61.5142.

[178] Jie Sun, Dane Taylor and Erik M. Bollt. 'Causal Network Inference by Optimal Causation Entropy'. In: *SIAM Journal on Applied Dynamical Systems* 14.1 (2015), pp. 73–106. DOI: 10.1137/140956166.

[179] Tomaso Aste and T. Di Matteo. 'Sparse Causality Network Retrieval from Short Time Series'. In: *Complexity* 2017 (2017), pp. 1–13. DOI: 10.1155/2017/4518429.

[180] Elsa Siggiridou and Dimitris Kugiumtzis. 'Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model'. In: *IEEE Transactions on Signal Processing* 64.7 (2016), pp. 1759–1773. DOI: 10.1109/TSP.2015.2500893.

[181] Elsa Siggiridou et al. 'Evaluation of Granger Causality Measures for Constructing Networks from Multivariate Time Series'. In: *Entropy* 21.11 (Nov. 2019), p. 1080. DOI: 10.3390/e21111080.

[182] Mateusz Iskrzyński et al. 'Cycling and reciprocity in weighted food webs and economic networks'. In: *Journal of Industrial Ecology* 26.3 (Dec. 2021), pp. 838–849. DOI: 10.1111/jiec.13217.

[183] Yoshi Fujiwara and Rubaiyat Islam. 'Bitcoin's Crypto Flow Network'. In: *Proceedings of Blockchain in Kyoto 2021 (BCK21)*. Journal of the Physical Society of Japan, 2021. DOI: 10.7566/jpscp.36.011002.

[184] Yuichi Kichikawa, Hiroshi Iyetomi and Yuichi Ikeda. 'Who Possesses Whom in Terms of the Global Ownership Network'. In: In *Big Data Analysis on Global Community Formation and Isolation: Sustainability and Flow of Commodities, Money, and Humans*, edited by Yuichi Ikeda, Hiroshi Iyetomi and Takayuki Mizuno. Singapore: Springer Singapore, 2021, pp. 143–190. ISBN: 978-981-15-4944-1. DOI: 10.1007/978-981-15-4944-1_6.

[185] Leonidas Sandoval. 'Structure of a Global Network of Financial Companies Based on Transfer Entropy'. In: *Entropy* 16.8 (2014), pp. 4443–4482. DOI: 10.3390/e16084443.

[186] Angeliki Papana et al. 'Financial networks based on Granger causality: A case study'. In: *Physica A: Statistical Mechanics and its Applications* 482 (2017), pp. 65–73. DOI: 10.1016/j.physa.2017.04.046.

[187] Stavros K. Stavroglou et al. 'Hidden interactions in financial markets'. In: *Proceedings of the National Academy of Sciences* 116.22 (2019), pp. 10646–10651. DOI: 10.1073/pnas.1819449116.

[188] Ken French. *US Research Returns Data. 49 Industry Portfolios [Daily]*. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Continuously updated. Accessed: July 2024. 2024.

[189] James Mackinnon. 'Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests'. In: *Journal of Business & Economic Statistics* 12 (Feb. 1994), pp. 167–76. DOI: 10.1080/07350015.1994.10510005.

[190] Eugene F. Fama and Kenneth R. French. 'Production of U.S. SMB and HML in the Fama-French Data Library'. In: *SSRN Electronic Journal* (2023). DOI: 10.2139/ssrn.4629613.

[191] Louis K.C. Chan, Josef Lakonishok and Bhaskaran Swaminathan. 'Industry Classifications and Return Comovement'. In: *Financial Analysts Journal* 63.6 (Nov. 2007), pp. 56–70. DOI: 10.2469/faj.v63.n6.4927.

[192] Eugene F. Fama and Kenneth R. French. 'Industry costs of equity'. In: *Journal of Financial Economics* 43.2 (1997), pp. 153–193. DOI: 10.1016/S0304-405X(96)00896-3.

[193] Michael Babyak. 'What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models'. In: *Psychosomatic medicine* 66 (3) (May 2004), pp. 411–21.

[194] Alexander Strang, Karen C. Abbott and Peter J. Thomas. 'The Network HHD: Quantifying Cyclic Competition in Trait-Performance Models of Tournaments'. In: *SIAM Review* 64.2 (2022), pp. 360–391. DOI: 10.1137/20M1321012.

[195] Tobias Wand. *Helmholtz-Hodge-Kodaira Decomposition on Financial Data by Ken French*. Available on Zenodo at https://zenodo.org/doi/10.5281/zenodo.13340981. 2024.

[196] Taichi Haruna and Yuuya Fujiki. 'Hodge Decomposition of Information Flow on Small-World Networks'. In: *Frontiers in Neural Circuits* 10 (2016). DOI: 10.3389/fncir.2016.00077.

[197] Yuuya Fujiki and Taichi Haruna. 'Hodge Decomposition of Information Flow on Complex Networks'. In: *Proceedings of the 8th International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*. Jan. 2015. DOI: 10.4108/icst.bict.2014.257876.

[198] Aric Hagberg, Pieter Swart and Daniel S Chult. *Exploring network structure, dynamics, and function using NetworkX*. Tech. rep. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[199] Greg Zimmerman. *Construction Materials Prices Increase More Than 20 Percent*. Available on https://www.facilitiesnet.com/designconstruction/tip/Construction-Materials-Prices-Increase-More-Than-20-Percent--49437 and accessed on 24/07/2024. 2022.

[200] Martin Heßler, Tobias Wand and Oliver Kamps. 'Efficient Multi-Change Point Analysis to Decode Economic Crisis Information from the SP500 Mean Market Correlation'. In: *Entropy* 25.9 (2023). DOI: 10.3390/e25091265.

[201] Peterson K Ozili. 'Causes and Consequences of the 2023 Banking Crisis'. In: *SSRN Electronic Journal* (2023). DOI: 10.2139/ssrn.4407221.

[202] Delia Diaconaşu, Seyed Mehdian and Ovidiu Stoica. 'The Global Stock Market Reactions to the 2016 U.S. Presidential Election'. In: *Sage Open* 13.2 (2023). DOI: 10.1177/21582440231181352.

[203] R. S. MacKay, S. Johnson and B. Sansom. 'How directed is a directed network?' In: *Royal Society Open Science* 7.9 (2020), p. 201138. DOI: 10.1098/rsos.201138.

[204] Haochun Ma et al. *Linear and nonlinear causality in financial markets*. 2023. arXiv: 2312.16185 [q-fin.ST].

[205] Mariusz Maziarz. 'A review of the Granger-causality fallacy'. In: *Journal of Philosophical Economics* Volume VIII Issue 2.Articles (May 2015). DOI: 10.46298/jpe.10676.

[206] Patrick A. Stokes and Patrick L. Purdon. 'A study of problems encountered in Granger causality analysis from a neuroscience perspective'. In: *Proceedings of the National Academy of Sciences* 114.34 (2017), E7063–E7072. DOI: 10.1073/pnas.1704663114.

[207] Hiroshi Iyetomi. 'Collective Phenomena in Economic Systems'. In: *Evolutionary Economics and Social Complexity Science*. Springer Singapore, 2020, pp. 177–201. ISBN: 9789811548062. DOI: 10.1007/978-981-15-4806-2_9.

[208] Wataru Souma. 'Characteristics of Principal Components in Stock Price Correlation'. In: *Frontiers in Physics* 9 (Apr. 2021), p. 602944. DOI: 10.3389/fphy.2021.602944.

[209] Alp Kustepeli. 'On the Helmholtz Theorem and Its Generalization for Multi-Layers'. In: *Electromagnetics* 36.3 (2016), pp. 135–148. DOI: 10.1080/02726343.2016.1149755.

[210] Ole Peters and Alexander Adamou. *Ergodicity Economics Lecture Notes*. 2018.

[211] Ole Peters and Alexander Adamou. *Insurance makes wealth grow faster*. 2017. arXiv: 1507.04655 [q-fin.RM].

[212] Ole Peters and Benjamin Skjold. *LMLhub/AAS_editorial_figures_July2023: v1.0.0 initial release*. Version v1.0.0. code published on zenodo. June 2023. DOI: 10.5281/zenodo.7994190.

[213] Tobias Wand, Oliver Kamps and Benjamin Skjold. 'Cooperation in a non-ergodic world on a network - insurance and beyond'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 34.7 (July 2024), p. 073137. DOI: 10.1063/5.0212768.

[214] P. Allen. 'Self-organization in Economic Systems'. In: Edward Elgar Publishing, July 2007, pp. 1111–1148. ISBN: 9781843762539. DOI: 10.4337/9781847207012.00080.

[215] John Foster. 'Economics and the Self-Organisation Approach: Alfred Marshall Revisited?' In: *The Economic Journal* 103.419 (July 1993), pp. 975–991. DOI: 10.2307/2234714.

[216] Martin A. Nowak and Robert M. May. 'Evolutionary games and spatial chaos'. In: *Nature* 359.6398 (Oct. 1992), pp. 826–829. DOI: 10.1038/359826a0.

[217] Katarzyna Sznajd-Weron and Józef Sznajd. 'Opinion Evolution in Closed Community'. In: *International Journal of Modern Physics C* 11.06 (2000), pp. 1157–1165. DOI: 10.1142/S0129183100000936.

[218] Didier Sornette. 'Physics and financial economics (1776–2014): puzzles, Ising and agent-based models'. In: *Reports on Progress in Physics* 77.6 (May 2014), p. 062001. DOI: 10.1088/0034-4885/77/6/062001.

[219] Jérôme Garnier-Brun, Michael Benzaquen and Jean-Philippe Bouchaud. 'Unlearnable Games and "Satisficing" Decisions: A Simple Model for a Complex World'. In: *Physical Review X* 14.2 (June 2024). DOI: 10.1103/physrevx.14.021039.

[220] I. Ferri, A. Gaya-Àvila and A. Díaz-Guilera. 'Three-state opinion model with mobile agents'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 33.9 (Sept. 2023). DOI: 10.1063/5.0152674.

[221] Jean-Philippe Bouchaud. 'Crises and Collective Socio-Economic Phenomena: Simple Models and Challenges'. In: *Journal of Statistical Physics* 151.3–4 (Jan. 2013), pp. 567–606. DOI: 10.1007/s10955-012-0687-3.

[222] Talia Baravi, Ofer Feinerman and Oren Raz. 'Echo chambers in the Ising model and implications on the mean magnetization'. In: *Journal of Statistical Mechanics: Theory and Experiment* 2022.4 (Apr. 2022), p. 043402. DOI: 10.1088/1742-5468/ac5d42.

[223] Martin Drechsler. 'Ising models to study effects of risk aversion in socially interacting individuals'. In: *Physica A: Statistical Mechanics and its Applications* 632 (Dec. 2023), p. 129345. DOI: 10.1016/j.physa.2023.129345.

[224] Thomas C. Schelling. 'Dynamic models of segregation'. In: *The Journal of Mathematical Sociology* 1.2 (1971), pp. 143–186. DOI: 10.1080/0022250X.1971.9989794.

[225] Lloyd, David J.B. and O'Farrell, Hayley. 'On localised hotspots of an urban crime model'. In: *Physica D: Nonlinear Phenomena* 253 (June 2013), pp. 23–39. DOI: 10.1016/j.physd.2013.02.005.

[226] Pierre-André Chiappori et al. 'Asymmetric Information in Insurance: General Testable Implications'. In: *The RAND Journal of Economics* 37.4 (2006), pp. 783–798. ISSN: 07416261. URL: http://www.jstor.org/stable/25046274 (visited on 05/03/2024).

[227] Kenneth J Arrow. 'The theory of risk aversion'. In: *Essays in the theory of risk-bearing* (1971), pp. 90–120.

[228] Ole Peters and Alexander Adamou. 'The ergodicity solution of the cooperation puzzle'. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 380.2227 (May 2022). DOI: 10.1098/rsta.2020.0425.

[229] David Meder et al. 'Ergodicity-breaking reveals time optimal decision making in humans'. In: *PLOS Computational Biology* 17.9 (Sept. 2021), pp. 1–25. DOI: 10.1371/journal.pcbi.1009217.

[230] A. Vanhoyweghen and V. Ginis. 'Human decision-making in a non-ergodic additive environment'. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 479.2278 (Oct. 2023). DOI: 10.1098/rspa.2023.0544.

[231] Benjamin Skjold et al. 'Are risk preferences optimal?' In: *Preprint available at* (July 2023). DOI: 10.31219/osf.io/ew2sx.

[232] Benjamin Skjold and Ole Peters. *Insurance as an ergodicity problem*. https://ergodicityeconomics.com/2023/08/08/insurance-as-an-ergodicity-problem/. [Online; accessed 11/08/23]. 2023.

[233] Tobias Wand. *Non-Ergodic Cooperation on a Lattice Network - Simulation*. May 2024. DOI: 10.5281/zenodo.11382048.

[234] Pauli Virtanen et al. 'SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python'. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[235] Maitreesh Ghatak. 'Theories of Poverty Traps and Anti-Poverty Policies'. In: *The World Bank Economic Review* 29.suppl 1 (2015), S77–S105. DOI: 10.1093/wber/lhv021.

[236] Carles Boix. 'Origins and Persistence of Economic Inequality'. In: *Annual Review of Political Science* 13.1 (2010), pp. 489–516. DOI: 10.1146/annurev.polisci.12.031607.094915.

[237] Frank Stollmeier and Jan Nagler. 'Unfair and Anomalous Evolutionary Dynamics from Fluctuating Payoffs'. In: *Phys. Rev. Lett.* 120 (5 Feb. 2018), p. 058101. DOI: 10.1103/PhysRevLett.120.058101.

[238] Frank Stollmeier. 'Evolutionary Dynamics in Changing Environments'. PhD thesis. Georg-August-Universität Göttingen, 2018. DOI: 10.53846/goediss-6920.

[239] Lorenzo Fant et al. 'Stable cooperation emerges in stochastic multiplicative growth'. In: *Phys. Rev. E* 108 (1 July 2023), p. L012401. DOI: 10.1103/PhysRevE.108.L012401.

[240] Athena Aktipis et al. 'Cooperation in an Uncertain World: For the Maasai of East Africa, Need-Based Transfers Outperform Account-Keeping in Volatile Environments'. In: *Human Ecology* 44.3 (May 2016), pp. 353–364. DOI: 10.1007/s10745-016-9823-z.

[241] C. Athena Aktipis, Lee Cronk and Rolando de Aguiar. 'Risk-Pooling and Herd Survival: An Agent-Based Model of a Maasai Gift-Giving System'. In: *Human Ecology* 39.2 (Jan. 2011), pp. 131–140. DOI: 10.1007/s10745-010-9364-9.

[242] Gudrun Dahl and Anders Hjort. *Having herds: Pastoral herd growth and household economy*. Department of Social Anthropology, University of Stockholm, 1976.

[243] Martin A. Nowak. 'Five Rules for the Evolution of Cooperation'. In: *Science* 314.5805 (Dec. 2006), pp. 1560–1563. DOI: 10.1126/science.1133755.

[244] Benjamin Skjold and Ole Peters. Personal communication. 2023.

[245] Pavel Chvykov and Jerome Nikolai Warren. 'Is salary just a negative insurance premium?' Presentation given at Ergodicity Economics Conference. 2024.

[246] Robert E. Lucas. 'Asset Prices in an Exchange Economy'. In: *Econometrica* 46.6 (Nov. 1978), p. 1429. DOI: 10.2307/1913837.

[247] Oswald Spengler. *Der Untergang des Abendlandes*. Verlag Braumüller, 1918.

[248] Isaac Asimov. *Foundation*. Gnome Press, 1951.

[249] Lukas Wittmann and Christian Kuehn. 'The Demographic-Wealth model for cliodynamics'. In: *PLOS ONE* 19.4 (Apr. 2024), pp. 1–21. DOI: 10.1371/journal.pone.0298318.

[250] Peter Turchin et al. 'An Introduction to Seshat'. In: *Journal of Cognitive Historiography* 5.1 (2018). Number: 1-2, pp. 115–123. DOI: 10.1558/jch.39395.

[251] Pieter François et al. 'A Macroscope for Global History. Seshat Global History Databank: a methodological overview'. In: *Digital Humanities Quarterly* 10.4 (2016). http://www.digitalhumanities.org/dhq/vol/10/4/000272/000272.html (accessed on 04/11/2024).

[252] Peter Turchin et al. 'The Equinox2020 Seshat Data Release'. In: *Cliodynamics* 11.1 (2020).

[253] Jaeweon Shin et al. 'Scale and information-processing thresholds in Holocene social evolution'. In: *Nature Communications* 11.1 (May 2020), p. 2394. DOI: 10.1038/s41467-020-16035-9.

[254] Peter Turchin et al. 'Disentangling the evolutionary drivers of social complexity: A comprehensive test of hypotheses'. In: *Science Advances* 8.25 (2022), eabn3517. DOI: 10.1126/sciadv.abn3517.

[255] James S Bennett et al. *Cliopatria - A geospatial database of world-wide political entities from 3400BCE to 2024CE*. Preprint published on SocArXiv. 2024. DOI: 10.31235/osf.io/24wd6.

[256] Tobias Wand. 'A Bayesian Approach to Survivorship Bias in Historical Data Analysis'. In: *Cliodynamics: The Journal of Quantitative History and Cultural Evolution* 15.1 (2024). DOI: 10.21237/c7clio15163477.

[257] Marten Scheffer et al. 'The vulnerability of aging states: A survival analysis across premodern societies'. In: *Proceedings of the National Academy of Sciences* 120.48 (2023), e2218834120. DOI: 10.1073/pnas.2218834120.

[258] Peter Turchin. *End times*. New York: Penguin Press, 2023.

[259] Daniel Hoyer et al. 'Navigating Polycrisis: long-run socio-cultural factors shape response to changing climate'. In: *Philosophical Transactions of The Royal Society B Biological Sciences* 378 (Sept. 2023). DOI: 10.1098/rstb.2022.0402.

[260] Daniel Hoyer et al. *All Crises are Unhappy in their Own Way: The role of societal instability in shaping the past*. Preprint on SocArXiv. Feb. 2024. DOI: 10.31235/osf.io/rk4gd.

[261] David Carballo, Paul Roscoe and Gary Feinman. 'Cooperation and Collective Action in the Cultural Evolution of Complex Societies'. In: *Journal of Archaeological Method and Theory* 21.1 (2014), pp. 98–133. DOI: 10.1007/s10816-012-9147-2.

[262] Peter J. Richerson and Morten H. Christiansen, eds. *Cultural Evolution*. Society, Technology, Language, and Religion. Boston: MIT Press, 2013. 499 pp. ISBN: 978-0-262-01975-0.

[263] Peter Turchin and Sergey Gavrilets. 'Tempo and Mode in Cultural Macroevolution'. In: *Evolutionary Psychology* 19.4 (2021). DOI: 10.1177/14747049211066600.

[264] Patrick Manning et al. 'Collaborative Historical Information Analysis'. In: *Comprehensive Geographic Information Systems*. Ed. by Bo Huang. 3 vols. Amsterdam: Elsevier, 2017.

[265] V. G. Childe. *Man Makes Himself*. Watts & Company, 1936.

[266] L. A. White. *The Evolution of Culture*. McGraw-Hill, 1959.

[267] Service, Elman R. *Origins of the state and civilization*. New York, NY: WW Norton, Apr. 1975.

[268] Patrick Vinton Kirch. 'From chiefdom to archaic state: Hawai'i in comparative and historical context'. In: *How Chiefs Became Kings*. University of California Press, Dec. 2010, pp. 1–28.

[269] David Engels. 'Kulturmorphologie und Willensfreiheit'. In: *Der lange Schatten Oswald Spenglers*. Ed. by David Engels, Max Otte and Michael Thöndl. Lüdinghausen: Manuscriptum, 2018, pp. 79–102.

[270] David Engels. *Oswald Spengler - Werk, Deutung, Rezeption*. Stuttgart: Kohlhammer Verlag, 2021. ISBN: 978-3-170-37495-9.

[271] Oana Borcan, Ola Olsson and Louis Putterman. 'Transition to agriculture and first state presence'. A global analysis. In: *Explorations in Economic History* 82 (2021), p. 101404. DOI: 10.1016/j.eeh.2021.101404.

[272] Elliot J. Carr. 'Characteristic time scales for diffusion processes through layers and across interfaces'. In: *Physical Review E* 97.4 (Apr. 2018). DOI: 10.1103/physreve.97.042115.

[273] Eva-Maria Wartha, Markus Bösenhofer and Michael Harasek. 'Characteristic Chemical Time Scales for Reactive Flow Modeling'. In: *Combustion Science and Technology* 193.16 (2021), pp. 2807–2832. DOI: 10.1080/00102202.2020.1760257.

[274] Todd L Parsons and Tim Rogers. 'Dimension reduction for stochastic dynamical systems forced onto a manifold by large drift: a constructive approach with examples from theoretical biology'. In: *Journal of Physics A: Mathematical and Theoretical* 50.41 (Sept. 2017), p. 415601. DOI: 10.1088/1751-8121/aa86c7.

[275] Peter Turchin et al. *seshatdb (Equinox Packaged Data)*. Version v.1. Zenodo, June 2022. DOI: 10.5281/zenodo.6642230.

[276] Donald B Rubin. *Multiple imputation for nonresponse in surveys*. Vol. 81. John Wiley & Sons, 2004.

[277] Peter Turchin et al. 'Explaining the rise of moralizing religions: a test of competing hypotheses using the Seshat Databank'. In: *Religion, Brain & Behavior* (2022), pp. 1–28. DOI: 10.1080/2153599X.2022.2065345.

[278] Peter Turchin et al. 'An integrative approach to estimating productivity in past societies using Seshat: Global History Databank'. In: *The Holocene* 31.6 (Feb. 2021), pp. 1055–1065. DOI: 10.1177/0959683621994644.

[279] Tobias Wand and Daniel Hoyer. *Seshat: Equinox Release with Continuous Polities*. Zenodo, July 2023. DOI: 10.5281/zenodo.8120128.

[280] David W. Hosmer and Stanley Lemeshow. *Applied Logistic Regression*. John Wiley & Sons, Ltd, 2000.

[281] P.-F. Verhulst. 'Notice sur la loi que la population poursuit dans son accroissement'. In: *Corresp. Math. Phys.* 10 (1838), pp. 113–121.

[282] Henri. J.M. Claessen. 'The emergence of pristine states'. In: *Social Evolution & History* 15.1 (Jan. 2016), pp. 3–57.

[283] Charles S. Spencer. 'Territorial expansion and primary state formation'. In: *PNAS* 107.16 (2010), pp. 7119–7126. DOI: 10.1073/pnas.1002470107.

[284] Timothy A Kohler, Darcy Bird and David H Wolpert. 'Social Scale and Collective Computation: Does Information Processing Limit Rate of Growth in Scale?' In: *Journal of Social Computing* 3.1 (2022), pp. 1–17. DOI: 10.23919/JSC.2021.0020.

[285] Tobias Wand. 'Analysis of the Football Transfer Market Network'. In: *Journal of Statistical Physics* 187.3 (2022). DOI: 10.1007/s10955-022-02919-1.

[286] Tobias Wand. 'Entropy scaling of high resolution systems—Example from football'. In: *Physica A: Statistical Mechanics and its Applications* 600 (2022), p. 127536. DOI: 10.1016/j.physa.2022.127536.

[287] Tobias Wand. *Estimating polynomial extensions to the Geometric Brownian Motion*. Sept. 2024. DOI: 10.5281/zenodo.13744865.

[288] Will Landecker et al. 'Interpreting individual classifications of hierarchical networks'. In: *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. 2013, pp. 32–38. DOI: 10.1109/CIDM.2013.6597214.

[289] Avanti Shrikumar, Peyton Greenside and Anshul Kundaje. 'Learning Important Features Through Propagating Activation Differences'. In: *Proceedings of the 34th International Conference on Machine Learning*. Ed. by Doina Precup and Yee Whye Teh. Vol. 70. Proceedings of Machine Learning Research. PMLR, Aug. 2017, pp. 3145–3153. URL: https://proceedings.mlr.press/v70/shrikumar17a.html.

[290] Jianming Zhang et al. 'Top-Down Neural Attention by Excitation Backprop'. In: *Int. J. Comput. Vision* 126.10 (Oct. 2018), pp. 1084–1102. DOI: 10.1007/s11263-017-1059-x.

[291] Francois Chollet et al. *Keras*. https://github.com/fchollet/keras (accessed on 04/11/2024). 2015.

Lebenslauf

Not included in the published version.