

## FINAL YEAR PROJECT REPORT

NAME:	Sohyun Park
DEGREE COURSE:	MSci Physics
PROJECT TITLE:	Statistical Mechanics of Foreign Exchange Markets
YEAR OF SUBMISSION:	2024
SUPERVISOR:	Dr. Thomas Machon
NUMBER OF WORDS:	8296



## Declaration

We did not independently collect all the data presented in this project. The currency data were generously provided by our supervisor, Dr. Thomas Machon, who obtained them from Yahoo Finance. All data preprocessing, parameter optimisation, and visualisation tasks were performed using Python by me, including the creation of a Bokeh application to visualise the currency network. The estimator, including the pseudolikelihood maximisation and the formulation of the Hamiltonian with a hidden variable, was discussed and developed under the guidance of Dr. Machon. I was responsible for calculating all equations and writing all necessary code for the project, conducting extensive research to support these efforts, including understanding the analytical tools to quantify phase transitions. Meanwhile, my project partner contributed valuable economic insights that enriched our analysis and understanding of the results.

## Acknowledgements

I would first like to thank Dr. Thomas Machon. He has been a fantastic supervisor, and I am delighted that I will be seeing him in the corridors for the next four years. His deep insights into multifaceted disciplines and his never-ending stream of ideas have enabled me to successfully explore this project. He has supported any ideas I came up with and rigorously examined the validity of my methods.

The feedback I received from Dr. Francesco Turci was extremely helpful in the interim report discussion, which has guided my direction into researching criticality in the currency markets. Although he wasn't a supervisor, whenever I met him in the corridor, he always smiled and offered great advice, and I am really excited to start the PhD with him in the coming years.

Also, I would like to express my gratitude to my personal tutor, Prof. Malcolm Bremer. Having initially pursued physics to understand how the universe operates, and now residing in complexity science, he has always listened to and encouraged every path that I have taken. If you are reading this, *thank you very much.*

Next, my dear friend Jonathan Burbach, who solidified my decision to choose this project. Since I had not taken condensed matter physics and was not particularly interested in the finance world, I was concerned about my choice. However, your encouragement has beautifully shaped me to learn about these new exciting areas over the year.

Lastly, I would like to dedicate this paper to my parents. My dad, who must have been waiting so long for me to graduate and see my first paper, here is my little work. Also, my mum, who says she sleeps well whenever she reads what I write. Remember how I always had a lot of questions since childhood? Even if she could not answer them, she always stood by me, pondering together, "Hmm... I wonder why as well?" and encouraging me to come up with my own ideas. You've been the source of my diverse curiosities about the natural world.

## **Abstract**

Financial datasets are difficult to estimate due to their inherent heavy-tailed distribution. A solution to this problem is inverse Ising inference, also known as pairwise maximum entropy modelling. The correlation structure of foreign exchange markets was reconstructed using this machine learning technique. Additionally, data-driven methods including pseudolikelihood maximisation and the Metropolis algorithm were employed to explore two decades of time series, to identify transformative shifts corresponding to known geopolitical and financial events, and to infer criticality in the dynamic market structure. This may suggest a role for self-correcting feedback mechanisms in maintaining a market steady state.

# 1 Introduction

The study of complex systems, particularly those exhibiting collective behaviour arising from interactions among their constituent elements, has been a focal point across various scientific disciplines. The Ising model, a cornerstone in statistical mechanics, has been pivotal in exploring these phenomena. Initially introduced to model ferromagnetism [1], where spins on a lattice interact to result in a phase transition at a critical temperature, the model has since transcended its origins to understand diverse systems, ranging from neural to social networks [2, 3].

In the realm of economics, the Ising model provides a framework for modelling the binary choices of agents, akin to the binary spin states in physical systems. This analogy extends naturally to financial markets, where the myriad of transactions and the interplay between various financial instruments can be thought of as a complex system with agents' decisions influencing market movements [4].

The intricate web of correlations present in financial data poses a challenge for traditional analysis techniques, which often fall short in capturing the underlying dynamics governing market behaviour. The Ising model, with its simplicity in modelling interactions and its power to explain emergent properties, offers an alternative approach. By framing market dynamics as a series of pairwise interactions within a network, the Ising model enables the identification of critical points akin to phase transitions, providing insight into the conditions leading to market stability or volatility [5].

The project at hand leverages these theoretical foundations and computational techniques to dissect the foreign exchange (forex) markets, a quintessential example of a complex, adaptive system. By applying the Ising model to the forex markets, the study seeks to unveil the subtle intricacies of currency interactions, paving the way for novel interpretations and strategies in financial analysis.

Moreover, the versatility of the Ising model, especially in the realm of inverse problems, extends its applicability far beyond traditional physics. This expansion into data science highlights the model's potential for extracting meaningful patterns from complex datasets [6], a principle that is central to our study of global financial markets. Similarly, the adaptability of the model in drawing parallels between seemingly disparate systems like neural networks and financial markets provides a unique perspective on understanding the stochastic nature of financial interactions [7], further justifying the application of the Ising model in our study. The insights from these works not only underscore the methodological robustness of our approach but also situate our research within a broader, interdisciplinary scientific discourse.

## 2 Background

### 2.1 Inverse Ising Inference

Inverse problems arise from observing a set of data and seeking to determine the underlying causes. This is the inverse of a forward problem, which begins with the causes and then predicts the results based on the input parameters.

As a practical application of solving inverse problems, inverse Ising inference is a technique primarily employed in statistical physics, computational biology, and machine learning to infer the parameters of an Ising model based on observed data.

The Ising model describes systems of interacting spins on a lattice, where each spin can take on one of two binary values, typically +1 or -1. The model's elegance lies in its simplicity, encapsulating the essence of phase transitions through local interactions between adjacent spins and an external magnetic field [8]. Mathematically, the energy of a configuration  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$  in an Ising model is given by the Hamiltonian function:

$$H = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N J_{ij} \sigma_i \sigma_j - \sum_{i=1}^N h_i \sigma_i \quad (1)$$

where  $J_{ij}$  is the coupling constant describing the strength of *pairwise* interaction between spins  $\sigma_i$  and  $\sigma_j$  (the scaling factor of 1/2 is introduced to compensate for double counting), and  $h_i$  is the external field acting on each spin  $\sigma_i$ . The goal of inverse Ising inference is to determine the values of parameters  $J_{ij}$  and  $h_i$  that best represent the observed data.

This inverse problem presents a computational challenge due to the large number of possible configurations ( $2^N$  for  $N$  spins), making it difficult to directly optimise the likelihood function through traditional Monte Carlo sampling. As a result, approximate estimators, such as pseudolikelihood maximisation (ordered logistic regression), mean-field approximation [9], and faster Monte Carlo methods [10], have been developed to infer these parameters.

The successful application of these methods not only opens the door to a deeper understanding of the structure of forex markets but also to potentially predicting critical transitions, contributing to a more robust financial analysis. Furthermore, in computational neuroscience, these methods aid in the modelling of neural dynamics through energy landscapes, revealing the collective behaviour of neurons and their implications for understanding altered cognitive functions [11]. In social network analysis, this approach illustrates the necessity for the adaptation of influence maximisation, treating individual opinions as spins, highlighting the importance of optimising the Ising network structure for marketing strategies [12].

### 2.2 Foreign Exchange Markets

The forex markets, the world's most voluminous and liquid financial markets, operate around the clock during business days, facilitating global currency trades. Participants in these markets exchange currencies, thereby determining the relative values of currencies, such as the Euro to Pound Sterling (EUR/GBP). These values are further influenced by a multitude of factors including economic indicators, news, and actions of large financial institutions.

In this dynamic environment, currency values exhibit inter-correlations due to shared influences affecting the market or entire sectors. For example, EUR/USD and GBP/USD are positively correlated currency pairs, often moving in tandem due to common market drivers, particularly shifts in the US dollar's strength.

Such changes might be prompted by economic announcements in the United States, impacting both the Euro and Pound simultaneously. While statistical measures like the linear correlation can quantify these relationships, a more sophisticated analysis can reveal the underlying network of connections, accounting for the complex, multi-layered interactions that define currency movements.

Drawing parallels from condensed matter physics, consider a one-dimensional system of spins, interacting only with their nearest neighbours. Figure 1 depicts how direct interactions between adjacent currency pairs can create a series of correlations, leading to a spillover effect where distant pairs exhibit indirect correlations. For instance, if A and B are energetically connected, and B and C are as well, then A indirectly influences C through B. We then measure the correlation between A and E even though they are not directly correlated. This chain of interactions implies that the observed statistical correlations among these independent components cannot accurately reflect direct relationships due to intermediary influences.

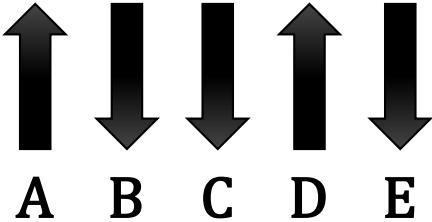


Figure 1: A simplified network of spins.

Our study employs inverse Ising inference to explore the network of correlations between various currencies and assess their temporal evolution. By fitting the Ising model to historical forex data, we aim to correct for this spillover effect and elucidate the true structure of currency relationships, providing insights into the underlying connections shaping global financial markets.

### 2.3 From Spins to Currencies

To adapt the Ising model for forex datasets, each *spin* can be represented as the logarithmic change in currency values from one day to the next, termed *daily log returns*. This approach standardises datasets, effectively normalising the diverse scales of global currency movements and allowing for a consistent comparison of their relative changes. For instance, as of 4 March 2024, significant differences in currency scales are observed, such as USD/GBP moving from 0.7609 to 0.8469 and USD/KRW from 1,257 to 1,365, as reported by Yahoo Finance. These variations, coupled with the disparities in their values, demonstrate the critical need for using log returns.

When transitioning to statistical modelling of the distribution of spins, at first sight it may seem feasible to employ a Gaussian process to model currency fluctuations as continuous variables, given its analytical tractability [13]. However, this continuum model is often based on normality assumptions from the central limit theorem, which fails to accurately represent the reality of the financial market where exchange rates are dominated by large fluctuations. The complexity of financial datasets is

characterised by their heavy-tailed distributions, even after logarithmic transformation, as displayed in Figure 2. This deficiency highlights the insufficiency of Gaussian behaviours in approximating parameters that govern critical market phases, as they fail to account for the significant outliers that drive market trends.

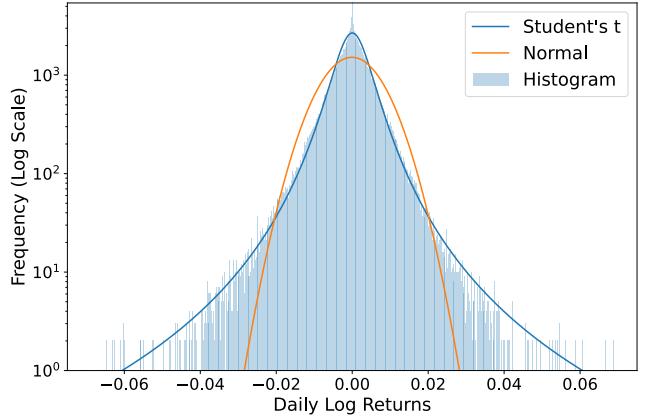


Figure 2: Histogram of daily log returns for all 26 currencies analysed in our study, which are roughly captured by the fat-tailed t-distribution, compared against a fitted normal distribution for benchmarking.

In response to these challenges, the Ising model offers a compelling alternative. By thresholding daily log returns to binary variables, this model sidesteps the traditional reliance on normality and the constraints of linear regression in the face of the standard error of the mean, which is proportional to the pronounced variances in currency fluctuations. In contrast, the variance of binary variables (+1 or -1, if equally distributed) simplifies to just 1.

To model the distribution of binarised currency fluctuations, which reproduces the observed correlation structure, we follow the maximum-entropy principle [14], which advocates selecting the probability distribution that maximises entropy under given constraints. This approach ensures the least biased estimation possible based on available data, avoiding unnecessary assumptions about the system's unknowns. Specifically, through pairwise maximum entropy modelling (MEM), an application of this principle, the most probable states of binary variables are determined by maximising entropy subject to constraints on observed means and pairwise correlations.

This framework was rigorously examined through its application to neuroscience for inferring network correlations in retinal neurons [2], attesting to the MEM's expansive applicability. In their seminal work, Schneidman et al. suggested that observed collective behaviours in complex networks are predominantly governed by pairwise correlations. This methodology justifies our analysis with inverse Ising inference, recognising it as MEM that best captures the observed pairwise correlations within the currency network. To articulate this approach, the statistical model which can reproduce the average rate of change for each sample of currencies  $\mathbf{x} = \{x_i\}$  introduces a set of Lagrange multipliers  $\mathbf{h} = \{h_i\}$ :

$$H(\mathbf{x}; \mathbf{h}) = - \sum_{i=1}^d h_i x_i$$

This energy function accounts for the independent statistical behaviour of each currency's increase or decrease, where  $\mathbf{h}$  influences the tendency of spins to align in a particular direction. By augmenting this model with pairwise Lagrange multipliers  $\mathbf{J} = \{J_{ij}\}$ , we encapsulate correlations between pairs of currencies, signifying that an increase or decrease in one currency tends to align with another's:

$$H(\mathbf{x}; \mathbf{h}, \mathbf{J}) = -\sum_{i=1}^d h_i x_i - \frac{1}{2} \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_i x_j \quad (2)$$

where we assume  $J_{ij} = J_{ji}$  and  $J_{ii} = 0$ . The symmetry of  $\mathbf{J}$  matrix implies that the influence between any two currencies is mutual and of equal magnitude in each direction. Setting the diagonal elements of  $\mathbf{J}$  to 0 indicates there is no self-interaction. This form is the Ising model (see Eq. 1). Higher-order correlations are unconstrained by the multipliers.

This context sets the stage for modelling the forex markets as a lattice, consisting of interconnected currencies whose correlated movements can be modelled by the coupling  $J_{ij}$  in Eq. 2. The external field  $h_i$  in financial context corresponds to the impact of broader economic factors influencing the sentiment toward an individual currency.

The critical insight provided by the Ising model is the emergence of collective behaviour from simple rules of interaction. This manifests as global market trends emerging from the aggregate of individual currency fluctuations, with transaction volumes exhibiting scale-free behaviour following market switches, indicating a collective response to changing conditions [15]. Identifying phases akin to ferromagnetic or paramagnetic states can signal periods of market stability or volatility, offering a predictive edge in financial analysis.

## 2.4 Computational Strategies

In pursuing inverse Ising inference, our primary objective is to efficiently and accurately estimate a set of unknown parameters of the Ising model, which reproduces the correlation structure of *foreign exchange markets*. The probability of observing a binary set of currencies  $\mathbf{x}$  is given by the Boltzmann distribution from *statistical mechanics*, which represents the maximum entropy distribution, conditioned on average energy and pairwise correlations:

$$P(\mathbf{x}; \mathbf{h}, \mathbf{J}) = \exp \left( \sum_{i=1}^d h_i x_i + \frac{1}{2} \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_i x_j \right) / Z \quad (3)$$

where the inverse temperature  $\beta$  is subsumed into parameters  $h_i$  and  $J_{ij}$ , and  $Z$  is the partition function that normalises the probability distribution. The set of parameters  $\{\mathbf{h}, \mathbf{J}\}$  which maximises the likelihood of observing all the data is the solution to this inverse problem.

However, accurately computing  $Z$  for maximum likelihood estimation (MLE) typically requires extensive Monte Carlo sampling. This process involves summing the exponential terms over all possible states of the system and evaluating their corresponding probabilities. As the number of variables increases, scalability becomes a significant

challenge, rendering the computation intractably complex. This is particularly pertinent for the 26 currencies in our study, corresponding to  $2^{26} \approx 67$  million terms at each observation. Consequently, the computational intensity and complexity inherent in these problems necessitate the consideration of an alternative strategy.

Pseudolikelihood maximisation (PLM) successfully addresses these challenges [16]. It reformulates the likelihood as the product of conditional probabilities, circumventing the direct computation of the partition function, thereby simplifying the computational process. The limitations of MLE and the derivation of this specific PLM formulation are detailed in Appendix 7.1.

This approach is encapsulated in the following objective function, where the aim is to maximise this expression of pseudolikelihood to estimate the model parameters:

$$\begin{aligned} & \ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) \\ & \approx \sum_{t=1}^N \sum_{i=1}^d -\ln \left( 1 + e^{-2x_i(t)[h_i + \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_j(t)]} \right) \end{aligned} \quad (4)$$

where  $\mathbf{X}$  represents all the data (binary matrix), while the parameter vector  $\boldsymbol{\theta}$  includes  $\mathbf{J}$  and  $\mathbf{h}$ , representing the upper triangular elements of the symmetric coupling matrix and external field vector, respectively. The summation over  $t$  spans different observational time points, and the summation over  $i$  encompasses the various currencies at those points.  $x_i(t)$  denotes the binary state of each currency at time  $t$  (e.g., a thresholded value: +1 for an increase or -1 for a decrease), indicating the logarithmic change between two successive days.

In conclusion, the strategic utilisation of PLM in our study significantly enhances our capability to decode the complex network of interactions in financial markets. This method efficiently handles the intricacies of correlation, correcting the spillover effect while providing a balance between computational feasibility and accuracy.

## 3 Experimental Methods

### 3.1 Data Preprocessing

In this phase, we transformed historical forex data from multiple currencies (see Table 1) into a binary format suitable for Ising model analysis. Initially, we standardised the dataset by aligning start dates, ensuring consistency for uniform analysis. Subsequently, we corrected potential discrepancies due to trading gaps or consecutive identical values<sup>1</sup> using linear interpolation and normalised the data by calculating daily log returns.

In outlier detection, there is no unified methodology; even if simply using the intrinsic properties of the dataset and filtering beyond a threshold might seem feasible. From an informational perspective, a contextual understanding of

<sup>1</sup>To circumvent consecutive identical values, which when logarithmically transformed result in 0 and thus inhibit thresholding to binary values, we selected forex data represented with more than 1 unit. For example, USD/IDR close prices were 15516, 15622, 15591, and 15631, whereas IDR/USD were 0.000064, 0.000064, 0.000064, and 0.000064 between 2023-10-03 and 2023-10-06. This approach ensures a more meaningful differentiation in daily log returns, facilitating effective binary conversion.

Date	AUD	EUR	GBP	JPY	NZD	CAD	CHF	CNY	CZK	...	SEK	SGD	THB	TRY	ZAR
2003-01-01	NaN	NaN	NaN	0.008	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
2003-01-02	NaN	NaN	NaN	0.008	NaN	NaN	0.121	NaN	...	0.113	NaN	NaN	NaN	NaN	NaN
2003-01-03	NaN	NaN	NaN	0.008	NaN	NaN	0.121	NaN	...	0.115	NaN	NaN	NaN	NaN	NaN
2003-01-06	NaN	NaN	NaN	0.008	NaN	NaN	0.121	NaN	...	0.115	NaN	NaN	NaN	NaN	NaN
2003-01-07	NaN	NaN	NaN	0.008	NaN	NaN	0.121	NaN	...	0.115	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2023-10-03	0.636	1.048	1.209	0.007	0.594	0.731	1.089	0.139	0.043	...	0.090	0.728	0.027	0.036	0.052
2023-10-04	0.631	1.047	1.208	0.007	0.591	0.729	1.086	0.139	0.043	...	0.090	0.728	0.027	0.036	0.052
2023-10-05	0.633	1.051	1.214	0.007	0.592	0.728	1.091	0.139	0.043	...	0.090	0.729	0.027	0.036	0.052
2023-10-06	0.637	1.055	1.219	0.007	0.597	0.730	1.096	0.139	0.043	...	0.091	0.731	0.027	0.036	0.051
2023-10-09	0.636	1.054	1.218	0.007	0.598	NaN	NaN	0.137	NaN	...	NaN	NaN	0.027	NaN	NaN

Descriptive Statistics															
Count	4528	5152	5164	5390	5153	5219	5217	5242	5149	...	5237	5163	5151	4880	5165
Mean	0.816	1.239	1.543	0.009	0.708	0.849	0.988	0.146	0.047	...	0.129	0.719	0.030	0.435	0.104
Median	0.772	1.227	1.545	0.009	0.702	0.810	1.017	0.147	0.046	...	0.128	0.732	0.030	0.463	0.097
Std	0.135	0.126	0.238	0.001	0.072	0.098	0.113	0.013	0.006	...	0.019	0.061	0.003	0.249	0.036

Table 1: Raw historical forex data from Yahoo Finance, showcasing daily closing values for various currencies over two decades. ‘NaN’ entries indicate days when trading data were unavailable. All values are denominated in USD, focusing on major freely floating currencies and excluding those under heavy governmental control.

large fluctuations provides additional information. Thus, we annotated significant financial events on the interactive plots (see Figure 4) to incorporate these events into our study, offering a more comprehensive approach to this process. These annotations covered economic crises, policy shifts, and other regional macroeconomic factors that have historically impacted currency volatility.

Although not included in this report, analyses of trends (identifying long-term movements), seasonality (repeating patterns or cycles at regular intervals), and volatility clustering (observing periods with high fluctuations) were also performed to reveal those regional historical events.

Subsequently, outlier detection and cleansing were executed using the median absolute deviation (MAD), a robust measure of spread superior to the standard deviation.

This methodical filtration process ensures that the final dataset reflects genuine market anomalies rather than mere reactions to major economic events, thereby purifying and enriching it with historical context.

Table 2 presents the preprocessed daily log returns for all 26 currencies analysed, each having exactly the same amount of data, yielding a total of  $4538 \times 26 = 117988$  data points. This implies that errors on the inferred parameters for different nodes within the network will be statistically the same.

The final step in preprocessing was to encode this refined dataset into a binary matrix, termed *trajectory*, with elements being thresholded values (+1 or -1) of daily log returns, hence conforming to the structural requirements of the Ising model.

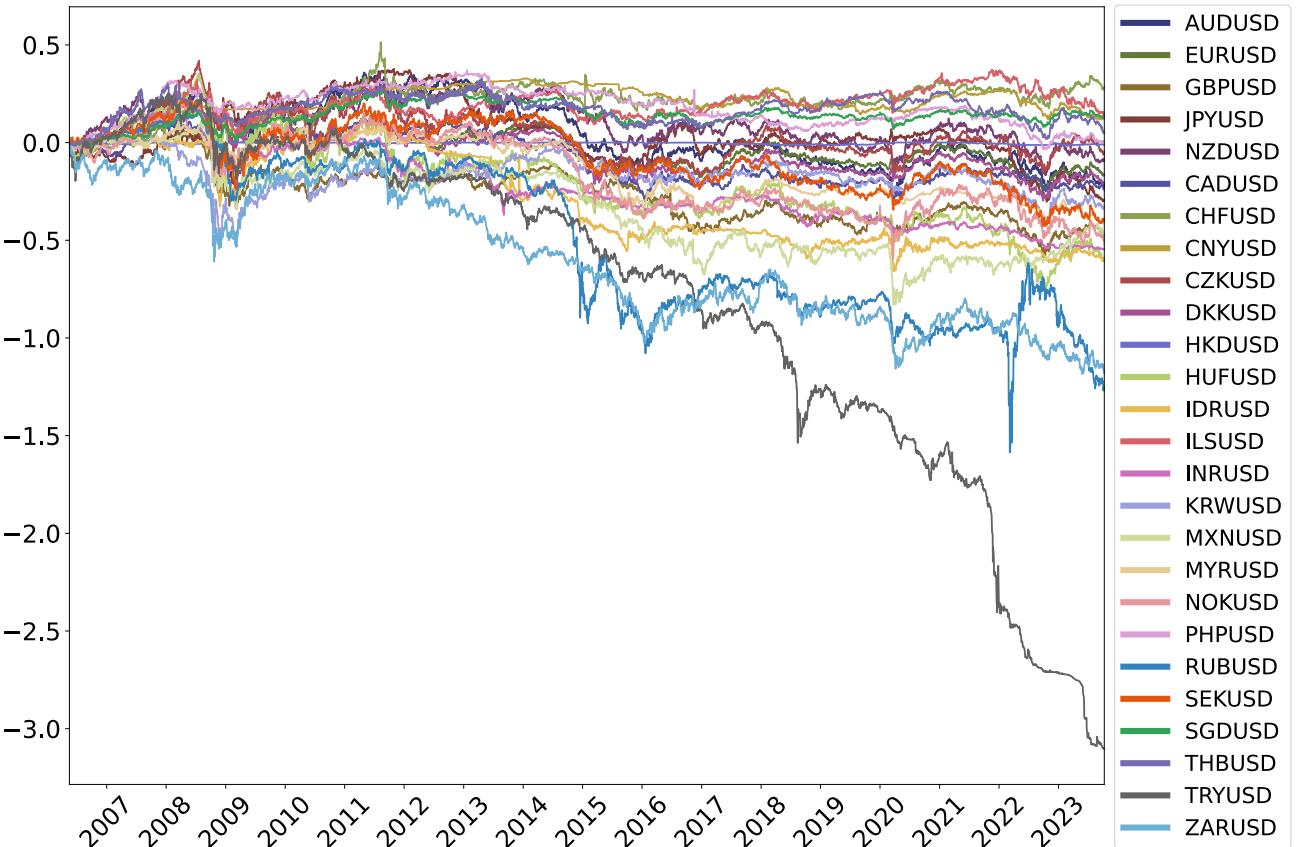


Figure 3: The preprocessed dataset presented as a cumulative sum of daily log returns for 26 currencies.

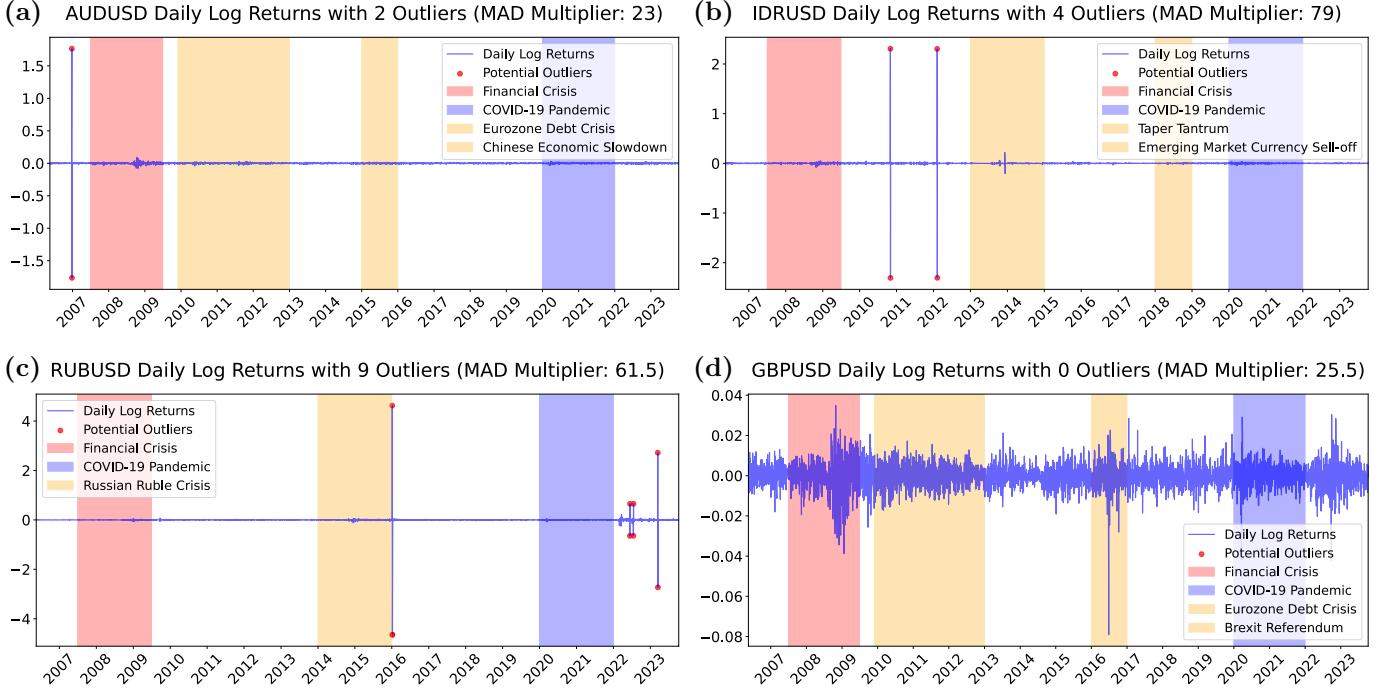
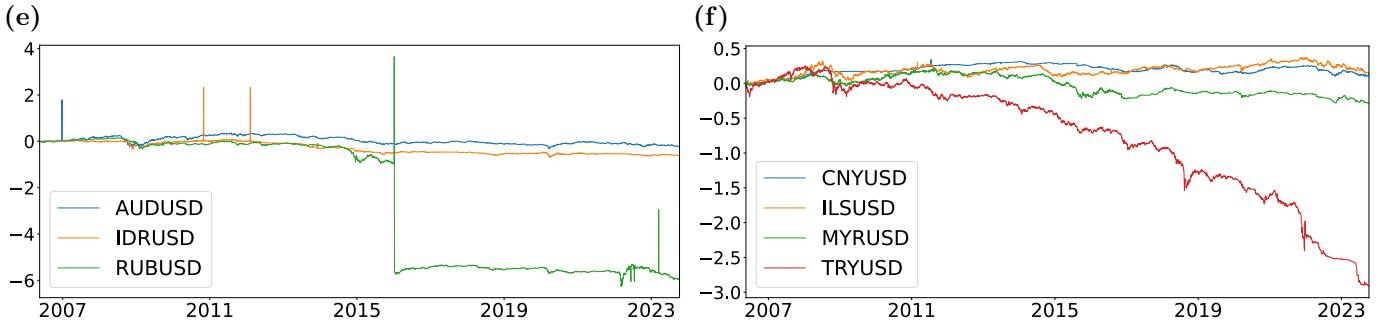


Figure 4: This composite plot shows outliers (placed outside the annotated periods) with characteristic of substantial spikes, which can be identified and filtered. An interactive plot is available in [17], which offers a slider for outlier identification amidst major economic events, detailed with annotations from various governmental sources. Obvious outlier examples are (a) AUD/USD and (b) IDR/USD, with IDR/USD requiring an elevated MAD (Median Absolute Deviation) due to Taper Tantrum volatility. (c) Post-February 24, 2022, RUB/USD analysis reflects the complexities due to geopolitical events, necessitating cautious outlier filtering. In contrast, (d) GBP/USD showcases no signs of outliers even amidst notable drops around 2016-2017, attributed to Brexit. This outlier detection task is inherently complex, demanding a meticulous, side-by-side examination of all currency charts, as depicted in the cumulative sums of daily log returns in (e) for prominent outliers and (f) for minor outliers, prior to filtering.



Date	AUD	EUR	GBP	JPY	NZD	CAD	CHF	CNY	CZK	...	SGD	THB	TRY	ZAR
2006-05-17	-0.010	-0.008	-0.003	-0.010	-0.005	-0.007	-0.006	2e-04	-0.016	...	-0.004	-0.007	-0.021	-0.016
2006-05-18	0.006	0.009	0.006	9e-05	-0.001	-0.004	0.008	-0.001	0.009	...	0.001	0.005	-0.014	-0.002
2006-05-19	-0.009	-0.007	-0.007	-0.007	-5e-04	0.002	-0.010	-0.001	-0.003	...	-0.003	-0.002	0.013	-0.005
2006-05-22	-0.005	0.007	0.003	0.002	0.002	6e-04	0.010	-8e-04	0.004	...	-8e-04	-0.004	-0.038	-0.007
2006-05-23	-0.003	-0.005	-0.004	-0.006	0.002	-0.007	-0.006	2e-04	-0.002	...	-2e-04	-4e-04	0.029	-0.015
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2023-10-03	-0.012	-0.008	-0.008	-0.002	-0.009	-0.008	-0.003	-0.002	-0.010	...	-0.005	-0.009	7e-04	-0.017
2023-10-04	-0.008	-8e-04	-0.001	0.004	-0.005	-0.002	-0.003	-0.002	-0.005	...	-7e-05	-0.002	-0.002	-0.004
2023-10-05	0.003	0.003	0.005	0.002	9e-04	-0.003	0.005	-8e-04	0.009	...	0.001	0.005	-10e-04	-0.001
2023-10-06	0.007	0.004	0.004	0.003	0.008	0.003	0.004	4e-04	0.002	...	0.003	-5e-04	-0.001	-0.010
2023-10-09	-0.002	-2e-04	-8e-04	-0.005	0.001	0.003	0.004	-0.013	0.002	...	0.003	-0.004	-0.001	-0.010

#### Descriptive Statistics

Count	4538	4538	4538	4538	4538	4538	4538	4538	4538	...	4538	4538	4538	4538
Mean	-5e-05	-4e-05	-1e-04	-6e-05	-2e-05	-5e-05	6e-05	3e-05	2e-06	...	4e-05	1e-05	-6e-04	-3e-04
Median	2e-04	2e-05	-10e-05	-2e-04	2e-04	-2e-05	-9e-05	2e-05	9e-05	...	6e-05	-6e-05	-3e-04	2e-04
Std	0.008	0.007	0.006	0.008	0.008	0.005	0.007	0.002	0.007	...	0.003	0.007	0.011	0.011

Table 2: Daily log returns for cleansed historical forex dataset, synchronised and processed from the raw dataset in Table 1, prepared before thresholding values to +1 or -1 for binary matrix formation. The balanced dataset is prepared with an identical count of observations for each currency from May 17, 2006, to October 9, 2023. All values are denominated in USD, featuring 4538 rows and 26 columns. Figures 2 and 3 were derived from this dataset.

### 3.2 Parameter Optimisation

This crucial stage identifies the optimal coupling matrix  $\mathbf{J}$  and external field vector  $\mathbf{h}$ , integral to modelling the intricate correlation structure of currency fluctuations. The process begins with the evaluation of the concavity of the pseudolikelihood function, spanning 351 dimensions. This includes 26 elements in  $\mathbf{h}$  and 325 distinct upper triangular elements in  $\mathbf{J}$  (since  $325 = 26(26 - 1)/2$ ) within an equilibrium setting where couplings between entities are symmetrical. This assessment was facilitated by examining the Hessian matrix of the pseudolikelihood to confirm it is negative-definite, thereby ensuring the uniqueness of model parameters regardless of initial conditions.

Using the vectorised expression of Eq. 5, the code efficiently maximises the likelihood alongside its gradients (Eqs. 6, 7) using a quasi-Newton algorithm.

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = \sum -\ln \left( 1 + e^{-2\mathbf{x} \odot (\mathbf{X} \cdot \mathbf{J} + \mathbf{h})} \right) \quad (5)$$

In this equation, the Hadamard product (denoted by  $\odot$ ) ensures element-wise multiplication, and  $\mathbf{h}$  is broadcast across all samples,  $\mathbf{X} = \{\mathbf{x}(t)\}$ , where each sample,  $\mathbf{x}(t)$ , represents the daily state of binary transitions at time  $t$ . Note that the likelihood obtained by the maximiser is completely blind to the order of samples within the trajectory,  $\mathbf{X}$ . This means that the random permutation of samples does not change the inferred parameters, implying that temporal correlation between samples *cannot* be inferred within this framework.

The gradient of the pseudolikelihood with respect to the external field vector  $\mathbf{h}$  is given by:

$$\nabla_{\mathbf{h}} \ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = \sum_{t=1}^N \frac{2\mathbf{x}(t)}{1 + e^{2\mathbf{x}(t) \odot (\mathbf{x}(t) \cdot \mathbf{J} + \mathbf{h})}} \quad (6)$$

Similarly, the gradient with respect to the coupling matrix  $\mathbf{J}$  is:

$$\nabla_{\mathbf{J}} \ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = (\mathbf{S}_{\mathbf{J}}^T \cdot \mathbf{X} + \mathbf{X}^T \cdot \mathbf{S}_{\mathbf{J}}) \odot (\mathbf{1} - \mathbf{I}) \quad (7)$$

where  $T$  denotes the transpose of the matrix to perform a symmetrising update of  $\mathbf{J}$ , and

$$\mathbf{S}_{\mathbf{J}} = \frac{2\mathbf{X}}{1 + e^{2\mathbf{x} \odot (\mathbf{X} \cdot \mathbf{J} + \mathbf{h})}}.$$

Here, the Hadamard product, combined with the mask matrix  $(\mathbf{1} - \mathbf{I})$ , which contains ones everywhere except for the diagonal, is critical for preserving the symmetry of  $\mathbf{J}$  and effectively zeroing out the diagonal.

These formulations significantly streamline computation by utilising matrix operations, enabling the processing of extensive datasets. During multivariate optimisation, we utilised the L-BFGS-B algorithm, favoured for its efficiency with large-scale problems and memory conservation [18], to navigate this high-dimensional complexity and balance computational speed and precision.

These methodical strategies facilitated the reliable estimation of parameters specifying the structure of the model across the complete dataset in less than 0.1 seconds, as illustrated in Figure 6. The estimation of these parameters across a set of timeframes allows for a detailed analysis of the temporal evolution of the market.

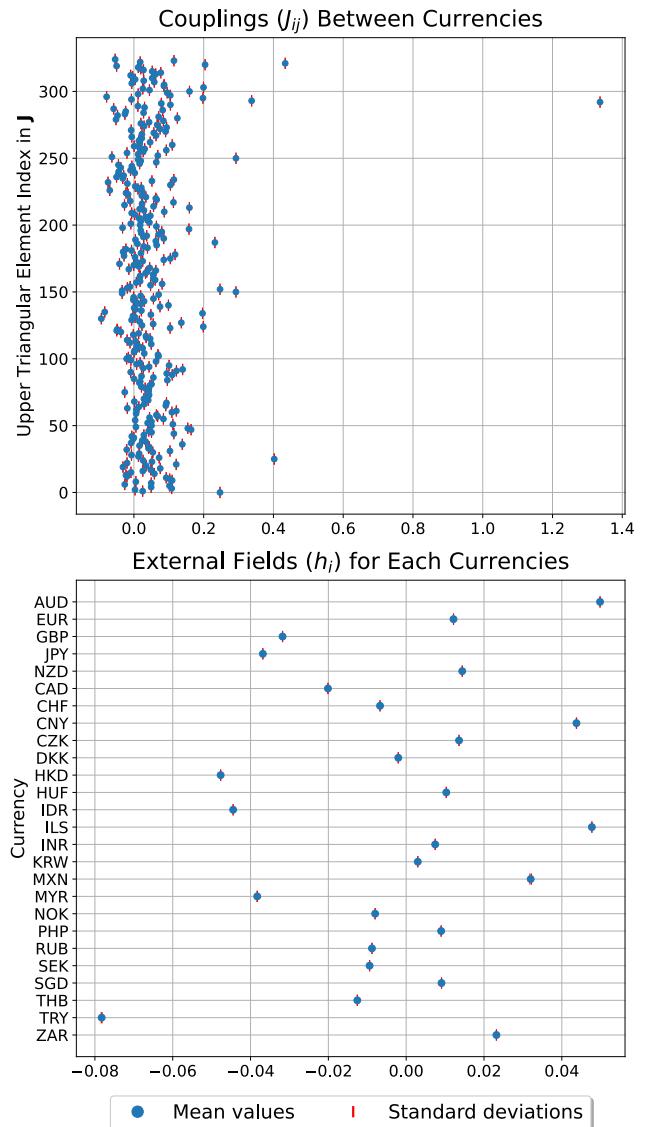


Figure 6: Scatter plots of optimised parameters from 4538 samples over 1,000 runs (execution time  $\approx 30s$ ). The standard deviations are minute, with the mean standard deviation being approximately 7.4e-06.

### 3.3 Visualisation

Our visualisation strategy harnessed the optimised parameters to forge compelling visual representations. This approach not only illuminated the complex interplay within the currency network but also facilitated a deeper dive into the temporal evolution of market dynamics spanning from 2006-05-17 to 2023-10-09.

These graphical insights effectively illustrate these market mechanics, and allow for the hypothesis testing and research directions we explore in the Results section. These visualisations elucidate the multifaceted relationships between currencies, as conceptualised within the Ising model framework<sup>2</sup>.

<sup>2</sup>For those interested in exploring the code behind the suite of these visual tools, as well as the entire pipeline of our methodologies, it is available in my GitHub repository [17]. This offers a comprehensive resource for further exploration and replication of our analytical approach.

## 4 Results

### 4.1 Currency Network

To visualise the inferred  $J$  matrix, a Bokeh application was developed for intuitive analysis of the intricate connectivities between currency nodes within the currency network (available on this website [19]). This application enables users to interactively adjust a threshold, filtering out lower-value couplings to explore the ranking of betweenness centrality of the currencies, and the network’s topological structure<sup>3</sup>. The adoption of absolute values for the couplings in this analysis aims to focus on weighted degree centrality, a method that emphasises the significance of both strong and weak ties within the network [21].

Figure 7 showcases a complex web of interactions, where currencies are largely clustered with geopolitical affiliations, such as Free Trade Areas (FTAs). As observed in Figure 8, these currency alignments do not necessarily correspond to geographical proximity.

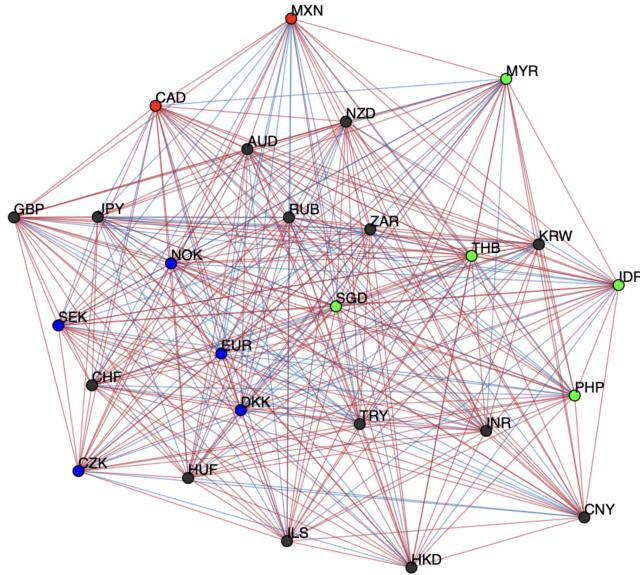


Figure 7: A fully connected graph at zero threshold. Each node represents a currency with the suffix USD. Red and blue edges indicate positive (alignment) and negative (opposition) couplings, respectively. Below, FTA groups and the top three currencies, ranked by betweenness centrality, are presented.

FTA	Colour
NAFTA	Red
EEA	Blue
ASEAN	Green
Others	Black
Currency	Betweenness Centrality Rank
CZK/USD	0.420
MXN/USD	0.240
THB/USD	0.220

<sup>3</sup>Within this application, users can explore various network properties. By simply hovering the mouse cursor over each currency, it displays the betweenness and connected weights between currencies, enabling a dynamic examination of the network. Each nodes are positioned using Fruchterman-Reingold force-directed algorithm [20].

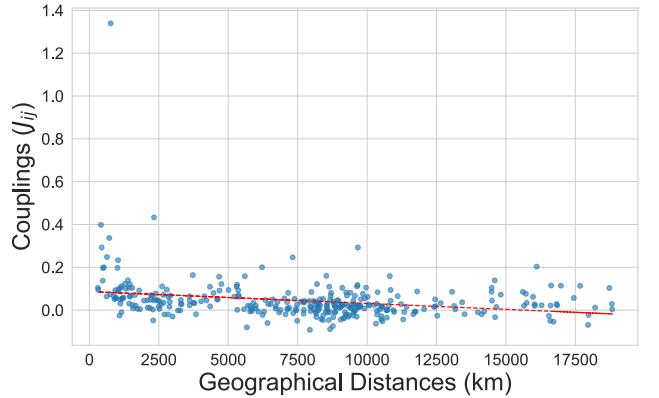


Figure 8: The correlation between couplings and geographical distances is weak, indicated by a correlation coefficient of -0.25. Reference locations were chosen to be the capitals of respective countries, with the exception of the euro, for which Brussels was used.

Increasing the threshold of  $J$  in this application reveals the central hub of each cluster, which can be further substantiated by examining various economic factors that reinforce its centrality. For instance, Figure 9 displays the evolution of the network when low-value couplings are excluded, highlighting significant interactions among currencies within the European Economic Area (EEA) and shedding light on their interdependence.

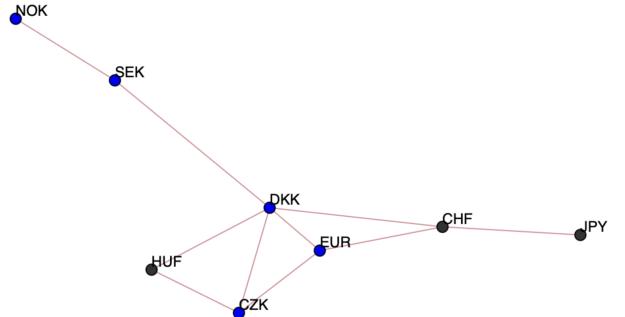


Figure 9: Network at a threshold value of  $J_{ij} \approx 0.16$ .

Currency	Betweenness Centrality Rank
DKK/USD	0.714
CHF/USD	0.476
SEK/USD	0.286

Surprisingly, at this threshold, EUR/USD exhibits zero centrality. This observation suggests that within the structure of this particular network and the distribution of weights, the euro does not act as a bridge on the shortest path between pairs of nodes—contrary to the expectations for a higher betweenness centrality score.

This phenomenon occurs in a weighted network where shortest paths are determined by the sum of edge weights. After the threshold is applied, ‘shorter’ alternative paths may emerge due to their lesser cumulative weights, thereby bypassing the euro. This finding warrants a cautious interpretation of the network’s properties, as it depends on its particular topology. For instance, the euro’s centrality is recovered when DKK/USD is excluded during optimisation and the network is reconstructed (see Figure 10).

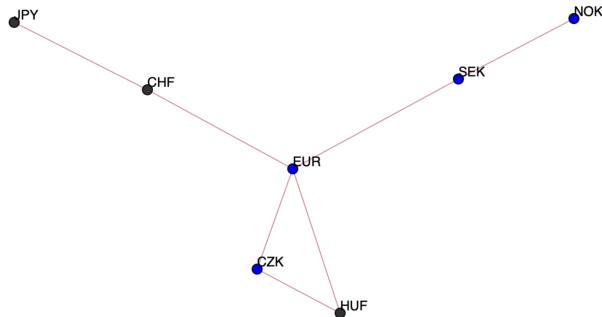


Figure 10: Network excluding the DKK/USD.

Currency	Betweenness Centrality Rank
EUR/USD	0.800
CHF/USD	0.333
SEK/USD	0.333

In Figure 11, maximising the threshold reveals the euro and Danish krone, elucidating the observed discrepancy in the centrality of the euro.

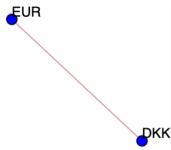


Figure 11: Network at a threshold of  $J_{ij} \approx 1.34$ , where the strength of their link is evident in the remarkable similarity of their charts, as seen in Figure 12.

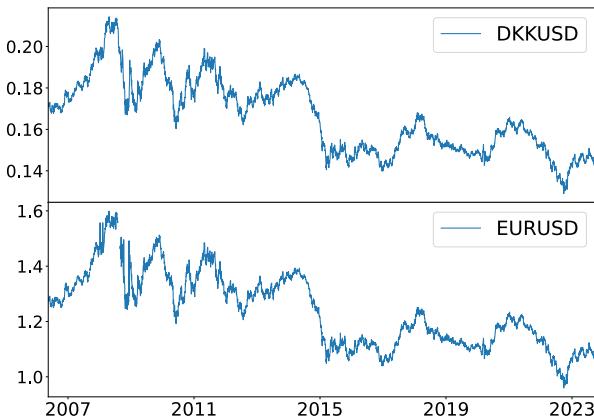


Figure 12: Comparison of EUR/USD and DKK/USD time series plots with their correlation coefficient at 0.95.

The synchronous movement of these currencies traces back to the European Exchange Rate Mechanism (ERM II), which tightly binds the Danish krone to the euro, enforcing a narrow fluctuation band and intertwining their economic policies and market behaviours [22].

With the exchange rate's fixed margin within  $\pm 2.25\%$  against the central rate of EUR 1 = DKK 7.46038 [23], the Danish krone effectively shadowed the euro's centrality in the network, due to their tight economic and exchange rate linkage, redistributing influence and obscuring the euro's sole centrality.

These outlined connectivity networks successfully prove the power of our methodology in reconstructing the correlation structure within the foreign exchange markets.

## 4.2 External Fields

Figure 13 delineates the inferred external fields vector  $\mathbf{h}$  for the 26 surveyed currencies, reflecting their overall long-term average response to the US dollar.

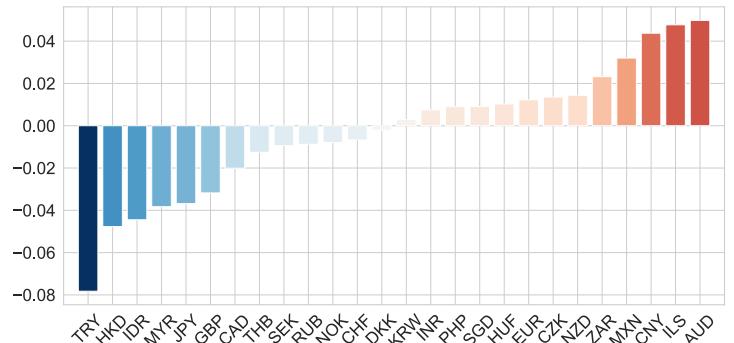


Figure 13: Bar chart of external fields, with a mean ( $\mu$ ) of  $-0.0029$  and a standard deviation ( $\sigma$ ) of  $0.0315$ .

Interestingly, despite the euro and Danish krone exhibiting nearly identical patterns, their external field values diverge ( $\text{DKK/USD} \approx -0.00204$  and  $\text{EUR/USD} \approx 0.0122$ ). In preprocessing stage, no significant outliers were identified to adjust their logarithmic transformations. However, a comparative analysis in Figure 14 sheds light on why the Danish krone's external field is lower than the euro's.



Figure 14: The cumulative sum of daily log returns for EUR/USD and DKK/USD, showing divergence from 2009 following the global financial crisis.

Most notably, TRY/USD exhibits the most substantial negative external field, signaling a continuous decline in its value, as shown in Figure 3 and 15.

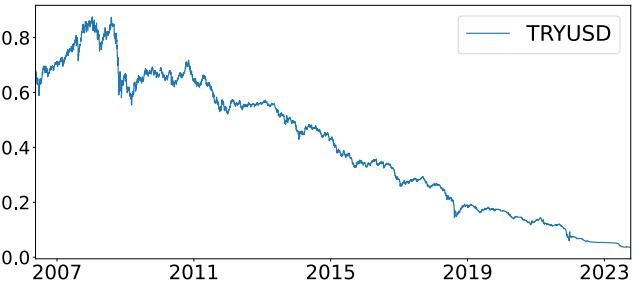


Figure 15: Time series plot of TRY/USD, highlighting the Lira's significant depreciation post-2008 financial crisis.

This underscores the necessity for an overall examination of the global trends shaping these currency movements. In this vein, the following section will delve into the temporal evolution of market dynamics.

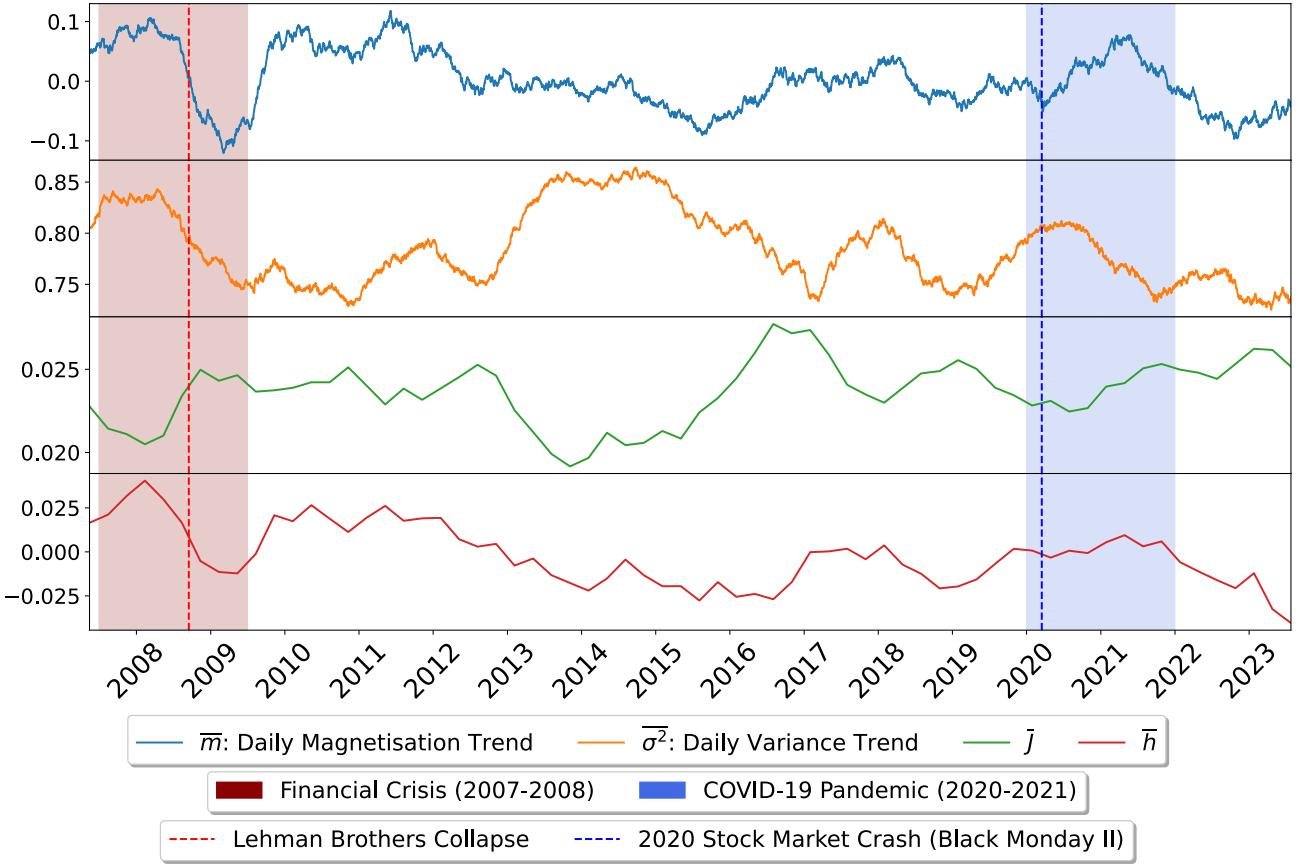


Figure 16: Forex markets time series plots, with samples drawn from 4538 observations between 2006-05-17 and 2023-10-09. Two major global financial crisis epochs are annotated alongside two pivotal moments in the stock markets: the collapse of Lehman Brothers on September 15, 2008, and the onset of Black Monday II on March 16, 2020. The table to the right presents the descriptive statistics for these moving averages.

	$m(t)$	$\sigma^2(t)$	$\bar{J}(t)$	$\bar{h}(t)$
Count	4279	4279	66	66
Mean	1.6e-3	0.784	0.024	-2.0e-3
Median	-2.8e-3	0.776	0.024	-3.5e-3
Std	0.049	0.036	1.9e-3	0.017

### 4.3 Market Dynamics Time Series

The moving averages depicted in Figure 16 illuminate the global movements and variability of currencies, referred to as the Daily Magnetisation Trend (DMT) and Daily Variance Trend (DVT), respectively. Alongside these, the granularly optimised parameters  $\bar{J}$  and  $\bar{h}$  were achieved through the computational efficiency of pseudolikelihood. An average of 260 trading days per year—adjusted for leap years—was derived from the dataset, setting the window size for each trajectory accordingly. A step size of 65 days, equivalent to a quarter of a trading year, was chosen to maintain analytical consistency over time.

The DMT and DVT represent the rolling mean of daily magnetisation and variance, respectively, across all 25 currencies analysed (excluding DKK/USD), with the window size matching the parameters. This exclusion is due to their coupled behaviour, which skews the overall average couplings  $\bar{J}$ . Daily magnetisation is quantitatively measured as the average value of spins at a time  $t$ :

$$m(t) = \frac{1}{d} \sum_{i=1}^d x_i(t) \quad (8)$$

where  $d$  is the number of currencies. A constant theme across the analysed period is the presence of positive  $\bar{J}$ , attributed to the role of the USD as a normalising factor

for all considered currencies. This observation signifies how the USD's fluctuating strength acts akin to a *time-varying external field*, synchronising currency movements and fostering alignment rather than opposition. The role of the US dollar as a unit will be investigated in the next section.

During the Great Recession (2007-07-01 to 2009-06-30, as marked on the plot), a noticeable trend was a significant drop in DMT, accompanied by an increase in  $\bar{J}$ , indicating a net alignment among the currencies. This emerging pattern suggests a crisis signature characterised by a pronounced inverse correlation between DMT and  $\bar{J}$  throughout the year.

The Forex markets' resilience in the COVID pandemic is noteworthy, in contrast to the stock market's significant downturn in March 2020. Despite the apparent drop in DMT and minor alignment suggested by increasing  $\bar{J}$ , these observations become pronounced only when contextualised within marked annotations. This leads us to aim for characterising their statistical behaviours in distinct economic periods, which will be probed in Section 4.5.

A striking feature of the plots is that the DVT and  $\bar{J}$  visually mirror each other, with their Pearson correlation coefficient (PCC) at -0.855. This however is explainable, as when all spins align with each other, the variance of the magnetisation must decrease. Furthermore, the PCC

between the average external fields and DMT, at approximately 0.854, guided the selection of the optimal window size for moving averages. This high PCC suggests that our analysis has been successful since the external field was introduced as a Lagrange multiplier to capture the average rate of change for each currency. However, altering the window size led to a reduced PCC, highlighting the sensitivity of the pseudolikelihood estimator to sample size. The challenges of using a large window include not only the tendency to overly flatten the overall trend of fluctuations but also to result in insufficient data points for plotting. Conversely, excessively small sub-sampling risks deviating from the maximum likelihood estimator’s accuracy [16], necessitating a balanced approach.

This dilemma underlines the need for a robust dataset with higher frequency observations, such as hourly data, to more effectively capture market evolution. However, it’s essential to note that exceedingly frequent data can also be dominated by noise, requiring careful analysis.

#### 4.4 Modelling the Hidden Variable

When considering the currency system, a dominant factor emerges as the strength of the US dollar. This influence, while substantial, remains obscured because all currencies in our analysis are denominated in dollars. To gauge the strength of currencies, one might draw comparisons with tangible metrics. For instance, comparing the economies of two countries typically involves examining GDP per capita [24]. However, this measure may not fully encapsulate economic reality due to the variance in living costs across different nations. For example, Norway’s high GDP per capita contrasts with its elevated living expenses. An alternative, more absolute measure might involve purchasing power parity (PPP) [25], assessing how much a currency can purchase in terms of goods, such as the number of apples or the Big Mac Index. Yet, even PPP presents challenges in accurately reflecting economic strength.

The concept of a hidden variable, or state, representing the fluctuating strength of the US dollar introduces an additional layer to our analysis. The assumption is that a rise in the US economy precipitates a decline in other currencies. Particularly, external fields are indicative of a long-term average response to the US dollar’s fluctuations. From an informational perspective, acknowledging a single, unifying factor that connects to all currencies—a factor deeply intertwined with the US economic landscape—becomes imperative. By considering this dynamic, we hypothesise a new Hamiltonian encompassing all observed currencies plus a hypothetical currency ‘ $u$ ’ (denoting the USD). This introduces a new, unobserved spin, ‘ $u$ ’, and the challenge becomes estimating its best representation.

Essentially, this hidden variable acts as a time-dependent external field, influencing all currencies within the system, despite the constancy of their mutual couplings. This consideration ensures that the inferred parameters more accurately reflect the intended representations, thereby justifying the inclusion of this ‘one unifying factor’ in our model. Specifically, it refines the skew of correlations, making them more representative of the true interconnections within the currency network.

This exploration introduces a revised Hamiltonian:

$$H(\mathbf{x}, u | \mathbf{J}, \mathbf{h}, \mathbf{w}, b) = -\sum_{i=1}^d h_i x_i - \frac{1}{2} \sum_{i=1}^d \sum_{j=1, j \neq i}^d J_{ij} x_i x_j - \sum_{i=1}^d w_i x_i u - bu \quad (9)$$

The initial terms mirror those of the Ising Hamiltonian applied in our prior analysis. The additional terms introduce a new facet to the Hamiltonian, where the  $\mathbf{w}$  vector symbolises couplings, akin to a weighted adjacency in network, between the observed currencies,  $\mathbf{x}$ , and the hidden currency,  $u$ , with  $b$  as a scalar reflecting bias which acts as the external field influencing this concealed currency.

The formulation below reshapes the likelihood into a form that illuminates the complex interplay between observed and hidden variables:

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) \approx \sum_{t=1}^N \sum_{i=1}^d -\ln \left( 1 + \frac{\cosh(-w_i x_i + S_u)}{e^{2x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)} \right) \quad (10)$$

where

$$S = h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j \quad S_u = b + \sum_{j=1, j \neq i}^d w_j x_j.$$

This expression, albeit complex, is manageable, with the derivation provided in Appendix 7.2. By incorporating an unobserved variable into the Ising model, this model falls within the class of Boltzmann machines. These models typically demand extensive computational resources associated with the expectation-maximisation algorithm, which iterates between estimating missing variables to calculate an effective likelihood (the expectation step) and then optimising this likelihood (the maximisation step) [26]. However, integrating over all possible values of hidden variables presents challenges due to the intractable partition function involved in conventional likelihood approaches.

Our methodology sidesteps these hurdles by directly employing pseudolikelihood, which does not require computation of the partition function, as well as integration over hidden states, thereby simplifying the analytical process. Intriguingly, by adhering to the principle of energy minimisation, we can reconstruct the entire set of samples for  $u$ , given the parameter vector  $\boldsymbol{\theta}$  is optimised. This notion is philosophically compelling if our assumption holds true that this unobserved variable accurately represents the fluctuating strength of the US dollar, effectively turning it into a measurable unit within our model.

However, it turned out that the Hessian matrix of this formulation is not negative-definite, indicating it does not effectively maximise for parameter inference. This finding was particularly perplexing because the pseudolikelihood is expected to be a concave function [27]. Our analysis revealed that the cosh term in the denominator of Eq. 10 results in the Hessian matrix having positive eigenvalues when second derivatives are taken with respect to  $\mathbf{w}$  and  $b$ .

Ultimately, the inclusion of this section aims to share our motivation—which drove several months of our work—and invite further investigation into this unresolved enigma.

## 4.5 Inferring Criticality

Exploring the realm of physics, we encounter the concept of phase transitions—dramatic systemic changes triggered when certain parameter thresholds are crossed.

### 4.5.1 Phase Transitions in Nature

Most phase transitions result from the competition between energy and entropy. On the one hand, attractive forces pull particles together, and on the other, entropy drives the particles to explore more configurations. This balance between energy and entropy manifests in the equation of Helmholtz free energy:

$$F = U - TS \quad (11)$$

where  $U$  is the internal energy of the particles,  $S$  is the entropy of the system, and  $T$  is the temperature. Low temperatures favour energy minimisation, while high temperatures favour entropy maximisation. A first-order phase transition occurs with a discontinuity in the free energy with respect to the order parameter. In a liquid-gas transition, the order parameter is density. At the point of discontinuity, two very different phases, a high-density liquid and a low-density gas, have the same free energy and can coexist on either side of a phase boundary.

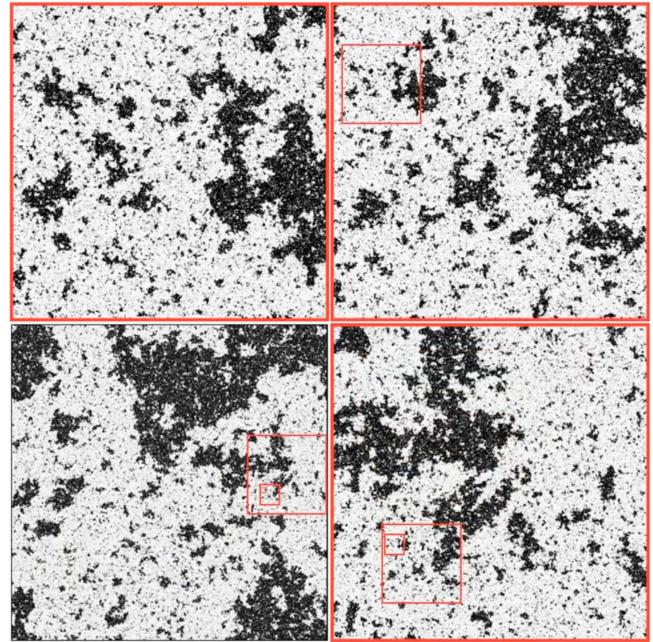
As temperatures rise, the boundary between the two begins to weaken, and fluctuations between the phases grow stronger and stronger. At the critical point, the fluctuations diverge. This is also known as a continuous phase, or second-order, transition because there is no jump in free energy between the phases.

The behaviour of the system around its critical point is one of the strangest, yet most beautiful results in statistical physics. The divergence of fluctuations creates a scale invariance with no length scale taking priority, as shown in Figure 17. This implies that the specific composition of a material, whether liquid, magnet, or otherwise, becomes less relevant. The rules of physics are driven by much simpler concepts such as dimensionality and symmetry. This universality enables the critical point to be one area where a simple model can produce exact results.

The simplest model for a system with attraction and randomness is the Ising model. The two-dimensional version is the classic model for demonstrating critical phenomena. While it has been solved analytically in zero field [29], it is also useful for testing theories such as the renormalisation group [30] and demonstrating behaviours such as scale invariance [31].

A prevailing hypothesis in complexity science posits that non-equilibrium systems often self-organise into a state of criticality, which is repeatedly claimed to be a universal theory of complex adaptive behaviours [32]. The complex system characterised by the Ising-spin network transitions from a disordered, high-temperature paramagnetic phase with weak correlations to a low-temperature spin-glass phase, marked by intense local interactions and large correlations [33]. At the critical state, the system exhibits unique properties like optimal input sensitivity [34], enhanced coordination [35], and maximised computational efficiency [36], embodying the edge-of-chaos principle.

This notion of criticality is mirrored in the field of neuroscience through the *critical brain hypothesis*, which states



**Figure 17:** 2D Ising model at criticality. This composite image, arranged to be viewed clockwise starting from the top left, showcases four very different length scales, each progressively zoomed out. It reveals the system's fractal structure, except this is not deterministic but a statistical distribution. Adapted from [28].

that the brain operates near criticality, a state that optimises its responsiveness and information processing [37]. While the scale of our system may not directly compare to the vast networks of the brain, this framework prompts us to question the proximity of our system to phase boundaries or critical points, where transformative changes occur, akin to crises.

By simulating thousands of spins, adhering to the distribution patterns observed in our analysis, we attempted to ascertain whether our modelled forex market system exhibits signs of phase transitions. This exploration bridges the realms of finance and physics, seeking to understand if the forex markets, like the brain, are poised on the edge of criticality, ready for rapid adaptation and optimisation in response to fluctuating global economic conditions.

### 4.5.2 Phase Transitions in Artificial Structures

Following this line of inquiry, we conducted Monte Carlo simulations across four distinct economic periods:

	Start Date	End Date	Volatility
Great Recession	2008-09-17	2009-09-15	0.0115
Euro Debt Crisis	2011-05-06	2012-05-03	0.0074
Stable	2017-05-31	2018-05-29	0.0044
Stablest	2019-03-08	2020-03-06	0.0040

Table 3: List of analysed economic periods.

Here, ‘Stablest’ marks the period of greatest stability, coinciding with the period before the 2020 stock market crash (Black Monday I, 9 March), while the most volatile period was identified as the Great Recession, im-

mediately following the bankruptcy of Lehman Brothers on 15 September 2008. These periods were confirmed by the volatility clustering analysis in Figure 18, where the subsequent most stable and volatile periods were also identified without overlapping with the previously identified extremes. The next most volatile period was recognised as the European Sovereign Debt Crisis.

All these distinct periods form trajectories, each with an equal amount of data, consisting of 260 trading days, equating to  $26 \times 260 = 6,760$  data points, to effectively eliminate sample size bias for estimators [38].

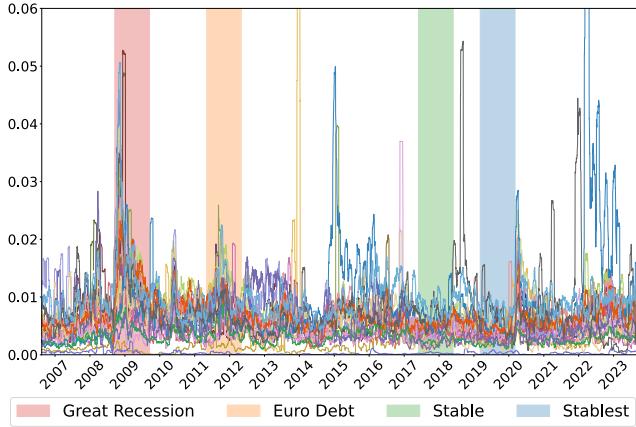


Figure 18: The rolling volatility of daily log returns for 26 currencies was calculated over a window of 260 trading days per year. The analysis focused on identifying distinct periods of market stability and volatility. The most stable period was chosen for its lowest mean rolling volatility, while the most volatile was identified by its highest. These periods are colour-coded in accordance with Table 3.

For the simulation, we sampled data using the Metropolis algorithm, which involves flipping spins according to the energy change  $\Delta E$  and the Boltzmann weight  $e^{-\Delta E/kT}$ , setting the Boltzmann constant  $k$  to 1. We monitored the autocorrelation time to ensure that samples were independent and identically distributed, facilitating proper system equilibrium for effective averaging of ensembles. Note that ‘temperature’ is conceptually absorbed into our model parameters  $\mathbf{J}$  and  $\mathbf{h}$  (see Eq. 3), which become dimensionless, specifying the system’s configuration. In this formulation, we could vary these parameters by the prefactor, which is the temperature,  $T$ , spanning from 0.5 to 1.5 to examine the proximity of our system to phase transitions. Since  $T = 1$  corresponds to where the system actually is, the range of  $T$  was decided as above; halving  $T$  has the effect of doubling  $\mathbf{J}$  and  $\mathbf{h}$ .

Given the parameter space is quantitatively divided by three distinct phases<sup>4</sup>, the first order parameter we chose was the magnitude of average magnetisation,  $|m|$ . Since the magnetisation  $M$  from the initial configuration was

<sup>4</sup>**Paramagnetic (P) phase:** Characterised by both  $m = 0$  and  $q = 0$  as  $d \rightarrow \infty$ , with spins ( $x_i$ ) fluctuating randomly between 1 and -1 without producing overall magnetisation.

**Ferromagnetic (F) phase:** Defined by  $m \neq 0$  and  $q > 0$ . Here, nearly all spins align in a single direction, either as  $x_i = 1$  or  $x_i = -1$ , leading to non-zero magnetisation.

**Spin-Glass (SG) phase:** Distinguished by  $m = 0$  and  $q > 0$ , indicating spins are locally magnetised (fixed in either direction) but lack global alignment, resulting in no net magnetisation.

updated as the spin flips, once the system reached equilibrium,  $|m|$  for each observation was obtained by  $|m| = |M/d|$ , where  $d$  is the number of spins. The reason that we opted to analyse the magnitude of this quantity is that averaging over ensembles would potentially lead to  $m \approx 0$  due to symmetry breaking, even if  $m \neq 0$  in theory, such as in the ferromagnetic phase. This decision also enables a straightforward assessment of the collective orientation intensity of the spins, independent of their initial directional bias since the initial configuration of each trajectory was derived from averaging the binarised states of each currency, positively biased as +1 and negatively as -1 based on their deviation from zero. Other observables were also measured according to the following equations [39]:

$$\begin{aligned} \text{susceptibility, } \chi &= \frac{1}{T} \left( \frac{1}{N} \sum_{i=1}^N m_i^2 - \left( \frac{1}{N} \sum_{i=1}^N m_i \right)^2 \right) \\ \text{spin-glass order parameter, } q &= \frac{1}{N} \sum_{i=1}^N m_i^2 \\ \text{spin-glass susceptibility, } \chi_{sg} &= \frac{1}{NT^2} \sum_{i,j=1}^N c_{ij}^2 \end{aligned} \quad (12)$$

Here,  $N$  is the number of ensembles after the system has reached equilibrium, and  $c_{ij} = \langle x_i x_j \rangle - m_i m_j$  is the covariance matrix, where  $\langle \cdot \rangle$  denotes the ensemble average. However, peaks in  $\chi_{sg}$  had critical temperatures  $T_c$  below  $T = 0.5$  for all simulations, which means that our system was *too hot* for the glassy type of phase transition due to finite size effects, so the results are omitted in this. The simulation results in Figure 19 display  $|m|$  and  $\chi$  across  $T = \{0.5, 0.525, 0.55, 0.575, \dots, 1.5\}$  for distinct periods.

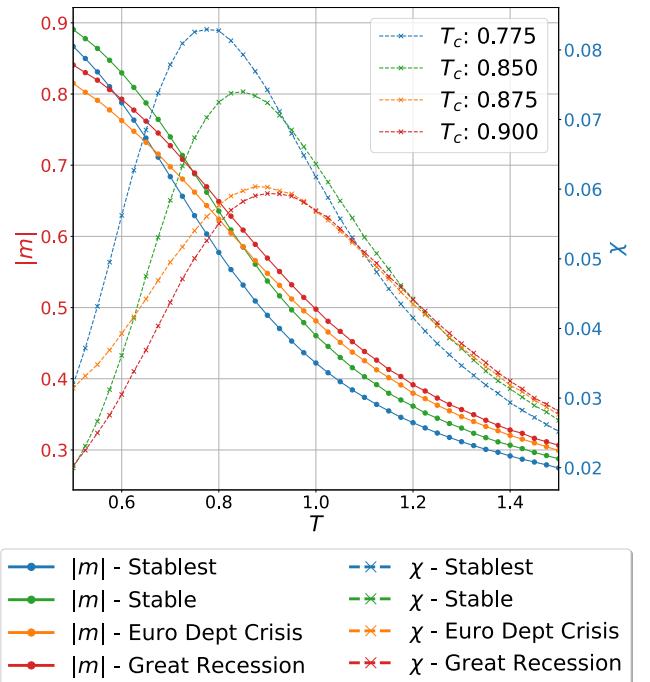


Figure 19: Magnitude of average magnetisation  $|m|$  and susceptibility  $\chi$  as functions of temperature  $T$ . The error bars of ensembles are minute since we allowed the system to reach equilibrium; therefore, they are excluded from the following figures, including this one.

Across all examined periods, we did not observe sharp phase transitions; rather, a smooth decrease in  $|m|$  was evident as  $T$  increased. Such a pattern typifies the characteristics of finite-sized systems, where transitions manifest smeared compared to the abrupt shifts observed at critical points in infinite systems. Surprisingly, the most volatile period features a susceptibility peak that is closest to  $T = 1$ , suggesting the global financial crisis is nearing a *critical state*, indicative of an approaching P-F phase boundary. Furthermore, the observation that the more volatile periods are closer to  $T = 1$  suggests underlying mechanisms<sup>5</sup> in the market that drive it away from a state of criticality. This is interesting since the phase transition structure only emerges in the limit of large systems. So, we presumed it was unlikely that our simulations would yield a signature of phase transitions. Figure 20 displays the unusual market behaviour during the Great Recession.

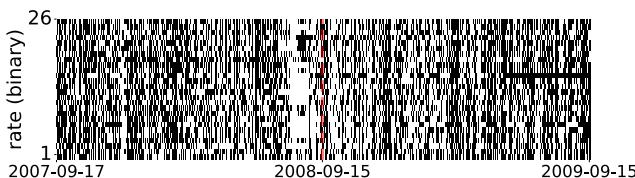


Figure 20: Raster representation of binarised exchange rates with black dots as  $+1$  and white as  $-1$ , respectively. Before the bankruptcy of Lehman Brothers, marked by the red dashed line, a distinct alignment of currencies is observed. One currency that has the noticeably peculiar increase between 2009-05-20 and 2009-09-15 is HKD/USD.

However, since all critical temperatures fall below  $T = 1$ , it is not possible to give an interpretation of our currency system. If we narrow down the period, our system might sit nearly at the critical temperature, but there is a tendency of bias toward criticality with parameters inferred using fewer samples [38], so we decided not to investigate narrower regions. Instead, another viable option arose: upscaling the system size to further investigate the phase boundaries.

This is motivated by the fact that we have a limited number of currencies in our model. If the system is scaled up proportionally where the spins have a similar distribution of couplings, we could check if the extensive Ising model has phase transitions. Figure 21 shows an example of a parameter matrix, where we could fit those parameters against a series of distributions to characterise their statistical behaviours. In this vein, we fitted couplings inferred from four trajectories analysed to all distributions available in the `scipy.stats` library, excluding those that solely require positive couplings.

Figure 22 reveals the top 2 distributions that passed the goodness of fit test by Kolmogorov-Smirnov (KS) statistics, which indicates the maximum difference between the cumulative distribution functions of the empirical data and the fitted distribution. All these distributions of  $\mathbf{J}$  across distinct economic periods are slightly right-skewed with a noticeable positive heavy tail. This signifies the direct influence of the US dollar as an unseen factor, which connects to all of the currencies and drives them to align

<sup>5</sup>perhaps homeostatic mechanisms.

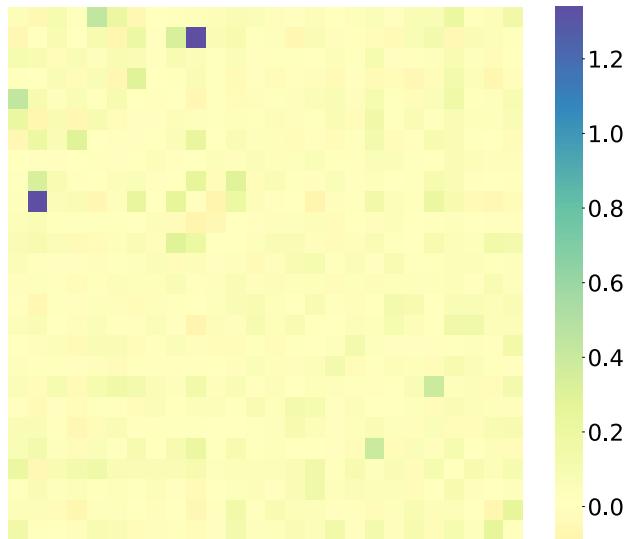


Figure 21: Representative heatmap of  $\mathbf{J}$  where the diagonal elements are filled with the corresponding  $\mathbf{h}$ , forming a  $26 \times 26$  symmetric parameter matrix, where  $\mathbf{J}$  and  $\mathbf{h}$  are inferred from the period of the Great Recession.

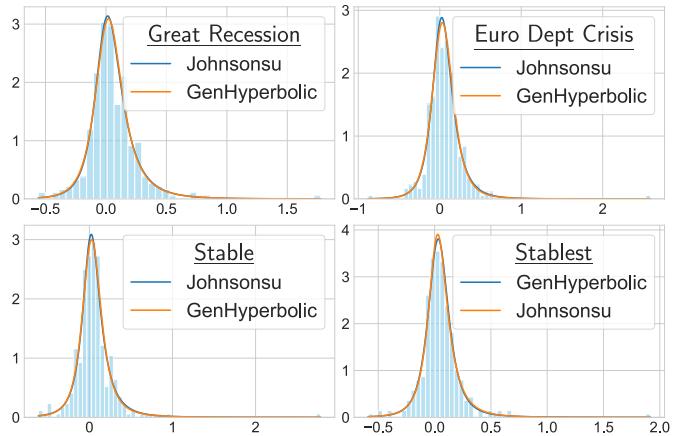


Figure 22: Probability density vs. couplings  $J_{ij}$  for four analysed economic periods. KS statistics and  $p$ -values for each distribution are written in the table below.

Distribution	Johnson's $S_U$		GenHyperbolic	
	KS	$p$ -value	KS	$p$ -value
Great Recession	0.035	0.82	0.035	0.80
Euro Debt Crisis	0.029	0.94	0.030	0.92
Stable	0.027	0.97	0.032	0.89
Stablest	0.032	0.89	0.030	0.92

rather than oppose each other. Also, this extreme positive tail is due to the tied coupling between the euro and Danish krone that we saw in Section 4.1.

Given that we obtained the distributions that best fit the couplings for each trajectory, we can generate couplings following these distributions to form a large parameter matrix. Nonetheless, different economic periods are best described by different fitting distributions. This implies that sampling from these assumed probability distributions may lead to inaccuracy. Therefore, the ideal method might be to directly sample from the empirical distribution to populate couplings. Thus, we used the kernel density to obtain a non-parametric estimate of the probability den-

sity function of couplings [40]. This work is still an ongoing process, therefore the rest of the findings are attached in Appendix 7.3.

## 5 Discussion

In statistics, probability distributions are used as models from which we sample data. Based on these samples, we utilise statistical estimators to estimate population parameters. A common example is the sample mean, which estimates the mean of the population. However, this estimation can be problematic for financial datasets, which often exhibit heavy-tailed distributions as we saw. This is peculiar in the context of the Central Limit Theorem (CLT). Various formulations of the CLT depend on specific assumptions, yet they generally converge to a universal result, demonstrating its broad applicability which is testable to the universality [41, 42, 43]. Notably, the CLT stipulates that for the theorem to apply, the random variables involved must be *independent* and possess *finite variance*. The universality of the normal distribution in representing the means of large samples stems from its ability to approximate the sum of numerous random influences, which are prevalent in many real-world processes.

In the currency market, individual participants decide independently whether to buy or sell, and the quantities involved are finite, leading to observed market fluctuations that essentially sum all these individual transactions. This scenario seems ideally *perfect* for the application of the CLT, which might suggest that currency fluctuations should approximate a normal distribution due to aggregates of large random ensembles of independent decisions. However, empirical observations show that currency fluctuations are not normally distributed but are instead heavy-tailed. This discrepancy arises because, contrary to the assumptions of the CLT, not all market interactions are independent and identically distributed. Many have inherent *correlations* and, as Figure 22 reveals, the distribution of  $\mathbf{J}$  is right-skewed with a pronounced positive heavy tail, even indicative of non-Gaussian interdependencies among currencies. This contributes to the heavy-tailed nature of financial data, making accurate estimation challenging due to the unpredictability posed by outliers.

By binarising the log returns, we circumvented this heavy tail problem. Utilising the Ising model, which is the maximum entropy model for binarised data while accounting for correlations, we successfully reconstructed the complex correlation structure of the currency network. Although limited network analysis was conducted due to the network's small size, further analysis could involve techniques such as community detection, clique (fully connected subgraph) identification, and network partitioning to mathematically delineate economic regions [44]. Additionally, altering the dynamic rules for thresholds, specifically by not taking the absolute values of  $\mathbf{J}$ , highlighted the prominent centrality of certain currencies like CHF/USD and JPY/USD, indicating their stability in trade relations over the analysed entire period. However, it is important, as noted in Section 4.1, to identify tied nodes within the network because they can mask other central nodes, such as the Danish krone's overshadowing of the euro.

Despite the ambition of making an application where the user selects the periods of interest and reconstructs the network of that time frame, the network illustrated in this document only covered the entire period due to time constraints. Nevertheless, the developed framework (available on the GitHub page referenced earlier [17]), enables users to construct targeted networks for particular periods using the comprehensive package for inverse Ising inference. Not only limited to currencies but also commodities and bonds could be injected as spin variables. By normalising as log returns, the scale is standardised making the data comparable across different asset classes since these log returns appear roughly like a random walk when aggregated into cumulative sums (see Figure 3). As an extreme example, the network that includes BTC/USD is shown in Figure 23, which is only illustrative since the first statement is that the fluctuations are much higher in cryptocurrencies in terms of the variance in log returns.

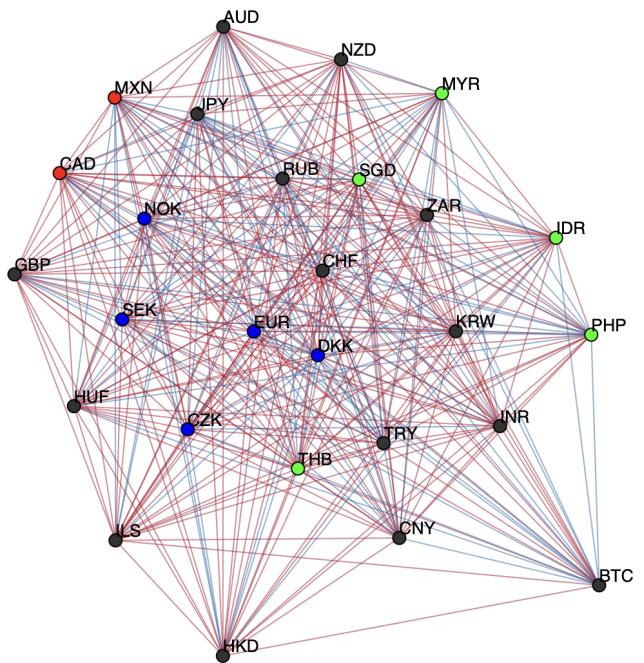


Figure 23: A fully connected graph including BTC/USD spanning from 2014-09-17 to 2023-10-09.

Currency	Betweenness Centrality Rank
BTC/USD	0.474
SGD/USD	0.228
MYR/USD	0.188

Notably, our inferred connectivity matrix is symmetric, under the assumption that detailed balance holds between spins. Yet, this assumption may be questioned in contemporary contexts where the forex markets are influenced not only by traditional exchange rates but also by cryptocurrencies and internet tokens. For instance, within certain periods above, BTC/USD exhibits the highest weighted degree centrality. Interestingly, we always performed post-inference symmetrising steps to estimate couplings, as outlined in Section 3.2. The fact that PLM estimates  $J_{ij}$  and  $J_{ji}$  independently is meaningful, as it allows kinetic and dynamic models with non-symmetric interactions to be more viable for generic network reconstructions. Also,

spite our Hamiltonian being symmetric, this asymmetry appears to suggest the potential to infer causality from the network structure.

Our system, comprising only 26 currencies, models a vast array of dynamics, analogous to observing a ‘shadow’ that only partially represents the actual form; a classical example of a map-territory relation. Therefore, the financial interpretation of inferred parameters is minimised in this report, considering that the data aggregates multiple economic factors. Additionally, analyses focusing on subsets of currencies, such as those from European countries, were not performed due to the potential for external fields to dominate the energy landscape. Such a limited number of elements can introduce an intrinsic bias [2], potentially dominating the characteristics of the inferred correlation structure; this is because of the nature of maximum likelihood, which finds the set of parameters which make observing the data most likely. For example, in a system with only one observed spin, the entire specification would hinge on a single large external field.

Our inverse Ising inference used incorporates a sometimes drastic simplification, by binarising the data. For instance, as shown in Figure 2, most daily log returns cluster around zero and can thus be categorised as the neutral state,  $n$ . This neutral state can be defined within the bounds of the order of one times the median absolute deviation. In this scheme, positive and negative deviations are designated as +1 and -1, respectively. Then, we could employ a Potts model, which allows for multiple (more than two) states in comparison to binary. Thus, the pseudolikelihood implementing this model includes three exponential terms in the denominator (see Appendix 7.2):

$$\begin{aligned} & P(x_i | \mathbf{x}_{\setminus i}; \boldsymbol{\theta}) \\ &= \frac{\exp(h_i x_i + \sum_{j=1, j \neq i}^d J_{ij} x_i x_j)}{\sum_{x'_i=-1, n, 1} \exp(h_i x'_i + \sum_{j=1, j \neq i}^d J_{ij} x'_i x_j)} \\ &= \frac{\exp(x_i(h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j))}{1 + 2 \cosh(h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j)} \end{aligned}$$

where we assigned  $n = 0$ . This formulation is simple; hence, a potential five-state Potts model could assign extreme values a threshold of  $\pm 2$ , potentially capturing more accurate connections in currency network. Such a model has proven effective in predicting the 3D structures of proteins by thresholding up to 21 states [45]. However, this process complicates the model assessment by assigning numbers to non-binary states, introducing additional assumptions. Thus, to maintain theoretical economy—the principle of deriving predictions from the most general and least detailed assumptions—I opted not to explore this alternative modelling approach.

In terms of the Boltzmann machine used as the Ising model with a hidden variable, it is also debatable to consider this variable as the US dollar. If this hypothetical currency ‘ $u$ ’ is intended to represent the USD, its value would be fixed at 1, equivalent to USD/USD, since all analysed currencies being denominated in USD. This conflicts with our foundational assumption of oscillation between  $\pm 1$  (refer to Appendix 7.3). Also, if we regard the

couplings  $\mathbf{w}$  between the observed currencies and the hidden currency as representing unobserved interactions, this hidden unit should symbolise *the rest of the world* rather than specifically the US dollar as initially proposed. This subtlety of the interpretation of this unobserved variable calls into question the validity of binarisation, as it would represent a myriad of complex global factors rather than a single entity (reminiscent of the local hidden variable theory in quantum mechanics that was introduced to represent the *underlying reality*; another theory that was subsequently invalidated). In situations where assumptions are questionable, bootstrapping can be useful for constructing hypothesis tests [46]. By incorporating the historical interest rates of the US dollar, we can represent this as a time-dependent external field. This model can be further refined by including the interest rates of all currencies, each subtracted from that of the USD. The Hamiltonian integrating these factors is given by:

$$H = - \sum_{i=1}^d (h_i + \sum_{a=1}^n h_{ai} f_a) x_i - \frac{1}{2} \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_i x_j$$

where  $h_i$  represents the external field for currency  $x_i$ , while the term  $\sum_{a=1}^n h_{ai} f_a$  aggregates the influence of  $n$  external factors on each currency  $x_i$ .  $h_{ai}$  are currency-dependent coupling constants, each coupled to external factors  $f_a$ , describing the individual currency’s response to each factor. The significance of each factor’s influence can then be evaluated by examining if any of these  $h_{ai}$  values are substantially different from zero.

During the initial stages of this research, substantial time was devoted to identifying outliers, a task complicated by the multifactorial nature of the data represented in spreadsheets. In the world of data, one might conjecture that automated systems could handle such tasks, given that log returns generally fit well to the t-distribution, the assumption that outliers could be systematically filtered out by machines is problematic. Such outliers may represent ‘black swan’ events in complex systems, which are inherently unpredictable and lie outside the range of empirical evidence [47]. The global financial crisis exemplifies such a black swan event, prompting inquiries into constructing systems robust enough to handle extreme outliers.

Our analyses consistently monitored whether our financial system is nearing criticality, indicated by our experience of periodic crises that drive the system to a critical point before abruptly destabilising. We initially hoped to develop a framework that would analyse phase transitions using a dynamic sliding window, rather than static, predetermined windows based on volatility analysis. This dynamic tool would assess in real-time whether the system is approaching the phase boundary. Unfortunately, this could not be implemented due to the extensive computational resources required. At low and critical temperatures, the autocorrelation time substantially lengthens, skewing the statistics without sufficient computational iterations. Nevertheless, we can still examine what the most recent data indicates, revealing a peak in susceptibility reminiscent of the European Debt Crisis period, as shown in Figure 24. This leads researchers to suggest that the reader “buckles up.”

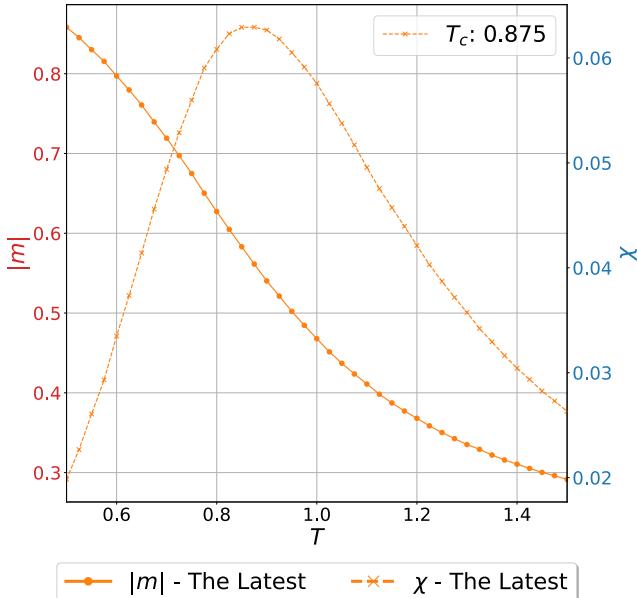


Figure 24:  $|m|$  and  $\chi$  across varied temperature ranges for the latest period between 2023-04-14 and 2024-04-11.

## 6 Conclusions

A question frequently asked during the course of this research was about the potential for predicting future market movements. It is crucial to clarify that inverse problems do not facilitate direct future predictions; they merely identify the most likely parameters that have shaped past market behaviours as we outlined in Section 3.2. Although analysing autocorrelation in inferred parameters can help predict future interactions, it does not determine the future state of the spins (i.e., the currencies).

A highlight of this research was uncovering signals of phase transitions within the financial system. This pursuit, initially motivated by the search for criticality in human-created structures, was philosophically intriguing because the concept of phase transitions is typically introduced to understand the stability of matter.

During this research, one question that occurred was if theoretical realisations inherently exhibit phase transitions in complex systems. In order to answer this question, we reconstructed parameters that specify the distribution of (binarised) currency movements, conducted time series analysis to reveal periods of significant transformations, and introduced a hidden variable to make the correlation more representative of true interactions. This final step was unsuccessful, but leaves room for future research on the nature of this hidden unit affecting the statistics of all currencies.

Much of the work was inspired by neuroscience research. The first significant application of this pairwise maximum entropy modelling was to reconstruct the network of retinal neurons. The hidden variable model was influenced by the inference techniques of unobserved neurons. Moreover, to qualitatively distinguish market phases, order parameters were employed. These are commonly applied in analysing the default mode network, attentional control, and task positive networks of the brain due to the strongly correlated nature of resting-state whole-brain networks. The absence of a signal for the P-SG phase transition in

forex markets led researchers to contemplate the existence of different universality classes in complex systems, further prompting future study into this branch of complexity science.

Drawing insights from neuroscience, it becomes clear that the framework of inverse Ising inference is perfectly suited for financial analysis. In neuroscience, taking derivatives of neuronal spiking data is unintuitive, leading researchers to often binarise the data based on the on and off states of neurons. Also, neurological data samples are usually not balanced, biasing towards criticality in their hypotheses. Importantly, this study does not fall prey to the widespread reproducibility issues found in biological research, increasing its viability.

This inquiry, which bridges the disciplines of physics, economics, and neuroscience, incites an exploration of fundamental questions about the natural world and human-designed systems alike: the search for underlying principles that govern behaviour on the brink of change. By applying analytical tools of one field to the data of another, we not only enrich our understanding of each but also discover universal patterns that transcend the specifics of any single system. The journey through connectivity networks and their evolution, hidden variables, and the quest for criticality in the forex markets thus becomes a metaphor for a broader scientific quest—understanding how complex systems navigate the delicate balance between stability, and the potential for transformative change.

## 7 Appendices

### 7.1 Pseudolikelihood

In the realm of parametric statistics, we encounter models defined up to some unknown parameters. The Ising model is one such example, characterised by parameters  $\mathbf{J}$  and  $\mathbf{h}$ . Maximum likelihood estimation (MLE) offers a methodology for estimating these parameters within the framework of an assumed probability distribution.

#### 7.1.1 Maximum Likelihood Principle

Consider a scenario where we have observed a sample  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ , with  $d$  representing the dimension of data. Given  $N$  samples observed over time  $t$ , the set of all samples,  $\mathbf{X} = \mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$ , forms a sample matrix, termed *trajectory*. The model assigns a probability  $P(\mathbf{x}; \boldsymbol{\theta})$  to observing  $\mathbf{x}$  given parameters  $\boldsymbol{\theta}$ , prompting the question: How do we optimally choose  $\boldsymbol{\theta}$ ?

If we assume each sample is independently drawn from the model's probability distribution, the likelihood of observing the entire dataset  $\mathbf{X}$  is the product of individual sample probabilities:

$$\mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = \prod_{t=1}^N P(\mathbf{x}(t); \boldsymbol{\theta}) \quad (13)$$

In this context,  $\boldsymbol{\theta}$  includes the Ising model's  $\mathbf{J}$  and  $\mathbf{h}$ . MLE advocates for selecting  $\boldsymbol{\theta}$  that maximises  $\mathcal{L}(\mathbf{X}; \boldsymbol{\theta})$ , underpinning the principle that parameter values that render the observed data most probable are optimal. However, this straightforward approach encounters numerous challenges.

### 7.1.2 Challenges of MLE

The likelihood  $\mathcal{L}(\mathbf{X}; \boldsymbol{\theta})$  in Eq. 13 diminishes exponentially with large  $N$ , as probabilities  $P(\mathbf{x}(t); \boldsymbol{\theta})$  are inherently less than one. This challenges computational feasibility, leading to the consideration of  $\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta})$  for maximisation. The monotonic nature of the logarithm function ensures that maximising  $\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta})$  corresponds to maximising  $\mathcal{L}(\mathbf{X}; \boldsymbol{\theta})$ .

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = \sum_{t=1}^N \ln P(\mathbf{x}(t); \boldsymbol{\theta})$$

In our case, the probability distribution is  $P(\mathbf{x}(t); \boldsymbol{\theta}) = \exp(-H(\mathbf{x}(t); \boldsymbol{\theta})) / Z(\boldsymbol{\theta})$ , leading to:

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) = \sum_{t=1}^N [-H(\mathbf{x}(t); \boldsymbol{\theta})] - N \ln Z(\boldsymbol{\theta})$$

Here, the practical difficulty arises in computing  $Z(\boldsymbol{\theta})$ , a sum over all possible states, which becomes infeasible for large datasets, even for a small system like ours with 26 currencies, equivalent to summing  $2^{26} \approx 67$  million terms for each observation.

### 7.1.3 Pseudolikelihood as an Approximation

To circumvent the impracticality of calculating  $Z(\boldsymbol{\theta})$ , we explore approximations. Given  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ , the probability of observing  $\mathbf{x}$  under  $\boldsymbol{\theta}$  can be dissected into a product of conditional probabilities,

$$P(\mathbf{x}; \boldsymbol{\theta}) = P(x_1|x_2, x_3, \dots, x_d; \boldsymbol{\theta})P(x_2, x_3, \dots, x_d; \boldsymbol{\theta})$$

This decomposition can continue down the chain.

$$\begin{aligned} & P(x_2, x_3, \dots, x_d; \boldsymbol{\theta}) \\ &= P(x_2|x_3, x_4, \dots, x_d; \boldsymbol{\theta})P(x_3, x_4, \dots, x_d; \boldsymbol{\theta}) \\ &\quad P(x_3, x_4, \dots, x_N; \boldsymbol{\theta}) \\ &= P(x_3|x_4, x_5, \dots, x_d; \boldsymbol{\theta})P(x_4, x_5, \dots, x_d; \boldsymbol{\theta}) \\ &\quad \vdots \\ & P(\mathbf{x}; \boldsymbol{\theta}) \approx \prod_{i=1}^d P(x_i|\mathbf{x}_{\setminus i}; \boldsymbol{\theta}) \end{aligned}$$

where  $\mathbf{x}_{\setminus i}$  signifies the set of all variables excluding  $x_i$ . This method approximates because it reintroduces  $x_1$ , and subsequently  $x_1, x_2$ , back into the conditional probabilities for each step's given variables, leading to an approximation of the joint probability distribution:

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) \approx \sum_{t=1}^N \sum_{i=1}^d \ln P(x_i|\mathbf{x}_{\setminus i}(t); \boldsymbol{\theta}) \quad (14)$$

This pseudolikelihood formulation calculates the likelihood of observing  $x_i$  given all other  $d-1$  spins, bypassing the need to compute  $Z(\boldsymbol{\theta})$  directly, as shown in:

$$P(x_i|\mathbf{x}_{\setminus i}; \boldsymbol{\theta}) = \frac{e^{-H(\mathbf{x})}/Z(\boldsymbol{\theta})}{\sum_{x'_i=\pm 1} e^{-H(x'_i, \mathbf{x}_{\setminus i})}/Z(\boldsymbol{\theta})} \quad (15)$$

Here,  $Z(\boldsymbol{\theta})$  cancels out. The denominator sums over the two possible states of  $x_i$  specific to the Ising model.

### 7.1.4 Final Expression and Gradients

Given the other  $x_j (j \neq i)$  values fixed to  $\mathbf{x}_{\setminus i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)$ , expanding the Hamiltonian from Eq. 2, Eq. 15 becomes (note that all terms cancel except for those that involve  $x_i$ ):

$$\begin{aligned} & P(x_i|\mathbf{x}_{\setminus i}; \boldsymbol{\theta}) \\ &= \frac{\exp(h_i x_i + \sum_{j=1, j \neq i}^d J_{ij} x_i x_j)}{\sum_{x'_i=\pm 1} \exp(h_i x'_i + \sum_{j=1, j \neq i}^d J_{ij} x'_i x_j)} \\ &= \frac{\exp(x_i(h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j))}{\exp(h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j) + \exp(-(h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j))} \\ &= \frac{e^{x_i S}}{e^S + e^{-S}} \end{aligned}$$

where

$$S = h_i + \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_j$$

The above expression can be further simplified by comparing conditional probabilities when  $x_i = +1$  and  $x_i = -1$ :

$$\begin{aligned} P(x_i = +1|\mathbf{x}_{\setminus i}; \boldsymbol{\theta}) &= \frac{e^{x_i S}}{e^{x_i S} + e^{-x_i S}} = \frac{1}{1 + e^{-2x_i S}} \\ P(x_i = -1|\mathbf{x}_{\setminus i}; \boldsymbol{\theta}) &= \frac{e^{-x_i S}}{e^{-x_i S} + e^{x_i S}} = \frac{1}{e^{-2x_i S} + 1} \end{aligned}$$

where we reintroduced  $x_i$  in denominator terms since  $+1 = x_i$  for the first expression and  $+1 = -x_i$  for the second, respectively, which yield the same results. This is a special case since  $x_i$  is binary. Consequently, Eq. 14 becomes:

$$\ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) \approx \sum_{t=1}^N \sum_{i=1}^d -\ln(1 + e^{-2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]})$$

This expression is Eq. 4. With a large number of samples ( $N \rightarrow \infty$ ), the pseudolikelihood approaches the accuracy of the maximum likelihood estimator [16], assuming the underlying model is correct. Moreover, this function is concave, thereby guaranteeing a unique solution upon maximisation.

Differentiating the pseudolikelihood with respect to  $h_i$  yields:

$$\frac{\partial \ln \mathcal{L}}{\partial h_i} = \frac{2x_i}{2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]} \frac{1 + e^{-2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]}}{1 + e^{2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]}}$$

Likewise, differentiating with respect to  $J_{ij}$  results in:

$$\frac{\partial \ln \mathcal{L}}{\partial J_{ij}} = \frac{2x_i \sum_{j=1, j \neq i}^d x_j}{2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]} \frac{1 + e^{-2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]}}{1 + e^{2x_i[h_i + \sum_{j=1, j \neq i}^d J_{ij} x_j]}}$$

For practical implementation, these expressions were vectorised to avoid nested loops in the code (see Eqs. 5-7), enhancing computational efficiency.

## 7.2 Derivation with the Hidden Variable

The new Hamiltonian (Eq. 9) simplifies to:

$$H = -H_{\mathbf{x}} - uH_u$$

where

$$H_{\mathbf{x}} = \sum_{i=1}^d h_i x_i + \frac{1}{2} \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_i x_j \quad H_u = \sum_{i=1}^d w_i x_i + b.$$

The foundational assumption is that this hidden currency can oscillate between two states: +1 or -1, akin to all observed currencies. Then the marginal probability distribution of  $\mathbf{x}$ , summing over  $u$ , is given by,

$$\begin{aligned} P(\mathbf{x}; \boldsymbol{\theta}) &= P(\mathbf{x}, u = +1; \boldsymbol{\theta}) + P(\mathbf{x}, u = -1; \boldsymbol{\theta}) \\ &= e^{H_{\mathbf{x}}+H_u}/Z(\boldsymbol{\theta}) + e^{H_{\mathbf{x}}-H_u}/Z(\boldsymbol{\theta}) \\ &= e^{H_{\mathbf{x}}}(e^{H_u} + e^{-H_u})/Z(\boldsymbol{\theta}) \\ &= 2e^{H_{\mathbf{x}}} \cosh(H_u)/Z(\boldsymbol{\theta}) \end{aligned}$$

where  $\boldsymbol{\theta}$  encapsulates all parameters  $\mathbf{J}$ ,  $\mathbf{h}$ ,  $\mathbf{w}$ , and  $b$ . Following the derivation of the pseudolikelihood outlined in Appendix 7.1,

$$\begin{aligned} P(x_i | \mathbf{x}_{\setminus i}; \boldsymbol{\theta}) &= \frac{2e^{H_{\mathbf{x}}(\mathbf{x})} \cosh(H_u(\mathbf{x}))/Z(\boldsymbol{\theta})}{\sum_{x'_i=\pm 1} 2e^{H_{\mathbf{x}}(x'_i, \mathbf{x}_{\setminus i})} \cosh(H_u(x'_i, \mathbf{x}_{\setminus i}))/Z(\boldsymbol{\theta})} \\ &= \frac{e^{x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)}{\sum_{x'_i=\pm 1} e^{x'_i S} \cosh(w_i x'_i + S_u)} \\ &= \frac{e^{x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)}{e^S \cosh(w_i + S_u) + e^{-S} \cosh(-w_i + S_u)} \end{aligned}$$

where

$$S = h_i + \sum_{\substack{j=1 \\ j \neq i}}^d J_{ij} x_j \quad S_u = b + \sum_{\substack{j=1 \\ j \neq i}}^d w_j x_j.$$

This expression can be further simplified following the comparison of conditional probabilities:

$$\begin{aligned} P(x_i = +1 | \mathbf{x}_{\setminus i}; \boldsymbol{\theta}) &= \frac{e^{x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)}{e^{x_i S} \cosh(w_i x_i + S_u) + e^{-x_i S} \cosh(-w_i x_i + S_u)} \\ &= \frac{1}{1 + e^{-2x_i S} \frac{\cosh(-w_i x_i + S_u)}{\cosh(\sum_{i=1}^d w_i x_i + b)}} \\ P(x_i = -1 | \mathbf{x}_{\setminus i}; \boldsymbol{\theta}) &= \frac{e^{x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)}{e^{-x_i S} \cosh(w_i(-x_i) + S_u) + e^{x_i S} \cosh(-w_i x_i + S_u)} \\ &= \frac{1}{e^{-2x_i S} \frac{\cosh(-w_i x_i + S_u)}{\cosh(\sum_{i=1}^d w_i x_i + b)} + 1} \end{aligned}$$

which yields the same results for both expressions. Consequently, the pseudolikelihood including the hidden variable is expressed as:

$$\begin{aligned} \ln \mathcal{L}(\mathbf{X}; \boldsymbol{\theta}) &\approx \sum_{t=1}^N \sum_{i=1}^d -\ln \left( 1 + \frac{\cosh(-w_i x_i + S_u)}{e^{2x_i S} \cosh(\sum_{i=1}^d w_i x_i + b)} \right) \end{aligned}$$

## 7.3 Critical Market Hypothesis

When the microscope was invented in the Netherlands, Antonie van Leeuwenhoek discovered ‘a universe in a drop of water’

In developing the new parameter matrix (PM), we aimed for it to closely resemble the original PM to accurately reflect reality, acknowledging that what we observe is merely an aggregation of many influences, given the complexity of real systems. Based on this, we made the following assumption:

*There is an underlying process generating the distribution of values of  $\mathbf{J}$ , and each parameter observed in the parameter matrix represents the average of a statistical ensemble of those values.*

For simplicity, each element in the new PM is referred to as a *drop*, and the constituents (generated couplings) forming each drop as *universes*. Following our assumption, the mean of these universes within each drop represents the value of the drop. The statistical parameters such as degrees of freedom, location, and scale of a drop dictate the distribution of the universes within it, where the *location* is the value of the drop. Therefore, the universes forming these drops are similarly distributed as drops. If each drop contains 100 universes, the new PM becomes a 260 x 260 symmetric matrix.

Figure 25 illustrates the new PM where each element of the original PM has been transformed into a drop of universes drawn from the distribution characterising the set of original elements. Despite the elements within this new matrix appearing blurred compared to the original PM, as shown in Figure 21, each drop retains the characteristics of the original element. Essentially, this process is the *inversion* of renormalisation, treating all inferred parameters as renormalisations of underlying aggregates.

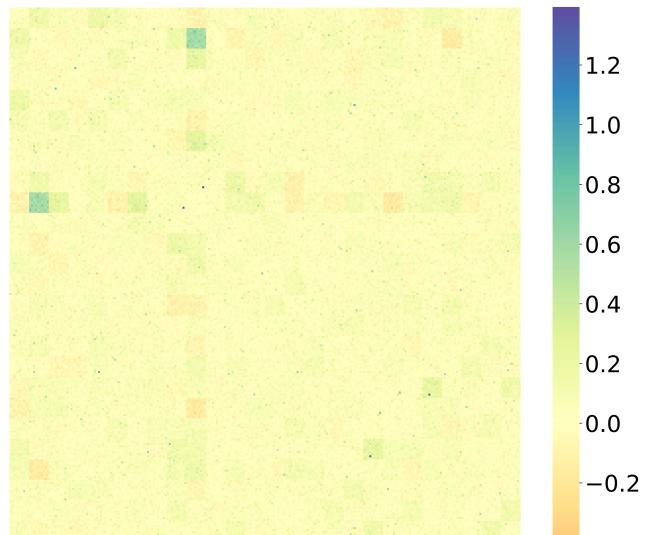


Figure 25: Representative heatmap filled with drops, where 100 universes form 1 drop, resulting in a 260 x 260 symmetric matrix. The fitted distribution is Johnson’s  $S_U$  for the period of the Great Recession.

Regarding the initial configuration for simulations, since each element in the original PM represents the parameters  $J_{ij}$  and  $h_i$ , which specify the statistics of each currency, this setup involves sampling from the trajectory of each currency, where those sampled spins will be specified by universes within each drop. If the system is scaled by a factor of  $N = 100$ , considering that the trajectory  $\mathbf{X}$  includes 26 columns corresponding to our system, the first  $N$  values from each column (currency) collectively form an initial array of  $N \times 26 = 2600$  spins, forming the basis of the initial configuration.

In the Ising model, to recover the thermodynamic limit, it is necessary to reduce the typical size of interactions  $J_{ij}$  as the number of spins increases. Given our network's fully connected structure, we applied a  $N^{-1/2}$  scaling of couplings following the Bethe-Peierls approach [48]. We confirmed that this scaling preserved tail-properties of couplings such as skewness and kurtosis as the system was scaled. The per-site magnetisation is represented as:

$$m_i = \tanh \left( \beta h_i + \beta \sum_j J_{ij} m_j - \beta^2 \sum_j J_{ij}^2 m_i (1 - m_j^2) \right)$$

where  $\beta$  is inverse temperature ( $1/T$ ). Figure 26 displays the results for a system scaled by a factor of  $N = 5$ .

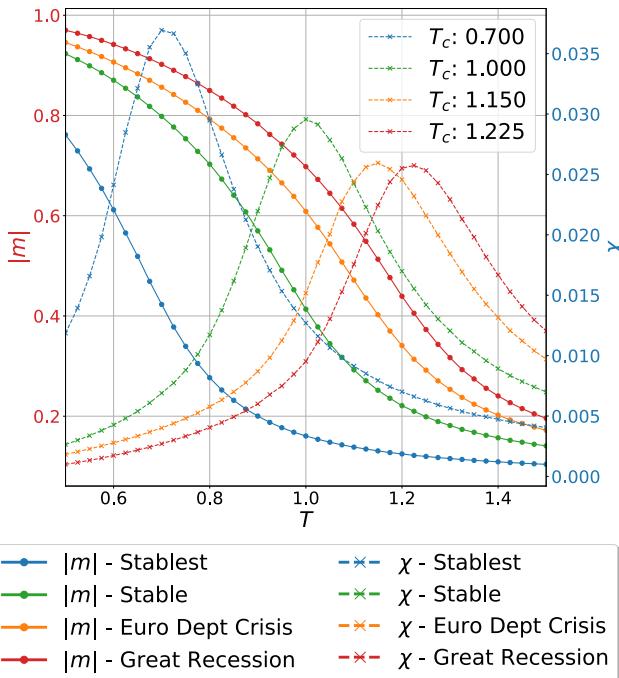


Figure 26:  $|m|$  and  $\chi$  across varied temperature ranges for distinct periods analysed with the system of 130 spins. Couplings  $J_{ij}$  are sampled from a generalised hyperbolic distribution for the ‘Stablest’ period and from Johnson’s  $S_U$  for all others, according to Figure 22, while per-site dependent external fields  $h_i$  are replicated for sampled spins specified by universes within each drop.

Surprisingly, this outcome indicates that financial crises do not themselves undergo a phase transition; instead, the stable market operates continuously at the edge of criticality, contradicting the conclusions drawn in Section 4.5.2. This is where the renormalisation group (RG) becomes relevant [30]. Inverting the renormalisation process

introduces assumptions into the model, and reversing the transformation should return the system to its original state.

In real-space RG theory, this process involves grouping a block of spins together to form a new super spin that reflects the dominant value of its constituents. It’s akin to the spins in the block collectively deciding on their representation, which allows us to treat them as a single entity. In our case, however, these ‘spins’ are universes (generated couplings) within each drop. Figure 27 shows the results from the renormalisation by block transformation for  $N = 5$ , enabling verification that the scaled-down system produces similar distributions of  $|m|$  and  $\chi$ .

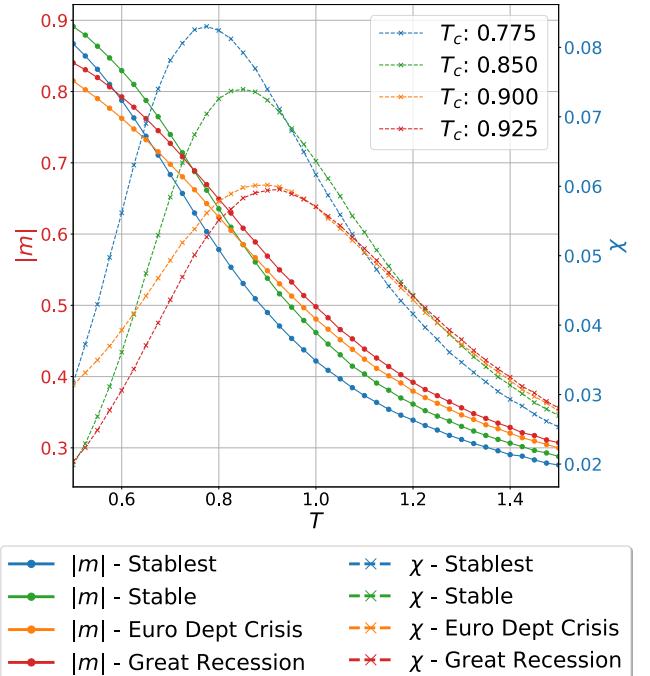


Figure 27:  $|m|$  and  $\chi$  across varied temperature ranges for analysed periods where a group of  $J_{ij}$  are renormalised from Figure 26 and the original  $h_i$  were used.

The location of peaks and the shape of all distributions are nearly identical to the original distributions we observed in Figure 19, demonstrating consistency. The only change is that critical temperatures  $T_c$  during crisis periods shifted slightly higher by 0.25 since we introduced a stochastic process for populating couplings.

Although this result from Figure 26 was renormalised successfully, the shape of distributions appears very different from each other. This may be due to an increased number of couplings. When there are more generated couplings in the scaled system, this might lead to a cluster exhibiting local magnetisations (such as spins specified by universes generated from  $J_{ij}$  between DKK/USD and EUR/USD). As this scale progressively increases, these locally cooperative regions would begin to appear at every scale — scale invariance. This defines a critical state. This effect may be responsible for the criticality observed in this expanded model.

Further consistency checks are necessary. This analysis is based on fitted distributions; therefore, we attempted to sample directly from the empirical distribution. Given that our couplings are unimodally distributed (see Figure

22), a kernel density estimator was used to infer the population of couplings, as depicted in Figure 28. However, time constraints prevented the development of algorithms to match up per-site dependent external fields with associated couplings. This process was more straightforward when using drops and universes populated by fitted distributions, as external fields were simply replicated according to each diagonal element of the original PM.

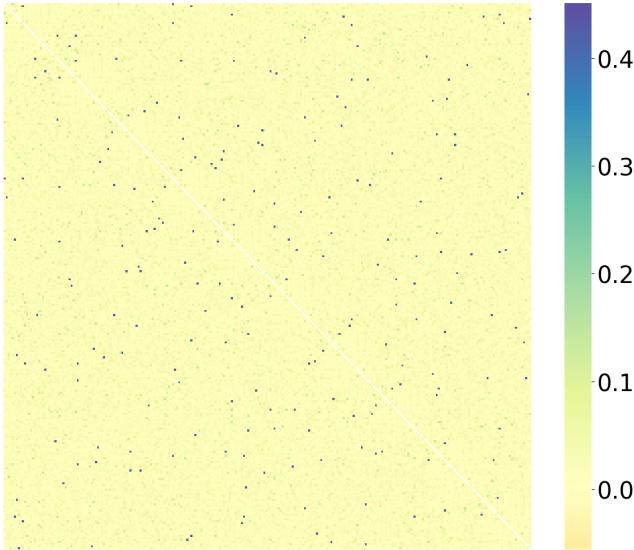


Figure 28: Representative heatmap filled with empirically populated couplings, resulting in a  $260 \times 260$  symmetric matrix for the period of the Great Recession.

For those interested in reproducing this process, it is critical to preserve higher-order moments to maintain the characteristics of the distribution of couplings. However, properly dividing generated couplings by an outlined scaling factor  $N^{-1/2}$  should be sufficient to ensure the preservation of these statistical properties. Initially, we hoped to simulate the model scaled by a factor of  $N = 100$  to accommodate 2500 spins, but the computational resources currently available to us limit this number to  $N = 5$ .

Overall, this finding is interesting, as it suggests that a version of the currency market with higher complexity owing to a larger number of currencies would be closer to phase transitions in its stable state. However, it remains unclear why the transition temperature of specifically the ‘stable’ market period lies at  $T = 1^6$ . Also, we cannot draw conclusions from this result, as the complication of the model is likely to reduce the efficacy of the mapping between the model and the actual system. If the realisation of an expanded network could be validated, that would lead to the framework of *critical market hypothesis*.

<sup>6</sup>The specific positioning of  $T_c = 1$  for the stable market is explainable by the power-law decay of the impact function [49]. Jean-Philippe et al. suggest that the market is at a critical point capable of quick reconfiguration to maintain equilibrium where competing market dynamics, specifically the actions of liquidity providers and takers, are balanced. This results in small changes propagating large effects, exhibiting power-law tails in return distributions.

## References

- [1] Ernst Ising. *Beitrag zur theorie des ferro-und paramagnetismus*. PhD thesis, Grefe & Tiedemann Hamburg, Germany, 1924.
- [2] Elad Schneidman, Michael J Berry, Ronen Segev, and William Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087):1007–1012, 2006.
- [3] Andrzej Grabowski and RA Kosiński. Ising-based model of opinion formation in a complex network of interpersonal interactions. *Physica A: Statistical Mechanics and its Applications*, 361(2):651–664, 2006.
- [4] Didier Sornette. Physics and financial economics (1776–2014): puzzles, ising and agent-based models. *Reports on progress in physics*, 77(6):062001, 2014.
- [5] Thomas Bury. Market structure explained by pairwise interactions. *Physica A: Statistical Mechanics and its Applications*, 392(6):1375–1385, 2013.
- [6] H Chau Nguyen, Riccardo Zecchina, and Johannes Berg. Inverse statistical problems: from the inverse ising problem to data science. *Advances in Physics*, 66(3):197–261, 2017.
- [7] Danh-Tai Hoang, Juyong Song, Vipul Periwal, and Junghyo Jo. Network inference in stochastic systems from neurons to currencies: Improved performance at small sample size. *Physical Review E*, 99(2):023311, 2019.
- [8] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.
- [9] Toshiyuki Tanaka. A theory of mean field approximation. *Advances in Neural Information Processing Systems*, 11, 1998.
- [10] Tamara Broderick, Miroslav Dudik, Gasper Tkacik, Robert E Schapire, and William Bialek. Faster solutions of the inverse pairwise ising problem. *arXiv preprint arXiv:0712.2437*, 2007.
- [11] Takahiro Ezaki, Takamitsu Watanabe, Masayuki Ohzeki, and Naoki Masuda. Energy landscape analysis of neuroimaging data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2096):20160287, 2017.
- [12] Christopher Lynn and Daniel D Lee. Maximizing influence in an ising network: A mean-field optimal solution. *Advances in neural information processing systems*, 29, 2016.
- [13] Zexun Chen, Bo Wang, and Alexander N Gorban. Multivariate gaussian and student-t process regression for multi-output prediction. *Neural Computing and Applications*, 32:3005–3028, 2020.
- [14] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, May 1957.

- [15] Tobias Preis, Johannes J Schneider, and H Eugene Stanley. Switching processes in financial markets. *Proceedings of the National Academy of Sciences*, 108(19):7674–7678, 2011.
- [16] Erik Aurell and Magnus Ekeberg. Inverse ising inference using all the data. *Physical review letters*, 108(9):090201, 2012.
- [17] Sohyun Park. Project repository for inverse ising inference. URL. Accessed: 10 March 2024.
- [18] Robert Malouf. A comparison of algorithms for maximum entropy parameter estimation. In *COLING-02: The 6th Conference on Natural Language Learning 2002 (CoNLL-2002)*, 2002.
- [19] Sohyun Park. Plm currency network. URL. Accessed: 10 March 2024.
- [20] Stephen G Kobourov. Spring embedders and force directed graph drawing algorithms. *arXiv preprint arXiv:1201.3011*, 2012.
- [21] Anuj Singh, Rishi Ranjan Singh, and SRS Iyengar. Node-weighted centrality: a new way of centrality hybridization. *Computational Social Networks*, 7:1–33, 2020.
- [22] Danmarks Nationalbank. Learn about inflation, interest rates and the fixed exchange rate policy. URL. Accessed: 4 March 2024.
- [23] European Central Bank. Foreign exchange operations. URL. Accessed: 4 March 2024.
- [24] International Monetary Fund. World economic outlook database, october 2023. URL. Accessed: 22 March 2024.
- [25] Paul R Krugman and Maurice Obstfeld. *International economics: Theory and policy*. Pearson Education, 2009.
- [26] Chuong B Do and Serafim Batzoglou. What is the expectation maximization algorithm? *Nature biotechnology*, 26(8):897–899, 2008.
- [27] Amir Globerson. Pgm lecture notes: pseudo-likelihood. URL. Accessed: 22 March 2024.
- [28] Douglas Ashton. Scale invariance in the critical ising model. URL, 2024. YouTube video.
- [29] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review*, 65(3-4):117, 1944.
- [30] GS Pawley, RH Swendsen, DJ Wallace, and KG Wilson. Monte carlo renormalization-group calculations of critical behavior in the simple-cubic ising model. *Physical Review B*, 29(7):4030, 1984.
- [31] H.E Stanley, L.A.N Amaral, P Gopikrishnan, P.Ch Ivanov, T.H Keitt, and V Plerou. Scale invariance and universality: organizing principles in complex systems. *Physica A: Statistical Mechanics and its Applications*, 281(1):60–68, 2000.
- [32] Roman Frigg. Self-organised criticality—what it is and what it isn’t. *Studies in History and Philosophy of Science Part A*, 34(3):613–632, 2003.
- [33] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.
- [34] Osame Kinouchi and Mauro Copelli. Optimal dynamical range of excitable networks at criticality. *Nature physics*, 2(5):348–351, 2006.
- [35] Andrea Cavagna, Alessio Cimarelli, Irene Giardina, Giorgio Parisi, Raffaele Santagati, Fabio Stefanini, and Massimiliano Viale. Scale-free correlations in starling flocks. *Proceedings of the National Academy of Sciences*, 107(26):11865–11870, 2010.
- [36] Nils Bertschinger and Thomas Natschläger. Real-time computation at the edge of chaos in recurrent neural networks. *Neural computation*, 16(7):1413–1436, 2004.
- [37] Jordan O’Byrne and Karim Jerbi. How critical is brain criticality? *Trends in Neurosciences*, 45(11):820–837, 2022.
- [38] Maximilian B Kloucek, Thomas Machon, Shogo Kajimura, C Patrick Royall, Naoki Masuda, and Francesco Turci. Biases in inverse ising estimates of near-critical behavior. *Physical Review E*, 108(1):014109, 2023.
- [39] Takahiro Ezaki, Elohim Fonseca dos Reis, Takamitsu Watanabe, Michiko Sakaki, and Naoki Masuda. Closer to critical resting-state neural dynamics in individuals with higher fluid intelligence. *Communications biology*, 3(1):52, 2020.
- [40] S Madeh Piryonesi and Tamer E El-Diraby. Role of data analytics in infrastructure asset management: Overcoming data size and quality problems. *Journal of Transportation Engineering, Part B: Pavements*, 146(2):04020022, 2020.
- [41] Iddo Eliazar and Joseph Klafter. A randomized central limit theorem. *Chemical Physics*, 370(1-3):290–293, 2010.
- [42] Claude Kipnis and SR Srinivasa Varadhan. Central limit theorem for additive functionals of reversible markov processes and applications to simple exclusions. *Communications in Mathematical Physics*, 104(1):1–19, 1986.
- [43] Yuta Koike. High-dimensional central limit theorems for homogeneous sums. *Journal of Theoretical Probability*, 36(1):1–45, 2023.
- [44] Mark Newman. *Networks*. Oxford university press, 2018.
- [45] Magnus Ekeberg, Cecilia Lökvist, Yueheng Lan, Martin Weigt, and Erik Aurell. Improved contact prediction in proteins: using pseudolikelihoods to infer potts models. *Physical Review E*, 87(1):012707, 2013.

- [46] Jerzy Neyman and Egon Sharpe Pearson. Ix. on the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694-706):289–337, 1933.
- [47] Nassim Nicholas Taleb. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications. *arXiv preprint arXiv:2001.10488*, 2020.
- [48] Sergey N Dorogovtsev, Alexander V Goltsev, and José FF Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275, 2008.
- [49] Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart. Fluctuations and response in financial markets: the subtle nature of random’ price changes. *Quantitative finance*, 4(2):176, 2003.

## Certification of ownership of the copyright

This Project Report is presented as part of, and in accordance with, the requirements for the degree of MSci at the University of Bristol, Faculty of Science.

I hereby assert that I own exclusive copyright in the item named below. I give permission to the University of Bristol Library to add this item to its stock and to make it available for consultation in the library, and for inter-library lending for use in another library. I also give consent for this report to made available electronically to staff and students within the University of Bristol. It may be copied in full or in part for any bona fide library or research work. No quotation and no information derived from it may be published without the author's prior consent.

Author	Sohyun Park
Title	Statistical Mechanicis of Foreign Exchange Markets
Date of submission	18 April

I agree that submission of this report constitutes signing of this declaration.

This project/dissertation is the property of the University of Bristol and may only be used with due regard to the rights of the author. Bibliographical references may be noted, but no part may be copied for use or quotation in any published work without the prior permission of the author. In addition, due acknowledgement for any use must be made.