Maximum entropy models for financial systems

Almog, A.

Chapter 3

Community Detection for Time Series

The mesoscopic organisation of complex systems, from financial markets to the brain, is the key intermediate level of organisation between the microscopic dynamics of individual units (stocks or neurons, in the mentioned cases), and the macroscopic dynamics of the system as a whole. The organisation is determined by "communities" of units whose dynamics, represented by time series of activity, is more strongly correlated internally than with the rest of the system. In the current literature, such emergent organisation is mainly detected through the measurement of cross-correlations among time series of nodes activity, the projection (usually via an arbitrary threshold) of these correlations to a network, and the subsequent search for denser modules (or so-called communities) in the network. It is well known that this approach suffers from an unavoidable information loss induced by the thresholding procedure. Another, less realized, limitation is the bias introduced by the use of network-based (as opposed to correlation-based) community detection methods. In this chapter we discuss an improved method for the identification of functional modules based on maximum-entropy. The method is threshold-free, correlation-based, and very powerful in filtering out both local unit-specific noise and global system-wide dependencies. The approach is guaranteed to identify mesoscopic functional modules that, relative to the global signal, have an overall positive internal correlation and negative mutual correlation.

The results presented in this chapter have been published in the following references:
A. Almog, F. Besamusca, M. MacMahon and D. Garlaschelli *PLoS ONE* **10** **7**. e0133679 (2015).
A. Almog, O. Roethler, R. Buijink, J.H. Meijer, J.H.T. Rohling and D. Garlaschelli ,in preparation (2016).
R. Buijink, A. Almog, C.B. Wit, O. Roethler, D. Garlaschelli, J.H. Meijer, J.H.T. Rohling and S. Michel *PLoS ONE*, under second revision (2016).

3.1 Introduction

One of the most important properties of complex systems is community structure. Real-world systems are organized in a modular way, with clusters of units sharing similar dynamics or functionality. However, while the clusters themselves are internally cohesive, externally they can maintain contrasting dynamics. This emergent structure is typically resolved from the recorded activity time series of the system's fundamental units (such as stocks, neurons, etc.). The problem of resolving and identifying these mesoscopic structures, without any prior information, is extremely challenging. In financial markets, the mesoscopic scale corresponds to sets of stocks that share similar price dynamics. The knowledge of the market structure is highly valuable, and can assist in hedging risks and for better understanding of the market. Consequently, over the past years, scientists have deployed and developed many time series techniques to retrieve qualitative information regarding the hierarchy and structure of financial markets [1, 2, 3, 4].

A promising approach is that of employing community detection techniques, developed in network theory [5, 6], on empirical correlation matrices (constructed from multiple time series). However, The attempts made so far have basically replaced network data with cross-correlation matrices, and are not adapted to deal with correlation matrices. More specifically, the methods enforce a network representation on the empirical correlation matrix, and later apply a network-based (as opposed to correlation-based) community detection method. Such a process also involves some type of thresholding procedure, which results in significant information loss. These methodological problems rise significant limitations, when one comes to analyse empirical correlation matrices.

Recently, a novel method was proposed, which has been specifically designed to detect communities from correlation matrices of multiple time series [7]. The method is able to filter out both local unit-specific noise and global system-wide dependencies, using random matrix theory as the null model. When applied to financial time series, the method was able to capture the dynamical modularity of real financial markets. It is able to identify clusters of stocks which are correlated internally, but are anti-correlated with each other.

In this chapter we discuss this improved method and generalize it using a more complete framework of maximum-entropy. The generalized maximum-entropy approach is able to capture both the global dynamics of the system and the random noise induced by the different units. More importantly, this improved null model is more "data-driven" in contrast to the original null model, enabling us to use more empirical information to construct the community structure. In this setting, the null models help us identify the non-random properties in the system to a higher level of resolution.

The rest of the chapter is organized as follows. In section 3.3 we analyse financial markets, exploring the mesoscopic structures induced by both the original weighted time series and their corresponding binary signatures. We show that both the binary and weighted information yield very similar community structures, suggesting that the binary signatures carry significant structural information. In section 3.4 we apply the method to the suprachiasmatic nucleus of mice, which is a complex network of oscillating neurons, revealing a remarkable core periphery structure. This application to functional brain networks comes as a by-product, as time series of brain activity have in common with financial time series the potential of being driven by a strong signal and also being subject to noise. In fact our method could have a great impact on the field of functional brain networks. Finally, in sec 3.5 we summarize our results and provide some conclusions.

3.2 Maximum-entropy approach to community detection

In general, the problem of community detection in network science consists of finding clusters of nodes that have dense connections internally and sparser connections externally. Such clusters, or communities, represent sub-units of the network like families in social networks or brain regions in structural brain networks. Identifying these modules and their boundaries is of great importance, and over the last years, many methods have been developed to resolve this type of organisation [5]. However, here we are interested in identifying functional modules from a correlation matrix, generated by multiple time series, rather than a conventional network.

We now describe the so-called modularity-based community detection methods but adapted to correlation matrices. This restricts us to undirected networks, given the symmetry property of correlation matrices. Let us consider a network with N nodes. One can introduce a number of partitions of the N nodes into non-overlapping sets. The different partitions will be represented by an N-dimensional vector $\vec{\sigma}$ where the i-th component σ_i denotes the set in which node i is placed by that particular partition. Now, we introduce the modularity measure $Q(\vec{\sigma})$ which indicates the quality of a particular choice of partition $\vec{\sigma}$ measured by a high degree of inter-community connectivity and a low degree of intra-community connectivity. So-called modularity optimization algorithms look for the specific partition that maximizes the value of $Q(\vec{\sigma})$, the objective function. The latter is defined as

$$Q(\vec{\sigma}) = \frac{1}{A_{tot}} \sum_{i,j} \left[A_{ij} - \langle A_{ij} \rangle\right] \delta(\sigma_i, \sigma_j) \tag{3.1}$$

where $\delta(\sigma_i, \sigma_j)$ is a delta function ensuring that only pairs of nodes within the

same community contribute to the sum, and A_{ij} is the adjacency matrix that indicates whether a link exists between the nodes, ($A_{ij} = 1$) or not, ($A_{ij} = 0$). The pre-factor A_{tot} serves to normalize the value of $Q(\vec{\boldsymbol{\sigma}})$ between -1 and 1, where $A_{tot} \equiv \sum_{i,j} A_{ij} = 2L$ is twice the number of total links in the network. The term $\langle A_{ij} \rangle$ is vital to the outcome of the community detection process. It represents the expectation of whether a link exists or not, according to the specific null model chosen. So far the majority of the methods use null models (hypotheses) which are suited only for networks. For example, the configuration model is a null model that preserves the degree sequence of the network. It has been shown that such null models can introduce biases when applied to correlation matrices [7]. The reason is that, while a null model for a network assumes independent links, correlation matrices have different metric properties that imply dependencies among their entries, even under the null hypothesis. A recent method proposed a redefinition of the modularity, which does take into account the existence of known properties of correlation matrices (see sec. 3.3.1). Here, we take the maximum-entropy approach, providing a specification of an appropriate null model.

Instead of the previous adjacency matrix A_{ij}, we input the empirical correlation matrix C_{ij}. The method defines the modularity as

$$Q(\vec{\boldsymbol{\sigma}}) = \frac{1}{C_{norm}} \sum_{i,j} \left[C_{ij} - \langle C_{ij} \rangle_{me} \right] \delta(\sigma_i, \sigma_j) \tag{3.2}$$

where $\langle C_{ij} \rangle_{me}$ is a maximum-entropy null model that needs to identify the random properties of empirical correlation matrices.

In this approach, the empirical correlation matrix is first decomposed and then reconstructed using only the eigenvalues (and eigenvectors) that are not reproduced by the random null model. Thus in this chapter, contrary to chapter 1, we are interested in the correlation matrix spectrum which the random model (multiple time series) generates. Once compared with the observed spectrum of the empirical correlation matrix, the model will identify the non-random eigenvalues (by elimination). The non-random eigenvalues will be later used to generate the new filtered matrix. In the next sections, we introduce two null models, which will serve us as the "random benchmark" in this new definition of modularity.

3.2.1 Random time series

Let us go back to the same multiple time series formalism as in chapter 1, where the system describes N time series with length T ($N \times T$ matrix). We recall that the entries of such a general matrix $\mathbf{M}$ are denoted by $r_i(t)$, where i labels the stock and t labels the time step. Before introducing the specific models, we need to make some necessary changes. Here, we extend our null models to weighted

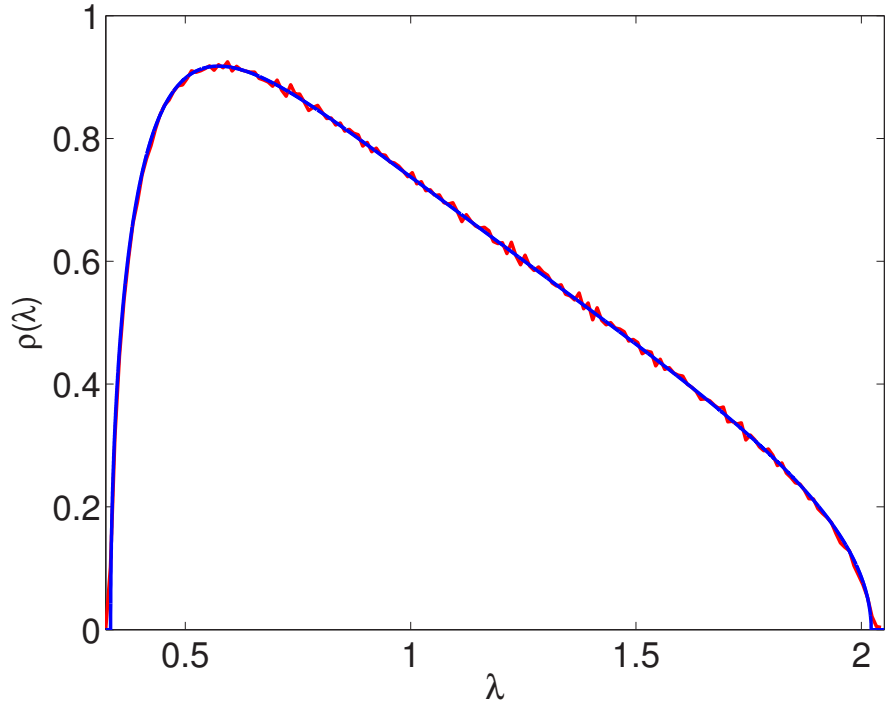

Figure 3.1: **Eigenvalues density distribution comparison: maximum-entropy versus Wishart matrix.** The eigenvalue density distribution of the correlation matrix of the time series generated by the solved maximum-entropy model for the $S\&P500$. In red is the sampled density distribution from 1000 runs, in blue is the theoretical Marchenko Pastur distribution. The figure present a good agreement between the maximum-entropy model and the known analytical curve of Wishart matrix.

returns, where $-\infty < r_i(t) < \infty$, and the constraints are enforced over the whole matrix. Moreover, since the system is weighted we want to apply a normalization condition on the distribution of the returns.

We start with the simplest case of N time series with no constraints. However, contrarily to the binary (± 1) case considered in chapter 1, here we are required to enforce one extra constraint over the second moment

$$\langle \frac{1}{NT} \sum_{i}^{N} \sum_{t}^{T} r_i^2(t) \rangle \equiv [r_i^*(t)]^2_{average} \tag{3.3}$$

to ensure that the probability distribution for $r_i(t)$ is normalised. The Hamilto-

nian reads

$$H(\mathbf{X}, \alpha) = \alpha \sum_{i=1}^{N} \sum_{t=1}^{T} r_i^2(t) \tag{3.4}$$

with α as a single free parameter, which leads us to a normal distribution.

Using maximum likelihood estimation as before, one finds

$$\alpha^* = TN \left(2 \sum_{i=1}^{N} \sum_{t=1}^{T} \left[r_i^*(t) \right]^2 \right)^{-1} \tag{3.5}$$

as the value of α.

Once solved, this model is a maximum-entropy equivalent to the generic case, where one measures the correlation between N independent random time series for T time steps (the observed period). In this specific case, the resulting correlation matrix would be an $N \times N$ Wishart matrix, whose statistical properties are well-known [19, 21]. In the limits where $N, T \to \infty$ and $T/N \geq 1$ the eigenvalues of the Wishart matrix are distributed according to a Marchenko-Pastur distribution

$$P(\lambda) = \frac{T}{N} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda} \qquad \text{if} \qquad \lambda_- \leq \lambda \leq \lambda_+ \tag{3.6}$$

and $P(\lambda) = 0$ otherwise. The boundaries λ_+ and λ_- are dependent on the data size and given by

$$\lambda_\pm = \left[1 \pm \sqrt{\frac{N}{T}} \right]^2 . \tag{3.7}$$

This analytic curve represents the boundaries of the bulk eigenvalues, which predominantly represent noise, and so have little meaning assigned to them.

In Fig 3.1 we plot the eigenvalue density distribution of the correlation matrix of the time series generated by the maximum-entropy model for the $S\&P500$. In red is the empirical density sampled from 1000 runs, in blue is the theoretical Marchenko Pastur distribution eq.(3.6). It is clear that both distributions are almost identical, validating our claim of correspondence between the models when measuring the eigenvalues spectrum. We should note that although we use the parameter σ which is extracted by eq. (3.5), it does not play any role in the eigenvalue distribution. The distribution is invariant to different σ, which explains the agreement with the theoretical curve.

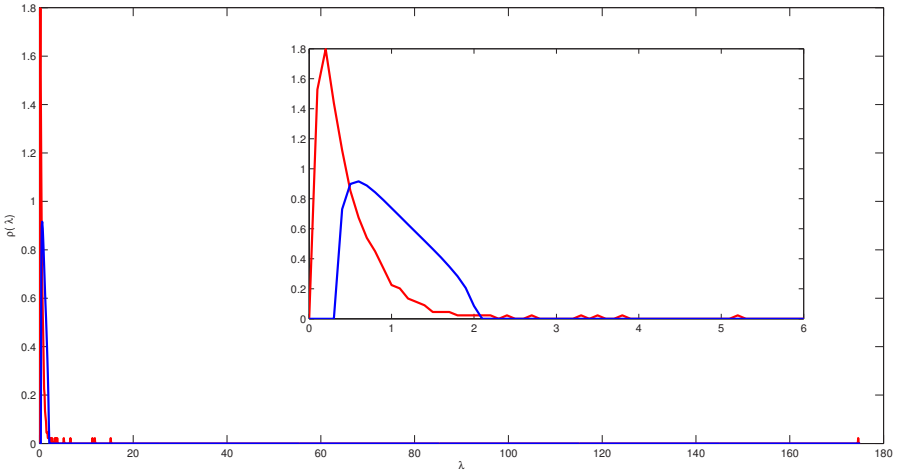

Figure 3.2: **Eigenvalues density distribution comparison: empirical correlation versus Wishart matrix.** The eigenvalue density distribution of the correlation matrix of the $S\&P500$. In red is the empirical eigenvalue distribution, in blue is the Marchenko-Pastur distribution. The figure shows the bad agreement of the Wishart matrix with respect to the empirical spectrum.

The empirical eigenvalues outside this range, however, have structural implications and relate to groups of correlated stocks [21]. As a result, any empirical correlation matrix C can be identified as a sum of of two matrices:

$$C = C^{(r)} + C^{(s)}, \tag{3.8}$$

where $C^{(r)}$ is the random part aggregated from the eigenvalues in the random spectrum $(\lambda_- \leq \lambda \leq \lambda_+)$, i.e.

$$C^{(r)} \equiv \sum_{i:\lambda_- \leq \lambda_i \leq \lambda_+} \lambda_i |v_i\rangle\langle v_i|, \tag{3.9}$$

and $C^{(s)}$ is the "structured" component, which is composed from those eigenvalues above the boundary of the bulk eigenvalues. $\lambda > \lambda_+$.

Moving forward, in financial markets, it is well established that stocks typically move up or down together, an effect known as the "market mode". This effect is indicated by the presence of a very large eigenvalue λ_m, orders of magnitude greater than the rest. This can be observed in Fig 3.2 where we compare the eigenvalue density distribution (of the cross-correlation matrix) of the $S\&P500$ index to the Marchenko-Pastur distribution. Since this eigenvalue represents a common

factor influencing all the stocks in a given market, from a structural perspective, the market mode eigenvalue signifies the presence of one single super-community, containing all the stocks in the market.

Thus, the other eigenvalues (not including the market mode), which deviate from the bulk, $\lambda_+ < \lambda_i < \lambda_m$ are the ones corresponding to mesoscopic clusters, i.e. groups of stocks with similar dynamics. This observation results in a further decomposition of the empirical correlation matrix

$$C = C^{(r)} + C^{(g)} + C^{(m)}, \tag{3.10}$$

where

$$C^{(m)} \equiv \lambda_m |v_m\rangle\langle v_m| \tag{3.11}$$

represents the market mode, and

$$C^{(g)} \equiv \sum_{i:\lambda_+<\lambda_i<\lambda_m} \lambda_i |v_i\rangle\langle v_i| \tag{3.12}$$

represents the remaining correlated groups. These sub-groups of correlated stocks comprise the mesoscopic structure of the market. They are also referred as "group modes" in the literature [1, 21].

To conclude, this approach actively filters both the null model spectrum and the market mode (largest eigenvalue) and has been introduced in [7]. However, there are a few drawbacks to this approach. Firstly, the global mode is not generated by the null model itself, and we are required to filter it "manually". The reason is that the null model only considers N random variables, and the presence of the market model is purely an empirical fact which is not being account for in the null hypothesis. Next, the random bulk distribution is only dependent on the size of the data (T and N), and not on the data itself. One can imagine that different systems would have different dynamics which requires different types of spectral filtering. Lastly, the null model cannot reproduce the so-called sub-random eigenvalues, where $\lambda_i < \lambda_-$, as we can see in Fig 3.2. This is quite alarming since the sub-random spectrum contains the majority of the empirical eigenvalues. In the next section, we will introduce a new null model that can overcome these limitations.

3.2.2 Random time series with global mode

Here we introduce a more sophisticated maximum-entropy null model, which enforce added "structural information", trying to overcome the random model limitations. The model is designed to enforce also the average daily return, hence

capturing the non-stationary nature of the system. Thus, the constraints are

$$\langle \frac{1}{NT} \sum_{i}^{N} \sum_{t}^{T} r_i^2(t) \rangle \equiv [r_i^*(t)]^2_{average} \qquad \langle \frac{1}{N} \sum_{i=1}^{N} r_i(t) \rangle \equiv \{r_i^*(t)\} \quad \forall t \tag{3.13}$$

recalling that we denote $\{r_i(t)\} \equiv \frac{1}{N} \sum_i r_i(t)$ as the average over stocks.

This step is motivated by section 1.6, where we impose the information separately for each day.

Combining all the above considerations, we finally generalize the model defined by eq.(1.59) to the matrix case as follows:

$$H(\mathbf{X}, \alpha, \vec{\theta}) = \alpha \sum_{i=1}^{N} \sum_{t=1}^{T} r_i^2(t) + \sum_{t=1}^{T} \theta_t \sum_{i=1}^{N} r_i(t) = \sum_{it} \left[\alpha r_i^2(t) + \theta_t r_i(t) \right] \tag{3.14}$$

where $\vec{\theta}$ it a T-dimensional vector with entries $\theta(t)$.

Using maximum likelihood estimation, in this case one needs to solve $T+1$ coupled equations

$$\alpha^* = TN \left(2 \sum_{i=1}^{N} \sum_{t=1}^{T} (r_i^*(t) - \{r_i^*(t)\})^2 \right)^{-1} \tag{3.15}$$

$$\theta_t^* = -2\alpha^* \{r_i^*(t)\} \tag{3.16}$$

in order to find the value of α and $\vec{\theta}$.

Once solved, this maximum-entropy model corresponds directly to a one-factor model defined in sec. 1.6.4. If we now define σ and μ as follows

$$\sigma^* \equiv \sqrt{\frac{1}{2\alpha^*}} \quad , \quad \mu_t^* \equiv \{r_i^*(t)\} \tag{3.17}$$

the probability of a matrix (multiple time series) is the product of Gaussian distributed random variables p_{it}

$$P(\mathbf{X}) = \prod_{it} \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{\frac{-1}{2\sigma^2} \sum_{\mathrm{it}} (\mathrm{r_i(t)} - \mu_\mathrm{t})^2}}_{p_{it}} \tag{3.18}$$

resulting from the integration (from $-\infty$ to ∞) of the exponential (partition function). The components of the multiple financial time series can thus be sampled as follows

$$r_i(t) \sim \mathbf{N}(\mu_t^*, \sigma^*), \tag{3.19}$$

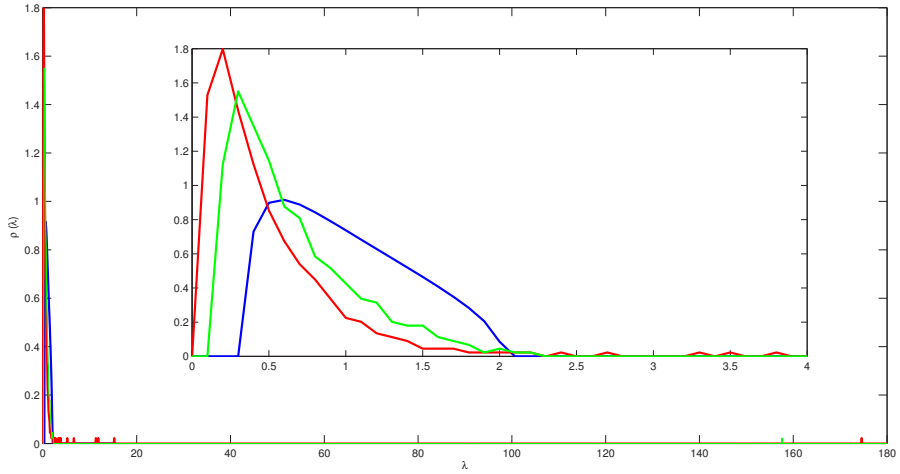

Figure 3.3: **Eigenvalues density distribution comparison: maximum-entropy model with global mode versus Wishart matrix.** The eigenvalue density distribution of the correlation matrix of the time series generated by the solved maximum-entropy model for the $S\&P500$. In red is the empirical eigenvalue distribution, in green is the sampled density distribution from 1000 runs (new null model), in blue is the Marchenko-Pastur distribution. The figure shows the spectrum of the improved null model with respect to the previous one, and how it can better replicate the random bulk and global mode that are observed in the empirical spectrum.

where $\mathbf{N}$ represents a random variable with normal distribution with mean μ_t^* specific to each time step and variance σ^*.
In Fig 3.3 we plot the eigenvalue density distribution (of the cross-correlation matrix) for the $S\&P500$ index. The green curve is the sampled eigenvalue distribution of the maximum-entropy model, and the red curve is the empirical eigenvalue distribution and the blue curve the Marchenko-Pastur distribution. We can see that the new null model can replicate the global mode (largest eigenvalue). Furthermore, there is a shift in the random bulk, which is in agreement with the empirical distribution. This results present an improved null model which is able to generate both the global mode and the random bulk of the empirical data,

$$\langle C \rangle_{me} = C^{(m)} + C^{(r)} \tag{3.20}$$

using partial information from the original data.

Returning to the modularity, we define the filtered empirical correlation matrix $C_{ij}^{(g)}$ filtering both the global mode $C^{(m)}$ and the random bulk $C^{(r)}$.

Once we input this result into the modularity eq.(3.2.2)

$$Q(\vec{\sigma}) = \frac{1}{C_{norm}} \sum_{i,j} \left[C_{ij} - \langle C_{ij} \rangle_{me}\right] \delta(\sigma_i, \sigma_j) = \frac{1}{C_{norm}} \sum_{i,j} (C_{ij} - C_{ij}^{(r)} - C_{ij}^{(m)}) \delta(\sigma_i, \sigma_j)$$

we see that this leads to

$$Q(\vec{\sigma}) = \frac{1}{C_{norm}} \sum_{i,j} C_{ij}^{(g)} \delta(\sigma_i, \sigma_j). \tag{3.21}$$

In other words, to clearly differentiate between the mesoscopic groups, one must subtract out the main drift of the system and the random correlation, using the maximum-entropy null model. The filtered matrix $C_{ij}^{(g)}$ constituted from the "non-random" eigenvalues $\lambda_+ < \lambda_i < \lambda_m$ and their corresponding eigenvectors v_i. The new method modified three modern community detection algorithms, customizing where necessary to be effective with correlation matrices [7]. The three algorithms we use in this paper are known as the Potts (or spin glass) method [12, 13], the Louvain method [14] and the spectral method [15]

3.3 Financial markets

Traditionally, the main object of time series analysis is the characterization of patterns in the amplitude of the increments of the quantities of interest (stock price in our case). The analysis requires a weighted description of the system, i.e. both the amplitude and the sign of the activity. Indeed, a time series of increments enclose complete information about the amplitude of the fluctuations of the original signal. However, a significant part of this information is encoded in the purely 'binary' projection of the time series, i.e. its sign. Recent studies have shown various forms of statistical dependency between the sign and the absolute value of fluctuations [8, 9, 10]. In chapter 1, we have shown a robust empirical relationship between binary and non-binary properties of real financial time series [11]. The results show that binary signatures, which retain only the sign of fluctuations, encode significant information regarding the full behaviour of the stock (both amplitude and direction). Motivated by these results, here we further explore the higher-order relations between financial time series and their corresponding binary signatures, in a more complex setting. In this section we study whether the binary signatures of assets can reproduce the same complex community organisation of financial markets, as the weighted information.

To this end, we use the daily closing prices of the stocks of three indexes (S&P500, FTSE100 and NIKKEI225) over the period 2001-2011. For each index, we restrict our sample to the maximal group of stocks that are traded continuously throughout the selected period. This results in 445 stocks for the S&P500, 78 stocks for the FTSE100 and 193 stocks for the NIKKEI225. Given a stock price

$P_i(t)$ where i denotes one of the N stocks in the index, and t denotes one of the T observed temporal snapshots (days), the log-return is defined as

$$r_i(t) \equiv \log\left[\frac{P_i(t)}{P_i(t-1)}\right], \tag{3.22}$$

where $2 < t < T$.
For each stock in the system we use the time series of it's log-returns for our analysis. This is the construct we refer to as the "weighted time series" throughout the rest of the chapter. In contrast, the "binary signatures" only reveal the direction of the fluctuation (sign) in the price and are defined as

$$x_i(t) \equiv \text{sign}[r_i(t)] = \begin{cases} +1 & r_i(t) > 0 \\ 0 & r_i(t) = 0 \\ -1 & r_i(t) < 0 \end{cases}. \tag{3.23}$$

In Fig. 1.1 from chapter 1 we show a simple example of a weighted time series, along with the corresponding binary projection.

The two types of information are in fact different descriptions of the same system, and are used to construct cross-correlation matrices. In turn, we deploy three popular community-detection algorithms [12, 13, 14, 15] specifically adapted, where necessary, for the correct use of cross-correlation matrices [7]. We examine and quantify similarities and variations in the organisation of the markets for these two representations. This approach reveals some interesting results. First, we can quantify the level of information encoded within the binary signatures, with respect to the full weighted time series. Secondly, we observe that both the binary and weighted representations yield very similar structures, which indicates that most of the information regarding the structure of financial communities is already encoded within the sign of a stock.

3.3.1 Spectral analysis

In this section we analyse the eigenvalue density distribution of the cross-correlation matrices for the two representations of the data (binary and weighted). When plotting the density distribution one can identify specific spectral properties that have structural implications. In other words, it is possible to identify distinct eigenvalues in the spectrum, which correspond to correlated clusters of stocks, and typically indicate a non-trivial structure (as defined in sec. 3.2.2) .

Our focus here is to detect these non-random eigenvalues in the spectrum of both the binary and weighted data. Moreover, we want to explore the similarities and differences between the two spectra to inform us about the corresponding structures yielded by each type of data.

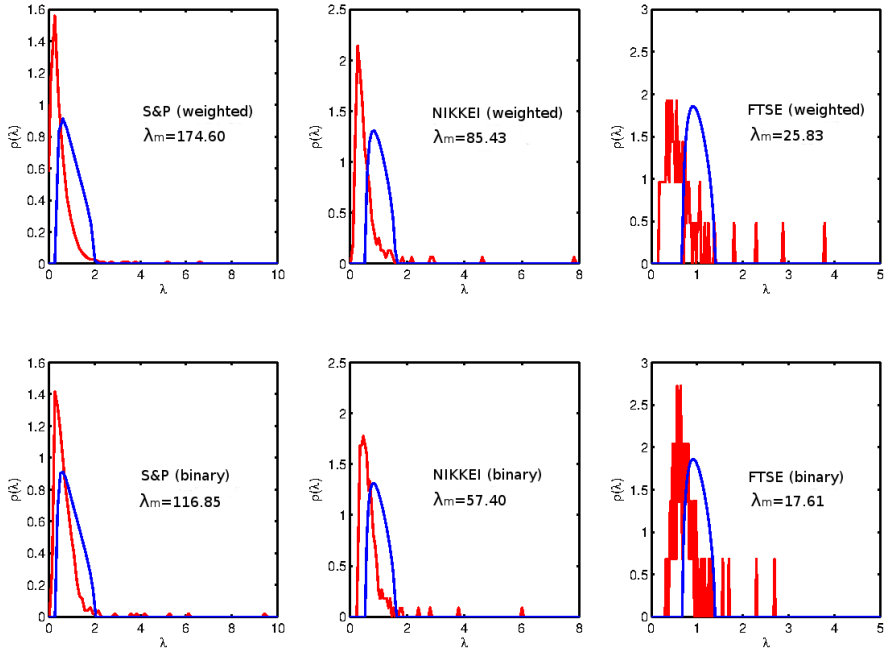

Figure 3.4: **Signatures of complex structure in the eigenvalue spectrum.**The eigenvalue density distribution (of the cross-correlation matrix) for the different indexes, where the upper panels are for the weighted series and the lower panels are for the binary series. The red curve is the empirical eigenvalue distribution and the blue curve the Marchenko-Pastur distribution. The largest empirical eigenvalue λ_m is not shown in the plots, but its value is reported in each panel. The figure shows that the complex structure, composed from global, group, and random modes, is also maintained in the binary representation.

In Fig. 3.4 we plot the eigenvalue density distribution for the three different indices. The top row corresponds to the weighted representation (log-returns), and the bottom row corresponds to the binary representation (binary signatures). We can observe the known structure of the financial markets in the weighted data, however this complex structure also exists in the binary data. This result is non trivial. We can observe a market mode, and several deviating eigenvalues also in the "simpler" binary data (with the same order of magnitude).

We also want to inspect whether both descriptions of the system function the same under randomization. The returns of each stock were separately permuted randomly, therefore preserving the total return of the stocks and destroying the daily correlation between the returns. Once the time series entries are shuffled, both binary and weighted correlation matrices end up as elementary random

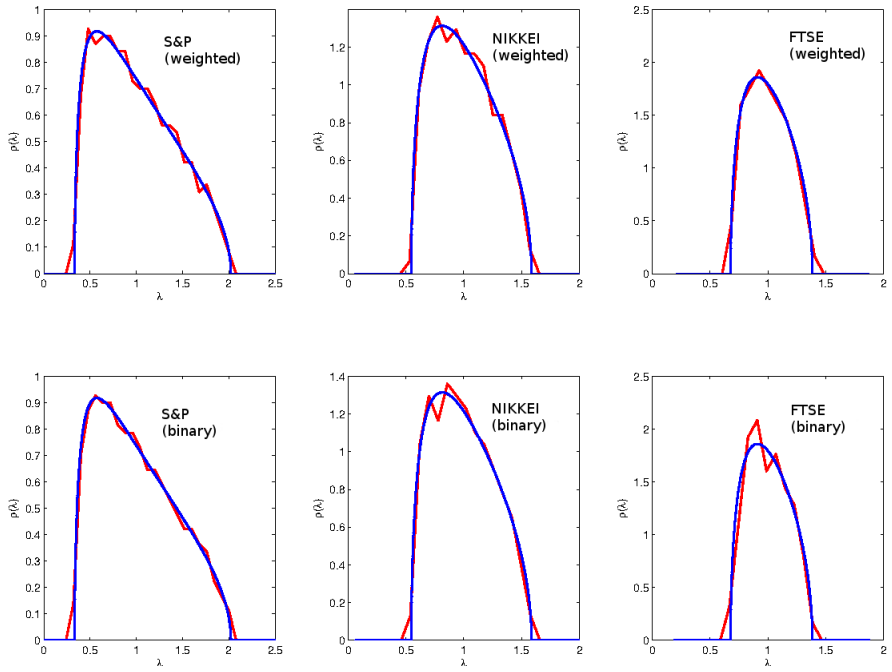

Figure 3.5: **Eigenvalues density distribution of randomized empirical data.** The eigenvalue density distribution of the Pearson correlation matrices where the upper panels are for the weighted series and the lower panels are for the binary series. The red curve is the empirical eigenvalue distribution and the blue curve the Marchenko-Pastur distribution. The figure shows the collapse to the random distribution once the original data is shuffled.

matrices. As discussed before, the eigenvalues of such matrices will be distributed with a Marchenko-Pastur distribution.

In Fig. 3.5 we plot the density distribution of the shuffled data for the three different indices. The top row corresponds to the weighted representation (log-returns), and the bottom row corresponds to the binary representation (binary signatures). As expected, in both cases we observed the known characteristics of a random matrix. The spectra of both representations collapsed to the known analytic curve.

To sum up this section, we identified a sub-group structure both in the weighted and the binary representation of the three indices. Each of the binary spectra we studied retain all the known properties of a "regular" (weighted) spectrum (random bulk, market mode and group modes). This result propels us to do a more refined analysis, and to further explore (and compare) the sub-group structure

Consumer Discretionary:		Consumer Staples:	
Energy:		Financials:	
Health Care:		Industrials:	
Information Technology:		Materials:	
Telecom. Services:		Utilities:	

Table 3.1: The 10 industry sectors in the Global Industry Classification Standard (GICS), with the colour representation used to highlight the sectors in the following Figures.

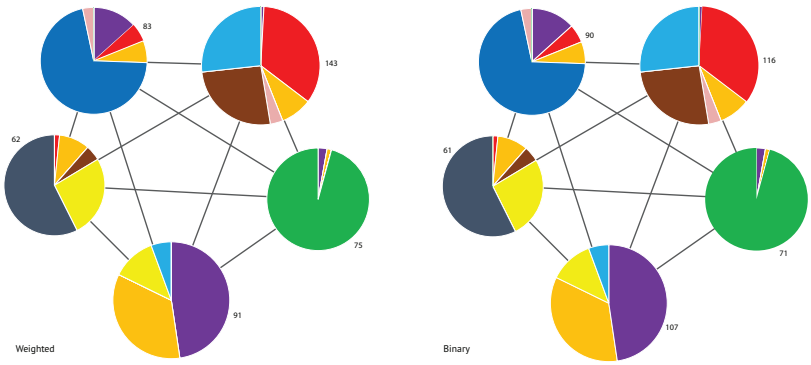

Figure 3.6: **S&P community structure.** Communities of the S&P 500 (daily closing prices from 2001 to 2011) generated using the modified Louvain algorithm [7]. Each community is labelled with the number of stocks and the pie chart represents the relative composition of each community based on the industry sectors of the constituent stocks (colour legend in Table 1). The inter-community link weights are negative, indicating that the communities are all residually anti-correlated. We can see that the binary partition is almost identical to the weighted one.

of the different indices. In the next section we will apply community detection algorithms to extract a more detailed structure for both representations, so that we can better quantify the similarities and the variations.

3.3.2 Community structure

In network theory, a community structure is the partition of the network into relatively dense sets of nodes, with respect to the rest of the network. More specifically, it is the organisation into clusters of nodes with dense connections internally, while the connections between the clusters are sparser. Community detection is the identification of such clusters of agents (nodes) in the system (network). In the last decade there has been a burst of research concerning this

topic, across a myriad of different fields [5].

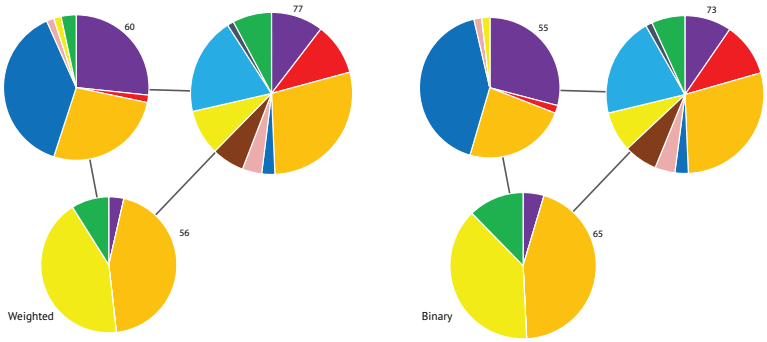

Figure 3.7: **Nikkei community structure.** Communities of the Nikkei 225 (daily closing prices from 2001 to 2011) generated using the modified Louvain algorithm [7]. Each community is labelled with the number of stocks, and the pie chart represents the relative composition of each community based on the industry sectors of the constituent stocks (colour legend in Table 1).The link weights are negative, indicating that the communities are all residually anti-correlated. We can see that the binary partition is almost identical to the weighted one.

This promising approach has also been applied to analyse time series data [22, 23, 24, 25] where the goal is to identify clusters of components with a similar dynamics. The attempts made so far have basically replaced network data with cross-correlation matrices as the input. However, this procedure suffers from a significant limitation. The null hypotheses used in the network-based algorithms are inconsistent with the properties of correlation matrices. As a result, these approaches can introduce an undesired bias when applied to the detection of communities in time series based networks.[7]. Here we adopt a new method [7], which is specifically shaped to deal with correlation matrices, based on the spectral properties we presented in the previous section. The method presents an improved and consistent way to cluster multiple time series, by leveraging a set of null models, specifically designed for use with correlation matrices.

Applying the new approach, we use three popular community detection algorithms, customizing where necessary to be effective with correlation matrices. The three algorithms we use in this chapter are known as the Potts (or spin glass) method [12, 13], the Louvain method [14] and the spectral method [15], and are modified in [7] to correctly deal with correlation matrices. The algorithms use a modularity optimization process, where modularity is a measure for how "optimal" your partition is. The algorithms attempt to choose a specific partitioning of the network into groups such that the corresponding value of the modularity is maximized (see Methods). The modularity implements a new null hypothesis, which

is fitted to time series (correlation matrices). More specifically, the hypothesis considers the empirical correlation matrix as a superposition of modes eq.(3.10), and decomposes it accordingly. Both the random mode and market mode are filtered out, and we are left with only the informative group mode, which is then used to extract the market structure.

Thus, we end up with three community detection algorithms that are consistent with time series data and represent the counterparts of the most popular techniques used in network analysis. The method allows us to explore the mesoscopic structure of different financial indices, and more specifically compare the different community structures resulting from the different representations (binary and weighted).

First, we perform community detection on both the binary and the weighted time series, using all three community detection methods, for the full time period of the data (2001-2011). We pick the division that maximizes the modularity (for a specific representation and algorithm), and compare the results for the two types of information. In the case where several divisions maximize the modularity (different runs result in different divisions), we take the division with the most occurrences over 1000 runs (the highest probability). Here, we want to identify groups with similar dynamics over the ten year period, with the rationale that such a long period will reduce the noise.

To help further explore the communities resulting from the different passes, we label each of the stocks according to its industry sector from the Global Industry Classification Standard (GICS). This classification represents a more "traditional" frame of mind where the different sectors are comprised of stocks conforming to a particular, qualitative description of the industry they represent. Recent results show that real markets have a more complex structure [2, 26, 27, 16], where different sectors are mixed in different sub-groups, i.e. the communities are assembled out of stocks from different sectors. Furthermore, the classification helps us to compare the results of the two representations in a very clear and visual way.

In Figs. 3.6 and 3.7 we plot the community structure of the S&P 500 and Nikkei 225 (daily closing prices from 2001 to 2011) generated using the modified Louvain algorithm. Each community is labelled with the number of stocks and the pie chart represents the relative composition of each community based on the industry sectors of the constituent stocks (colour legend in Table 1). The links between the communities represent "residual" (i.e filtered) anti-correlation relations [7]. We can see that the binary partition is very similar to the weighted one. For both indexes about $7 - 8$ percent of the stocks switch community. In the next section we will give a more quantitative measure for the dissimilarities of the different partitions.

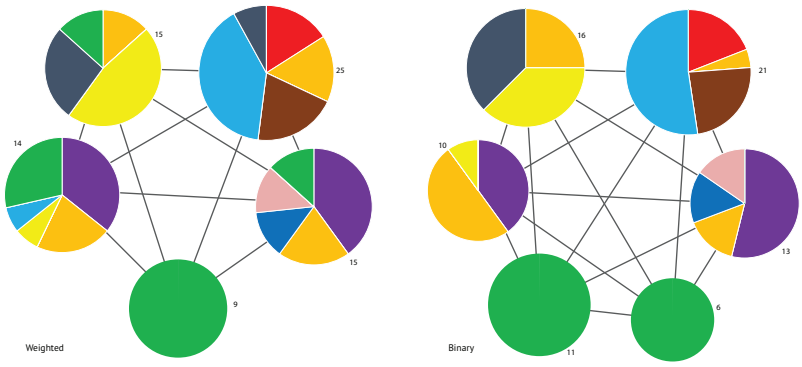

Figure 3.8: **FTSE community structure.** Communities of the FTSE 100 (daily closing prices from 2001 to 2011) generated using the modified Louvain algorithm [7]. Each community is labelled with the number of stocks, and the pie chart represents the relative composition of each community based on the industry sectors of the constituent stocks (colour legend in Table 1). The inter-community link weights are negative, indicating that the communities are all residually anti-correlated.

In Fig. 3.8 we observe a more complex result. Again, we plot the community structure of the FTSE100 (daily closing prices from 2001 to 2011) generated using the modified Louvain algorithm. Now, the binary representation consistently identifies one more community than the weighted representation (for all the algorithms). Later, we further explore these differences in community structure. Most notably, the binary information results in a cluster configuration where the Financials sector (green) was partitioned into two communities, whereas the weighted representation created only one sole Financials community, spreading the rest of the stocks among other clusters.

We should note that, purely from a community detection perspective, there is no "correct" partition. Each partition is generated from different data and so maximizing modularity should not be expected to yield the same partitions. We are comparing the end results of these processes, and in this setting (this chapter) we treat the weighted partition as the "truth" since it is using a priori more information to establish the partition. Thus, our aim is merely to examine the degree to which the "binary" community structure matches the "weighted" community structure, despite having less information for the algorithms to work with. That said, it would be interesting to determine if the binary information can yield different points of view, or added structural information with respect to its weighted counterpart. This would be useful in particular because binary time series are more robust to noise and errors in the data. Such a line of exploration is however

beyond the scope of this chapter.

In this section we showed that the binary description leads us to a very similar market structure to the weighted description. This results suggest that the information regarding the community partition is mainly encoded in the binary signature of the fluctuations, i.e. just from the knowledge of the direction of movement, one can practically reproduce the "correct" structure. In the next section we will quantify the similarities and deviations of the different partitions for the different algorithms. Furthermore, we will explore the evolution in time of the variations between the two representations.

3.3.3 Variation of information analysis

Index	Method	**Q** weighted	**Q** binary	Frequent VI	Switching stocks (%)	Minimal VI	Switching stocks (%)
S&P	Potts	0.4035	0.4134	0.3543	8.81%	0.3198	6.74%
	Louvain	0.4070	0.4134	0.3477	8.09%	0.3192	7.64%
	Spectral	0.4006	0.3932	0.6955	61.57%	0.6955	61.57%
NIKKEI	Potts	0.4551	0.4525	0.3689	6.78%	0.2598	4.66%
	Louvain	0.4551	0.4525	0.3711	7.25%	0.2604	4.15%
	Spectral	0.4481	0.4424	0.4521	8.29%	0.4521	8.29%
FTSE	Potts	0.4641	0.4988	0.5031	28.42%	0.4026	26.14%
	Louvain	0.4635	0.4988	0.4995	21.79%	0.3981	17.95%
	Spectral	0.4597	0.4903	0.6919	69.23%	0.6919	69.23%

Table 3.2: The Variation of Information measured between the binary and weighted partitions, with the maximal modularity **Q**, for the period 2001-2011. "Frequent VI" is the variation of information between the most common partitions (that maximize the modularity), and "Minimal VI" is the variation of information between the most similar (that maximize the modularity). The "Switching stocks" is the percentage of stocks that moved to different communities.

Once we obtain the community structure using the different algorithms, our goal is to quantify the dissimilarities (or similarities) between the different partitions (binary and weighted). For this task we apply the Variation of Information (VI) measurement [28, 29]. The variation of information is an information-theoretic measure of the distance between two partitions. The different partitions $\vec{\boldsymbol{\sigma}}^1$ and $\vec{\boldsymbol{\sigma}}^2$ represent N-dimensional vectors where the i-th component σ_i denotes the set in which node i is placed by that particular partition.

The variation of information involves the mutual information $I(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2)$ which is defined as

$$I(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2) = \sum_{i=1}^{N}\sum_{j=1}^{N} p(\sigma_i^1, \sigma_j^2)\log\left(\frac{p(\sigma_i^1, \sigma_j^2)}{p(\sigma_i^1)p(\sigma_j^2)}\right), \tag{3.24}$$

where $p(\sigma_i^1, \sigma_i^2)$ is the joint probability distribution, and $p(\sigma_i^1)$ the marginal distribution of σ_i^1. The mutual information measures the overlap between the two partitions, however it is not a metric (does not obeys the triangle inequality) nor is it normalized. Thus, for this study we use the (normalized) variation of information which is defined as

$$VI(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2) = 1 - \frac{I(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2)}{H(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2)} \tag{3.25}$$

where $H(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2)$ is the joint entropy and is defined as

$$H(\vec{\boldsymbol{\sigma}}^1 : \vec{\boldsymbol{\sigma}}^2) = \sum_{i=1}^{N} \sum_{j=1}^{N} p(\sigma_i^1, \sigma_j^2) \log \left(p(\sigma_i^1, \sigma_j^2) \right). \tag{3.26}$$

The variation of information ranges from 0 to 1, where 0 indicates two identical partitions, and 1 a complete dissimilarity between the partitions.

First, we measure the variation of information between two partition vectors, generated by the weighted and binary time series. Respectively, this approach enables us to quantify the difference in group structure (for 2001-2011), and compare the performances of the different algorithms.

In Table 3.2 we plot the measured VI between the different partitions, which resulted from the binary and weighted data. These measurements are for the community structures resulting from 10 years (2500 time steps) of data, for each of the three different indices. We run the algorithms 1000 times (for Louvain and Potts, while the spectral is deterministic) and extract the partitions that maximize the modularity. We measure the VI between two different partitions: the most frequent one, and the one that minimizes the VI (the most similar ones). One can consider this to be the best result (subject to the best partition). Again we should note that all the partitions maximize the modularity, and therefore are optimal. Since the VI is not linear (and not intuitive), we also included a simpler measurement of the percentage of stocks that are not occupying the same community in the different partitions.

We can observe that both the Potts and Louvain algorithms consistently perform better then the spectral method, i.e. they yield partitions with higher modularity. We should note that, one can only compare the value of the modularity for the same representation (binary or weighted), while there is no meaning in comparing modularity between the different representations. Generally, we can observe that the binary information and the weighted information result in very similar structures, as we showed in the previous section. The exception to this is the FTSE index, where the binary information consistently yields a greater number of communities. It is interesting to note that the binary information always results in either the same or a greater number of communities over the weighted time series.

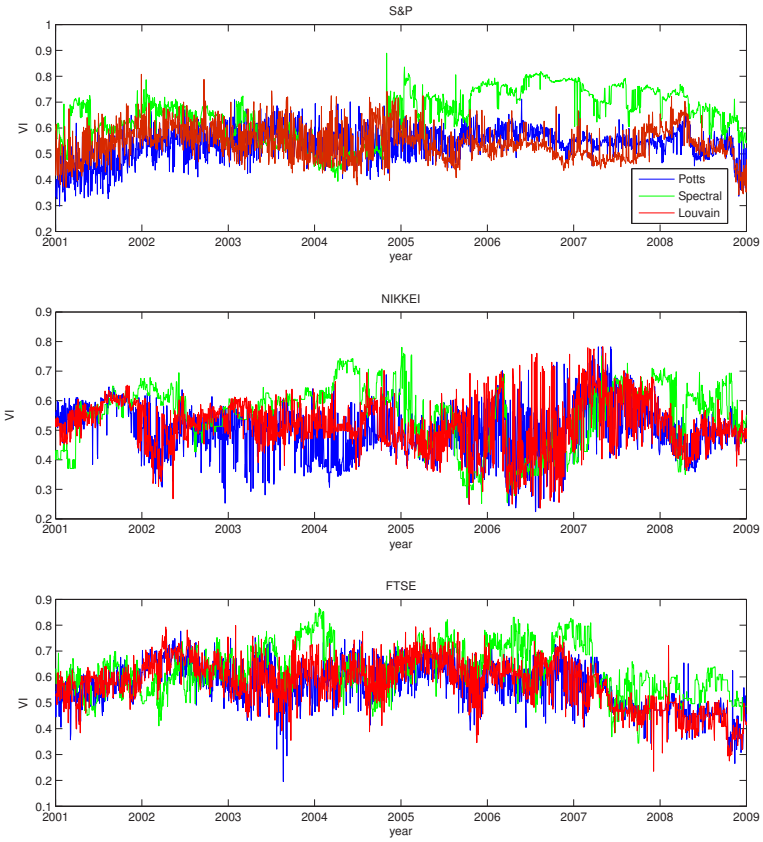

Figure 3.9: **Variation of information analysis for the different clustering algorithms.** variation of information between the binary and weighted partitions for a sliding window of 600 trading days (approximately 28 moths) starting at Q3 2001. The VI is measured between the frequent partitions for the different algorithms: Potts (blue), Louvain(red) and Spectral (green).

Next, we will explore the evolution in time of the VI between the different types of information. We considered a sub-period of 600 time steps (about two and a half years), and apply the same procedure as before. However, here we use a sliding window technique, where in each step we input a new day and ignore the previous information. This results in 1900 time steps (from the original 2500), where each point is the frequent VI calculated from the correlation matrix using

the given 600 time step. In Fig. 3.9 we plot the different measurements for the different algorithms. The Potts and the Louvain methods present a more stable dynamics, while the Spectral method yield higher VI. Furthermore, there is no systematic effects of the financial crisis on the similarity between the binary and the weighted representations.

The analysis reveals that in financial markets both the binary and weighted information yield very similar structures, which also manifest themselves in similar spectral properties. The algorithms find complex mesoscopic structure of internally correlated clusters, which are residually anti-correlated with each other [7]. Moreover, the clusters are populated by stocks from various sectors. Remarkably, we show that the simple knowledge of the direction of increments of each stock can reproduce this complex structure very successfully.

3.4 Functional brain networks

In this section, we move from financial markets to brain systems. Remarkably, from a pure "network science perspective", these different systems share very similar features. Their dynamics is monitored in terms of multiple time series of activity of the constituent units, such as stocks or neurons respectively. In both systems, we have no complete information or exhaustive understanding about the underlying mechanisms at play, and we resort to information extraction from observations and empirical measurements. This illustrates why the use of correlation matrices is so popular both in financial systems and in neuroscience. Resolving the structure of brain networks from time series data (via the calculation an empirical correlation matrix), exactly as we did for financial markets in the previous section, is the object of the research field that goes under the name of 'functional brain networks.' In neuroscience, functional networks are defined by the correlated dynamical activity of neurons, as opposed to structural networks which are instead defined as the real physical connections between neurons [32, 33, 34].

In functional brain networks, the mesoscopic structure is the key intermediate level of organisation bridging the microscopic dynamics of individual neurons with the macroscopic dynamics of the brain as a whole. At this mesoscopic level, brain activity tends to be organized in a modular way [41], with functionally related units being positively correlated with each other, while at the same time being relatively less (or even negatively) correlated with dissimilar ones. Such emergent organisation is mainly detected through the aforementioned functional networks. In this setting, the network displays the correlated dynamical activity between pairs of units in the brain. Dependent on the type and resolution of the data, the units can vary from single neurons to brain regions. Over the last years, the statistical characterization of these functional networks developed into a very

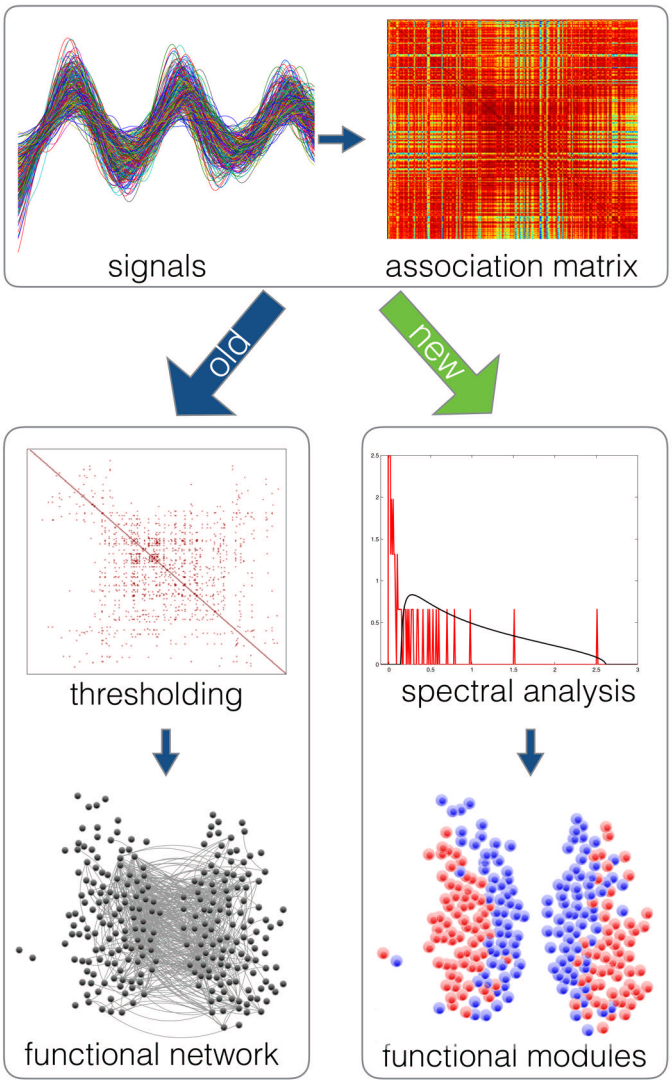

Figure 3.10: **Functional modules without functional network, a new method to identify emergent organisation.** A compression between the general steps of functional network construction and our new method. On the left scheme, we can see the introduction of a threshold procedure, once the majority of the data is neglected one can create a network representation. On the right scheme, we apply our spectrum filtering by comparing the empirical data to the random null model. By directly producing a partition of the original time series into communities, our new method bypasses the functional network projection, avoiding the use of a threshold procedure.

popular and powerful tool to study brain organisation [33, 35, 36].

Recent studies have used functional networks analysis to quantify global and local properties of brain networks [43, 44]. The use of functional networks yields very promising results. They are expected to encode the physiological mechanism for information processing and mental representations [37, 38, 39, 40]. Moreover, various case studies have shown that neurological and psychiatric disorders, such as Alzheimer's disease and schizophrenia, lead to variations in the network topology [45, 46]. Despite these encouraging results, it is well known that the current approaches suffer from specific methodological problems [42]. Firstly, the methods suffer from an inevitable information loss induced by the thresholding procedure. Another, less realized, limitation is the bias introduced by the use of network-based (as opposed to correlation-based) community detection methods [7]. Finally, the presence of (anti)correlated modules is obfuscated by the presence of a global mode of brain activity which imparts an overall positive correlation, a problem that becomes particularly evident when searching for communities in small brain regions, where the global signal is strong.

As we demonstrated in the previous section, our maximum-entropy approach can overcome precisely these methodological limitations. Furthermore, our method is not limited to a set of random walks with a common global factor. In financial markets, which we analysed in the previous section, the global factor translates to the similar reaction of stocks to news or events (the known market mode). However, once generalized, our new null model as described in section 3.2.2 can account for various systems with a global mode. In biological systems, the global mode can be a result of the units sharing a similar environment, or being affected by the same periodic signal (e.g. blood pressure or heartbeat rates). Remarkably, the removal of the global mode in our method entails by definition the removal of all common factors which affect the units of the system. This filtering capability demonstrates why our method has a great potential in the biological setting.

3.4.1 The suprachiasmatic nucleus

As we mentioned, we analyse the suprachiasmatic nucleus (SCN), located in the hypothalamus in the brain. The SCN is a complex structure comprising of a heterogeneous collection of genetically and electrically rhythmic neurons that synchronize their activity to the external world to form a central pacemaker. This complex network needs to produce stable output signals and at the same time remain flexible to changes in the environment. Thus, it is essential for the regulation of our daily and seasonal rhythms. The network is organized in a complex way for it to coordinate these daily rhythms at the tissue level [47], where the central pacemaker is a 24 hour rhythm that is synchronized to the external light-dark cycle. In response to a shift in the external cycle, neurons of the SCN resyn-

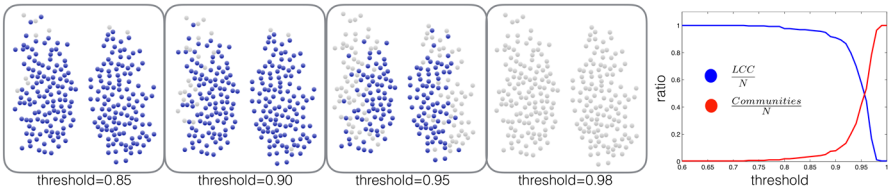

Figure 3.11: **The community structure of the SCN as resolved by a standard threshold approach.** On the left, we plot the community structure, resolved by the standard method, for different thresholds. In blue are the nodes that belong to the large cluster, while in gray are isolated nodes (communities composed out of one node). In the right panel, we plot the fraction of nodes in the largest connected component $\frac{LCC}{N}$ in blue, and the fraction of communities detected $\frac{Communities}{N}$ in red.

chronize with a different frequency. Recent studies have shown that the neuronal network organisation of the SCN changes in different seasons [48], however, the mechanisms behind these changes are still elusive. Currently, not much is understood about emergent organisation of the SCN, and if it is shaped by the external light-dark cycle. Here, we apply our community detection method to identify the functional community structure in different light-dark conditions.

This set-up investigates gene transcription activity, on a cellular scale, in SCN slices of male mice. This process is part of the molecular clock and is rhythmically expressed with a period of 24 hours. The target protein, which we monitor, is fused with the light emitting firefly luciferase. For each cell in the sample, the luminosity levels are recorded, generating the gene transcription "activity" time series of the units in the system. We should note that this "single cell resolution" of the data is unique for functional analysis. Currently, brain networks are most often derived from data acquisition techniques that do not have the possibility to perform recordings at the single cell level. Techniques such as Functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG) or Electroencephalography (EEG) use brain regions as nodes in the network and fiber bundles between these regions as edges. Our setup enables us to investigate a brain network at the micro-scale where nodes are single cells and edges are functional connections between the cells. The bioluminescence time series were obtained from individual cells (for exact method REF: [50]). In the next section, we apply the popular standard methods for constructing functional networks, highlighting their known limitations.

3.4.2 Standard approach to functional networks

Before we present the output of our method, we preliminary preform a standard analysis based in the mainstream method for detecting communities via functional networks. This is a useful reference as a comparison with our own method. In Figure 3.10 we present the standard ("old") approach versus our new approach. Once we measure the empirical correlation matrix, the conventional method requires a threshold value to exclude negative values and to neglect "weak" links simultaneously. Despite the significant loss of information, this thresholding procedure is an integral part of all current methods. Next, we adapt the thresholded matrix as an adjacency matrix, which corresponds to a network structure. The resulting network can either be a binary one, where all links are 1 or 0, or a weighted one, where the values from the correlation matrix remain as the weights of the links. Finally, we can apply various community detection methods [12, 13, 15, 14], to detect the community structure. In this analysis, we use the Louvain method [14].

In Figure 3.11 we present the community structure, resolved by the standard method, for different thresholds. In blue are the nodes that belong to the large cluster, while in gray are isolated nodes (communities composed out of one node). In the right panel, we plot the fraction of nodes in the largest connected component $\frac{LCC}{N}$ in blue, and the fraction of communities detected $\frac{Communities}{N}$ in red. It is evident that applying different thresholds essentially detaches isolated nodes from the large cluster, and there is no optimal value for the threshold. Therefore, the standard method can only observe a "radial gradient" of connectivity, and there is no large scale left-right symmetry, which is one of the signatures of functional as opposed to structural connectivity. This poor performance of the method is a known limitation when applied to very dense networks.

In the next section, we employ our new method described in section 3.2, which results in the new scheme portrayed in Figure 3.10. By directly producing a partition of the original time series into communities, our method bypasses the functional network projection, avoiding the use of a threshold procedure. Also, the method removes noise and filters out the common trend in the system.

3.4.3 SCN analysis

In Figure 3.12 we present the community structure detected by different recursive runs of the method. In the bottom panels are the partitions detected, where each community is marked with a different colour. In the top panels are the corresponding resolved filtered correlation matrices (for each run) displaying the resolved structure as a block matrix. In the left panel, we can observe that all the nodes in the system belong to one community when the global mode is not filtered out yet. Next, once we filtered out the global mode we found a consistent core-periphery structure. This result is robust to all the samples we have

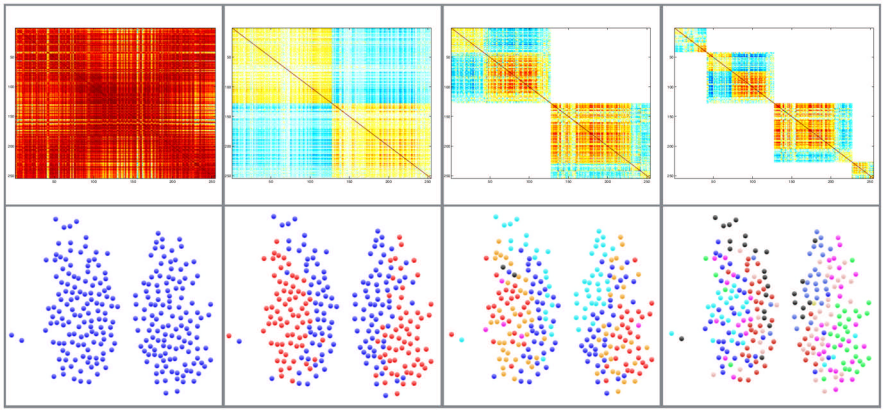

Figure 3.12: **The hierarchical community structure of the SCN as resolved by our method.** The community structure detected by different recursive runs of the method. In the bottom panels are the partitions detected, where each community is marked with a different colour. In the top panels are the corresponding resolved filtered correlation matrices (for each run) displaying the resolved structure as a block matrix.

analysed. Further iterations of our method within each of the identified modules, can detect hierarchy in the community structure. Note that for the second run the functional modules still retain the left-right symmetry, which implies a real functional structure. However, the hierarchical decomposition of the two clusters is difficult to interpret, because the resulting sub-clusters are small in size, which troubles reliable detection of significant differences between these sub-clusters. As a result, in this chapter, we will focus on the first partition that our method yields since it is providing the most impressive and consistent results.

The core-periphery structure confirms the ventral-dorsal distinction within the SCN. The ventral part of the SCN receives light input and adjusts quickly to changing light schemes, while the dorsal part of the SCN lags behind [51]. Our approach is able to identify the two communities in all the different experimental conditions, i.e. different the light-dark conditions simulating summer conditions (long days, short nights: L16D8) and winter day conditions (short days, long nights: L8D16). This results present a very robust functional structure, which is not dependent on external conditions. However, a more detailed analysis of the signals within the communities presented some variations in the distribution of the signals (within the communities) which depend on the light-dark cycles [50]. These findings motivate further research into this functional structure, and

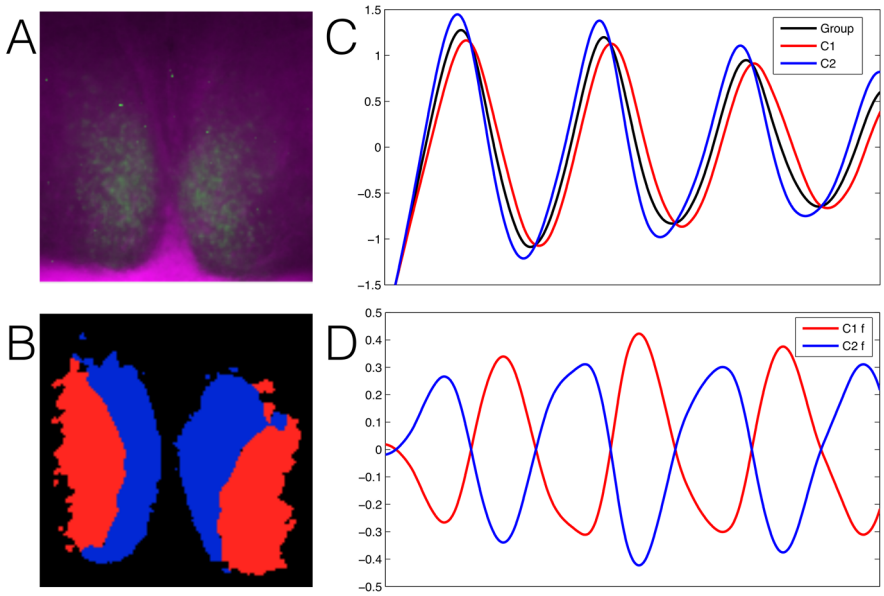

Figure 3.13: (A) The bioluminescence image of one SCN sample. (B) The plotted average partition over all the samples. (C) The plotted average signal of the whole system (in black) versus the mean signals of the two detected communities (in red and blue). (D) The plotted average residual signals of the two communities, once the global signal is subtracted.

possibly the next hierarchical partition.

We should note that regional analyses of the SCN has been performed by other groups. Evans and co-workers used a similar approach to identify single-cell-like regions of interest, but do not use clustering algorithms and choose the regions by hand [52]. Silver and co-workers also used regions of interest, called superpixels, but these were not necessarily identified as single cells. Based on these superpixels they use threshold methods to determine regional differences in the SCN [53, 54]. Abel and co-workers used a threshold method based on mutual information on single-cell-like regions of interest [55]. However, all studies encounter the known limitations as described in the previous section using the threshold method. They only find one large cluster and many isolated cell-like communities (in the shell or dorsal part).

In Figure 3.13 we present the overall findings our method provides. In panel A we can observe the experimental setup and the original sample with the illuminating neurons. In panel B we plot the average partition over all the samples

(approximately 40); these results highlight the robustness of our findings and the clear core-periphery partition. In C we plot the average signal of the two communities (blue and red) with the mean signal of the whole system (black). In panel D we plot the deviations of each community signal from the common signal. We found a perfect anti-correlation with respect to the global mode. Our method is indeed guaranteed to resolve mutually anti-correlated modules, as the result of the maximization of the modularity. Note however, that the anti-correlation is defined in terms of the "residual" signals with respect to the average system-wide signal (see [7]).

3.5 Conclusions

Over the last few years community detection methods have revealed themselves as useful tools to study the structure of complex systems. In this chapter, we first have introduced a new approach aiming at analysing structural dependencies, which result from different descriptions (weighted and binary activity) of a complex system. Our approach enables us to quantify the level of "structural information" encoded within the binary projection of weighted time series, and measure variations and similarities between the different partitions.

Our findings suggest that the binary signatures of financial time series carry significant structural information. These results are far from trivial, as one might expect that the full knowledge of the amplitudes of price fluctuations is a key component in clustering the markets into correlated groups. However, here we explicitly showed that purely binary information can replicate the main features obtained from complete information. Thus, we conclude that the key features of the market structure are induced by the binary dynamics of the stocks. Even when the two representations differ by some extent, the binary description provides very sensible information (as exemplified by the Financials sector in the FTSE).

Typically, the binary signature of a time series is obtained as a projection from the full weighted information, and is therefore known only if the latter is also known. This means that there are not many practical situations in which the full weighted information is unknown, while the binary projection is known. However, what can occur quite often (and indeed typically occurs) is that the weighted amplitudes of time series increments, much more than their signs, are affected by noise or errors. For instance, in many financial institutions (e.g. hedge funds or investment banks) there are specific departments in charge of cleaning the data and verifying their reliability. Generally, errors in the data are detected in the form of increments with anomalously large or small absolute value. By contrast, the sign of the increment itself is very robust to errors. Therefore our results indicate that analyses based on binary projections are likely to be much more robust

to noise than analyses based on the full original data. Since the extraction of the binary signatures is an extremely simple procedure, while the verification of the quality of weighted data can be very demanding and time consuming, our findings suggest that, at least for the purpose of identifying non-trivially correlated groups of stocks, the simpler procedure can be safely adopted.

Alongside the financial markets analysis, we have also applied our method to brain data. This step was possible due to the maximum-entropy generalization of the null model 3.2.2. By enforcing the total increments of each time step in the data, under the maximum-entropy formalism, we were able to expand our method to any system of signals with a common global factor, i.e. not be limited to a system with N random walks with an ad hoc market mode. Moreover, the new null model uses the optimal amount of information (eigenvalues) based on the particular data structure and not the data size (i.e. N and T) as before. In this general setting, the method resolves the partition of communities with maximum mutual anti-correlation, with respect to the global signal.

From a neuroscience perspective, our threshold-free method has a high potential to refine the search for functional modules in the brain. The fact that the approach does not require a "traditional" functional network representation is enabling the use of complete data sets and not just parts of it. Another aspect of our method that has significant implications for the research of brain networks is the multi-resolution quality. Since our method can recursively filter any global factors that are present in the system, we were able to detect single-cell-like regions of interest. This result pushes the resolution limit beyond the state-of-the-art methods, when identifying functional modules.

We applied our method to the SCN, revealing a core-periphery functional structure of two communities. Once the global signal is filtered from the system, the two clusters present a distinct anti-correlated dynamics. The global signal in this system corresponds to the common 24 hours cycle in all the cells, where the two communities describe a phase leading and a phase lagging community. This core-periphery structure confirms the ventral-dorsal distinction within the SCN. The ventral part of the SCN receives light input and adjusts quickly to changing light schemes, while the dorsal part of the SCN lags behind. Our community detection approach enhances the identification and the subsequent functional characterization of neuronal clusters in the SCN, possibly paving the way for more elaborate network analysis on the level of single cells in other brain regions.

Bibliography

[1] Sinha S, Chatterjee A, Chakraborti A, Chakrabarti BK. *Econophysics: An Introduction*, Wiley-VCH, Weinheim; 2010.

[2] Bouchaud JP, Potters M. *Theory of Financial Risk and Derivative Pricing*, Cambridge University Press, 2nd ed; 2003.

[3] Mantegna RN, Stanley HE. *Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press; 1999.

[4] Mantegna RN. *Hierarchical Structure in Financial Markets*, Eur Phys J B, 1999; 11, 193.

[5] Fortunato S. *Community detection in graphs*, Phys Rep. 2010; 486, 75.

[6] Newman MEJ. *Networks, an Introduction*, Oxford University Press; 2010.

[7] MacMahon M, Garlaschelli D. *Community detection for correlation matrices*, Phys Rev X. 2015; 5, 021006.

[8] La Spada G, Farmer J, Lillo F. *The non-random walk of stock prices: The long-term correlation between signs and sizes*, Eur Phys J B. 2008; 64, 607-614.

[9] Petersen AM, Wang F, Havlin S, Stanley HE. *Quantitative law describing market dynamics before and after interest rate change*, Phys Rev E. 2010; 81, 066121.

[10] Boguna M, Serrano MA. *Generalized percolation in random directed networks*, Phys Rev E. 2005; 72, 016106.

[11] Almog A, Garalaschelli D. *Binary versus non-binary information in real time series: empirical results and maximum-entropy matrix models*, New J Phys. 2014; 16, 093015.

[12] Reichardt J, Bornholdt S. *Detecting Fuzzy Community Structures in Complex Networks with a Potts Model*, Phys Rev Lett. 2004; 93, 218701.

[13] Reichardt J, Bornholdt S. Ntatistical, *Mechanics of Community Detection*, Phys Rev E. 2006; 74, 016110.

[14] Blondel JVD, Guillaume J, Lambiotte R, Lefebvre E. *Fast Unfolding of Communities in Large Networks*, Journal of Statistical Mechanics: Theory and Experiment. 2008; 10008.

[15] Newman MEJ. *Modularity and Community Structure in Networks*, Proc Natl Acad Sci USA. 2006; 103, 8577.

[16] Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Guhr T, Stanley HE. *Random Matrix Approach to Cross Correlations in Financial Data*, Phys Rev E. 2002; 65, 066126.

[17] Wigner EP. *Characteristic Vectors of Bordered Matrices With Infinite Dimensions*, The Annals of Mathematics. 1955; 62, 548.

[18] Mehta ML. *Random Matrices*, Elsevier; 2004.

[19] Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Stanley HE. *Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series*, Phys Rev Lett. 1999; 83, 1471.

[20] Laloux L, Cizeau P, Bouchaud JP, Potters M. *Noise Dressing of Financial Correlation Matrices*, Phys Rev Lett. 1999; 83, 1467.

[21] Heimo T, Kumpula JM, Kaski K, Saramaki J. *Noise Dressing of Financial Correlation Matrices*, Phys Rev Lett. 1999; 83, 1467.

[22] Heimo T, Kumpula JM, Kaski K, Saramäki J. *Detecting Modules in Dense Weighted Networks with the Potts Method*, Journal of Statistical Mechanics: Theory and Experiment. 2008; 08, P08007.

[23] Fenn DJ, Porter MA, Mucha PJ, McDonald M, Williams S, Johnson NF, Jones NS. *Dynamical Clustering of Exchange Rates*, Quantitative Finance. 2012; 10, 1493.

[24] Isogai T. *Clustering of Japanese Stock Returns by Recursive Modularity Optimization for Efficient Portfolio Diversification*, J Complex Networks. 2014; 2, 557.

[25] Piccardi C, Calatroni L, Bertoni F. *Clustering Financial Time Series by Network Community Analysis*, Int J Mod Phys. 2011; 22, 35.

[26] Utsugi A, Ino K, Oshikawa M. *Random Matrix Theory Analysis of Cross Correlations in Financial Markets*, Phys Rev E. 2004; 70, 026110.

[27] Potters M, Bouchaud JP, Laloux L. *Financial Applications of Random Matrix Theory: Old Laces and New Pieces*, Acta Physica Polonica B. 2005; 36, 2767.

[28] Meila M. *Comparing Clusterings, An Information Based Distance*, Journal of Multivariate Analysis. 2007; 98, 873.

[29] Meila M. *Comparing Clusterings by the Variation of Information*, Learning Theory and Kernel Machines. 2003; 2777, 173.

[30] http://www.lorentz.leidenuniv.nl/ garlaschelli/.

[31] M. MacMahon, http://www.mathworks.com/matlabcentral/fileexchange/49011.

[32] Eguiluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV. Scalefree *brain functional networks*, Phys Rev Lett. 2005; 94, 018102.

[33] Bullmore E, Sporns O. *Complex brain networks: graph theoretical analysis of structural and functional systems*, Nature Reviews Neuroscience. 2009; 10, 186–198.

[34] Park H, Friston k. *Structural and Functional Brain Networks: From Connections to Cognition*, Science. 2013; 342, 1238411-1238411.

[35] Singer, W. *Neuronal synchrony: a versatile code for the definition of relations?*, Neuron **24**, 49–65, (1999).

[36] Fries, P. *A mechanism for cognitive dynamics: neuronal communication through neuronal coherence*, Trends Cogn. Sci. **9**, 474–480 (2005).

[37] Bressler, S. L. *Large-scale cortical networks and cognition*, Brain Res. Brain Res. Rev. **20**, 288–304 (1995).

[38] Mesulam, M. M. *From sensation to cognition*, Brain **121**, 1013–1052 (1998).

[39] McIntosh, A. R. *Towards a network theory of cognition*, Neural Netw. **13**, 861–870 (2000).

[40] Buzsáki, G. *Rhythms of the Brain*, (Oxford Univ. Press, New York, 2006).

[41] Meunier, D., Lambiotte, R., Fornito, A., Ersche, K.D., Bullmore, E.T. *Hierarchical modularity in human brain functional networks*, Front. Neuroinformatics **3**, 37.

[42] RubinovaM. Sporns O. *Weight-conserving characterization of complex functional brain networks*, NeuroImag, **56**, Issue 4, 2068–2079, (2011).

[43] Stam, C.J., Reijneveld, J.C. *Graph theoretical analysis of complex networks in the brain*, Nonlinear Biomed. Phys. **1**, 3, (2011).

[44] Rubinov, M., Sporns, O. *Complex network measures of brain connectivity: Uses and interpretations* Neuroimage **52**, 1059–1069, (2010).

[45] Supekar, K., Menon, V., Rubin, D., Musen, M., Greicius, M.D. *Network analysis of intrinsic functional brain connectivity in Alzheimer's disease*, PLoS Comput. Biol. **4**, e1000100 (2008).

[46] Lynall, M.E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., Muller, U., Bullmore,E. *Functional connectivity and brain networks in schizophrenia*, J. Neurosci. **30**, 9477–9487 (2010).

[47] Liu, A. C., Welsh, D. K., Ko, C. H., Tran, H. G., Zhang, E. E., Priest, A. A. *Intercellular coupling confers robustness against mutations in the SCN circadian clock network*, Cell, **129(3)**, 605-616 (2007).

[48] VanderLeest, H.T., Houben, T., Michel, S.,Deboer, T.,Albus, H., Vansteensel, J.M., Block, G.D, and Meijer, J.H. *Seasonal Encoding by the Circadian Pacemaker of the SCN*, Current Biology, **17**, Issue 5, 468-473 (2007).

[49] Rohling J.H.T., vanderLeest, H.T., Michel, S., Vansteensel, M.J., Meijer, J.H. *Phase resetting of the mammalian circadian clock relies on a rapid shift of a small population of pacemaker neurons*, PLoS ONE **6**, 2011, e25437.

[50] Buijink, R., Almog, A., Wit, C.B., Roethler, O., Garlaschelli, D., Meijer, J.H., Rohling, J.H.T, Michel, S. *Exposure to long photoperiods induces changes in coupling between single neurons of the mouse suprachiasmatic nucleus*, PLoS ONE, accepted pending (2016).

[51] Albus, H., Vansteensel, M.J., Michel, S., Block, G.D., Meijer, J.H. *A GABAergic mechanism is necessary for coupling dissociable ventral and dorsal regional oscillators within the circadian clock*, Curr Biol ,**10** :886-93 (2005).

[52] Evans J.A., Leise T.L., Castanon-Cervantes O., Davidson A.J. *Intrinsic regulation of spatiotemporal organization within the suprachiasmatic nucleus*, PLoS One. **6(1)**:522 e15869, (2011).

[53] Foley N.C., Tong T.Y., Foley D., Lesauter J., Welsh D.K., Silver R. *Characterization of orderly 536 spatiotemporal patterns of clock gene activation in mammalian suprachiasmatic nucleus*, Eur J Neurosci **33(10)**, 1851-65, (2011).

[54] Pauls S., Foley N.C., Foley D.K., LeSauter J., Hastings M.H., Maywood E.S. *Differential contributions of intra-cellular and inter-cellular mechanisms to the spatial and temporal architecture of the suprachiasmatic nucleus circadian circuitry in wild-type, cryptochrome-null and vasoactive intestinal peptide receptor 2-null mutant mice*, Eur J Neurosci, **542**, 2528-40 (2014).

[55] Abel J.H., Meeker K., Granados-Fuentes D., St John P.C., Wang T.J., Bales B.B. *Functional network inference of the suprachiasmatic nucleus*, Proc Natl Acad Sci U S A, **19** ,4512-7 (2016).