



A look at financial dependencies by means of econophysics and financial economics

M. Raddant^{1,3}  · T. Di Matteo^{2,3,4}

Received: 27 January 2023 / Accepted: 13 May 2023 / Published online: 1 June 2023
© The Author(s) 2023

Abstract

This is a review about financial dependencies which merges efforts in econophysics and financial economics during the last few years. We focus on the most relevant contributions to the analysis of asset markets' dependencies, especially correlational studies, which in our opinion are beneficial for researchers in both fields. In econophysics, these dependencies can be modeled to describe financial markets as evolving complex networks. In particular, we show that a useful way to describe dependencies is by means of information filtering networks that are able to retrieve relevant and meaningful information in complex financial datasets. In financial economics these dependencies can describe asset comovement and spill-overs. In particular, several models are presented that show how network and factor model approaches are related to modeling of multivariate volatility and asset returns, respectively. Finally, we sketch out how these studies can inspire future research and how they contribute to support researchers in both fields to find a better and a stronger common language.

Keywords Networks · Asset markets · Information filtering

JEL Classification C58 · C22 · G12

✉ M. Raddant
raddant@csh.ac.at

T. Di Matteo
tiziana.di_matteo@kcl.ac.uk

¹ Department for Knowledge and Communication Management, University for Continuing Education Krems, Dr.-Karl-Dorrek-Straße 30, 3500 Krems, Austria

² Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK

³ Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria

⁴ Centro Ricerche Enrico Fermi, Via Panisperna 89 A, 00184 Rome, Italy

1 Introduction

Dependencies in financial markets are complex and can be found at many levels. These markets are influenced by changing economic fundamentals and economic policies around the world. There are also numerous interactions within the asset markets themselves, and there is feedback from market participants who have varying expectations about future developments. Many of these processes do not result in measurable streams of data but only manifest themselves in latent variables, like the observable changes in asset prices and other financial data. In order to understand the functioning of financial markets we therefore have to analyze the structures that are present in market data, which in turn means to measure and explain their dependencies. This is of course not a new endeavor, and the literature in the field of financial economics has dealt with this for several decades and has developed models that explain the differences in asset returns. Network science has however brought a new approach to the field, as it has enabled us to describe the dependencies of large sets of assets jointly—and not only factors that explain differences.¹

Research on the quantitative aspects of financial markets has always been interdisciplinary, at the intersection of economics, mathematics and statistics. With the advent of networks science however, even more fields became interested in financial markets, especially researchers from physics, computer science and engineering. This has led to many new and important insights into financial markets, but has also made it more difficult to understand the diversity of approaches and to evaluate what the most powerful methods are to study financial dependencies. This review therefore specifically aims at showing different approaches to measuring financial dependencies across disciplines and connecting them to the foundations in financial economics and statistics. The presentation will therefore sometimes be rather wide in scope than deep. This is intentional because we believe that interdisciplinary research is typically only impact-full if it reaches researchers in different disciplines, and when it shows that its authors understand and address research questions across disciplinary borders. We will focus on the analysis of asset returns, in particular the dependencies between stocks, measured by their daily price changes. Most of the presented material will however also apply to any other financial time series in any other frequency.² We are guided by four questions: how are the returns modeled? How are dependencies measured? How are filtering or dimensional reduction achieved? How are the measurements mapped into a network representation?

Hence, we first touch upon cornerstones in financial economics and new approaches from econophysics in Sect. 2. In Sect. 3, we summarize statistical regularities of asset returns. In Sect. 4, we will explain the basics of calculating correlations and its challenges together with some of its pitfalls. Section 5 shows the spectrum of correlation matrices and how it is related to finding stocks with similar behaviors. We then proceed to present recent approaches from econophysics and financial econometrics. In Sect. 6, we discuss how networks can be derived from

¹ See also Caldarelli (2012) and Sect. 2.

² See also Tsay (2010) and Dacorogna et al. (2001) for general introductions to the analysis of financial market data.

correlation measures and we introduce two filtering methods that apply structural constraints on the network. Further we show how clusters and emerging properties can be identified in such networks. We then turn to present GARCH models, pairwise estimation methods, and the variance decomposition approach in Sects. 7, 8 and 9. In the last section, we present some conclusions.

2 The broader picture of financial market research

2.1 Cornerstones in financial economics

Before discussing recent studies on financial dependencies and networks, it is important to understand the origin of the field, together with its main motives and the scientific language used.

In financial economics, there are two important viewpoints when it comes to asset markets. One is the analysis of market efficiency; the other one is the analysis of the returns of an asset and how it relates to the development of the market and that of other factors. We will focus on the latter. A well-known model in this class is the so-called single index model by Sharpe (1963). Here we assume that a vector of returns r_i of an asset i can be described by

$$r_i = \alpha_i + \beta_i r_M + \varepsilon_i, \quad (1)$$

where r_M is the market return and β_i measures how sensitive the stock's return is with respect to the market. α_i is a constant, ε_i is the residual error. Since we assume that the errors are uncorrelated, i.e., $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ for all $i \neq j$, any correlation between stocks can only come from the market factor r_M .

A slight variation of this idea is the capital market pricing model (CAPM, Sharpe 1964; Lintner 1965), which includes a risk-free reference interest rate r_F , and where the β then describes the risk of an investment relative to the market portfolio. Additionally, there is also a large literature that discusses to which extent the β s are changing over time (see for example Bollerslev et al. 1988; Bodurtha and Mark 1991; Raddant and Wagner 2017).

Variations with more factors have been developed over the years, including the well-known model by Fama and French (1992) which includes factors that account for size and value effects.

$$r_i - r_F = \alpha_i + \beta_{1,i}(r_M - r_F) + \beta_{2,i}\text{SMB} + \beta_{3,i}\text{HML} + \varepsilon_i. \quad (2)$$

Here, β_1 measures the sensitivity of a stock to market risk. β_2 and β_3 measure the sensitivity to the factors SMB (Small Minus Big) and HML (High Minus Low). These factors represent the average return of stocks with certain characteristics in terms of firm size, value, and growth. Hence, implicitly this model assumes that stocks can be clustered according to these three characteristics.

While much of the differences in asset returns can be attributed to the actions and performance of the firms themselves, there is also ample evidence for the influence of investors. They sometimes evaluate stocks with certain characteristics in a too

homogeneous way, disregard differences and thus influence stock prices (Barberis et al. 2005; Green and Hwang 2009). Investors are also not always good in spotting firms that are economically linked, thus missing opportunities to predict returns (Cohen and Frazzini 2008).

It is important to note that the analysis of the comovement on the global level uses more aggregated data (see, e.g., Rigobon 2003). Most approaches have focused on the analysis of stock market indices or other smaller samples including sectoral indices. A wide range of methods has been applied; among these are unit root and cointegration tests, vector autoregression models, correlation-based tests (Forbes and Rigobon 2002; Fry et al. 2010), causality tests (Billio et al. 2012), multivariate GARCH models (Engle 2002), and models of variance decomposition (Diebold and Yilmaz 2009). We will have a closer look at the latter in the following sections.

Another aspect in the literature about financial market dependencies is the question about its determinants (Bracker et al. 1999). Studies often find that structural similarity of the countries' economies explains only partially the level of comovement of their financial markets. This resulted in a debate about the influence of global sectoral factors (Dutt and Mihov 2013; Bekaert et al. 2009, 2011). Previous results hint at an increase in the importance of these factors. Forbes and Chinn (2004) find that cross-country factors and global sectoral factors both are important determinants of stock returns. They also note that changes in global linkages over time might make it difficult to disentangle different influences on asset market comovement.

2.2 Statistical physics, econophysics, and complex networks

The contribution of statistical physics and the development of econophysics approaches have much to do with the data revolution in social science and economics (see also Pentland 2014; Kutner et al. 2019; Caldarelli et al. 2018; Steinbacher et al. 2021). Large-scale investigations on the scaling-properties of asset returns (Plerou et al. 1999, see also Sect. 3) opened the field for many similar investigations of financial data (Mantegna and Stanley 1999; Bouchaud and Potters 2009b). These approaches have rather quickly led to the realization that important challenges in the field are the dependencies of financial data and the complexity of the financial system in general (Laloux et al. 1999; Plerou et al. 2002).

Over the course of a decade, the analysis of the cross-correlations of asset returns has developed into an analysis of asset markets seen as complex networks, studied with techniques from network theory (Mantegna 1999; Song et al. 2011a). This has been accompanied by a general influx of network science into the analysis of financial markets in general, fueled of course by the financial crisis that peaked in 2008 (Bardoscia et al. 2021; Summer 2013; Acemoglu et al. 2015; Battiston et al. 2016; Elliott et al. 2014). Further extensions in the methodology are the analyses of bi-partite networks, systems with couplings, and multiple layers of networks (Tumminello et al. 2011; Gualdi et al. 2016; Buldyrev et al. 2010).

Describing the details of this evolution is outside the scope of this paper. The developments of new methodology however show that collaboration across disciplines is happening, especially when important problems arise to which established models have too little to offer. The connecting element between disciplines is the data. Yet, in order to respond to economic and social challenges, the observed data have to be explained by models in a way that researchers from different disciplines appreciate.

3 Preliminaries about asset returns

Dependencies between financial assets are mostly analyzed based on their log daily returns $r_{i,t}$, which are calculated from the stock prices $p_{i,t}$ of asset i at the days t and $t - 1$,

$$r_{i,t} = \ln p_{i,t} - \ln p_{i,t-1}, \quad (3)$$

where $i = 1, \dots, N$ and $t = 1, \dots, T$. N denotes the number of stocks, T denotes the number of observations, in this case these correspond to trading days.

Daily closing prices are the most common data used in the analysis of asset returns. For some applications, e.g., when different markets are involved and trading hours are not perfectly aligned, lower frequencies, for example weekly returns, can be used. Note that the use of logarithmic returns allows to calculate these by summation of the higher frequency returns. On the other end, there is a trend toward the use of higher frequency data in the trading industry that utilizes advances in scalable computing and often employs machine learning.

For most of the following analyses we use daily data covering the time period from April 2006 until March 2021 ($T = 3,775$ trading days) for $N = 404$ stocks that are constituents of the S &P 500 index. Three exemplary time series of daily prices and returns are plotted in the left panels of Fig. 1. For the purpose of comparison, we plot normalized prices \tilde{p} , such that

$$\tilde{p}_{i,t} = \frac{p_{i,t}}{p_{i,1}}. \quad (4)$$

While stock prices show generally some upward trend in the long run, we can see that the returns shown in the bottom left panel fluctuate around zero with some visible spikes and clusters.

The volatility of returns can be defined by their sample variance $\hat{\sigma}^2$ as

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{r})^2. \quad (5)$$

One can also define a variance in terms of the expectation of the time series. A common notation is to state that

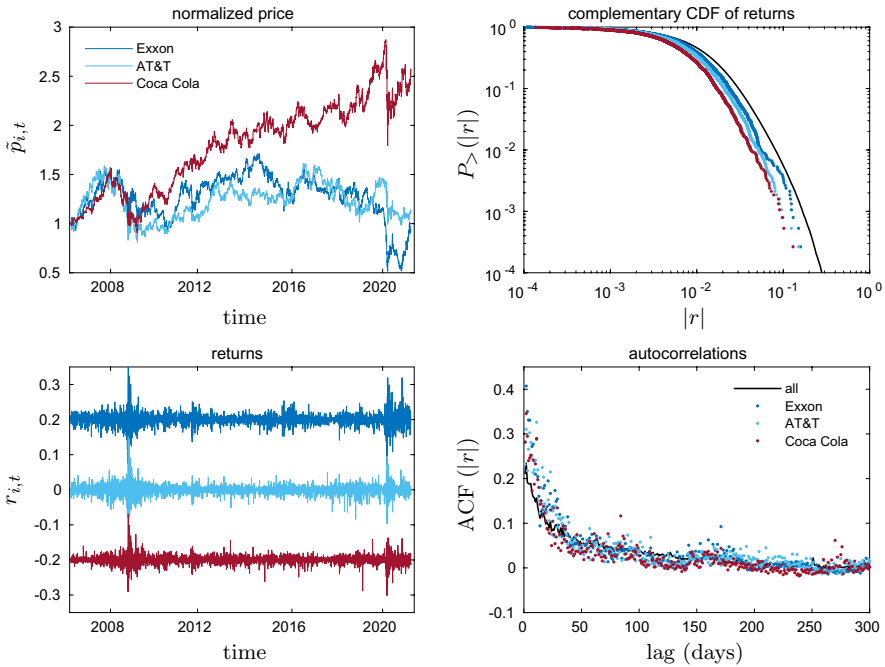


Fig. 1 Daily stock returns and their properties. The top left panel shows the normalized stock price \bar{p} for three stocks over time. The bottom left panel shows the corresponding daily returns of these three stocks (two of them plotted with an offset of ± 0.2). The top right panel shows complementary CDFs (CCDFs) of absolute returns on log-log scales for the three stocks, as well as the CCDF for all stocks (black line, with the CCDF = $1 - F(x)$, where $F(x) = Pr[X \leq x]$). The bottom right panel shows the corresponding auto-correlations dependent on the lag in days

$$\sigma^2 = \langle r^2 \rangle - \langle r \rangle^2, \tag{6}$$

where $\langle \rangle$ denote the expectation. There are also instances where $|r|$ or r^2 are used as measures for volatility. This is of course motivated by the fact that $\langle r \rangle$ is typically close to zero and that the intercept of r is in many models estimated separately from the volatility component (more on volatility models in Sect. 7).

Hence, one aspect of the returns is that their distributions are non-Gaussian, heavily skewed, and fat-tailed. In fact, the tail of these distributions is well represented by a power-law, in our case with an exponent of 2.94, as can be seen in the top right panel of Fig. 1 (see also Plerou et al. 1999).

Another aspect of the returns is that the volatility is not constant over time. Times with high volatility are often followed by more high volatility. This phenomenon is also called *long memory* (see also Cont 2001). We can measure this aspect by calculating the autocorrelations, which are the correlations of the time series with lagged versions of itself. Their entirety is also called the autocorrelation function (ACF), which is shown in the bottom right panel of Fig. 1. The autocorrelations of the absolute returns are typically significantly positive for

more than 100 days (Ding et al. 1993). Hence, while we cannot forecast tomorrow's stock prices, some information about tomorrow's volatility is included in today's prices.

These findings are important for two reasons. First, because they influence how we can analyze the time series of returns and their dependencies. The changing volatility can make it difficult to determine whether two time series are in fact significantly related. Properties of the distribution of stock returns imply that many parametric tests do not apply for this data. The second aspect of these findings is that they have of course been discussed in streams of literature in different fields and that it is therefore important to understand what the state-of-the-art models are that deal with them – so that future research on financial networks can benefit from and contribute to them.

4 Correlations of asset returns

The correlations between all pairs of returns can be represented by a $N \times N$ correlation matrix C . Each element C_{ij} of this matrix describes the type of relationship (dependence) between a pair of stocks i and j .

Pearson's correlation coefficient is the most common measure to describe correlations that are linear. For each pair of stocks i and j it can be calculated as

$$C_{ij} = \frac{\langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle}{\sigma_i \sigma_j}, \quad (7)$$

where $\sigma_i \sigma_j$ is the product of the standard deviations of the returns of stocks i and j (see also Eq. 6).

Correlation coefficients will always be in the range $[-1, 1]$ since we divide by the product of the standard deviations of both variables. An alternative way of measuring dependencies is the covariance, which is given by the numerator of the above expression. This can be useful if one wants to preserve the differences in the scale of the returns for further calculations.

Figure 2 shows distributions of correlation coefficients for the S &P sample of stocks. While the distribution of the correlation coefficients of stocks from different sectors (between-sector) shows similarity to a Gaussian distribution, the distribution of within-sector correlations has a higher mean and shows skewness.

In many applications, we are interested in a dynamical analysis of these correlations. Pearson's correlation coefficients are therefore often calculated by applying a running window approach, as illustrated in Fig. 3. Equation 7 is in this case applied only to the returns inside a time window with length Δt . Here the first three panels show pairs of stocks from different sectors, while the last panel shows an example for a pair of stocks which both belong to *Non-cyclical consumer goods*.

Using running windows can however lead to spurious results when the window size Δt is too small. Correlations are then likely to be exaggerated. Also, the comparison of these correlations is difficult since the normalization by the changing variance in Eq. 7 will drive the results.

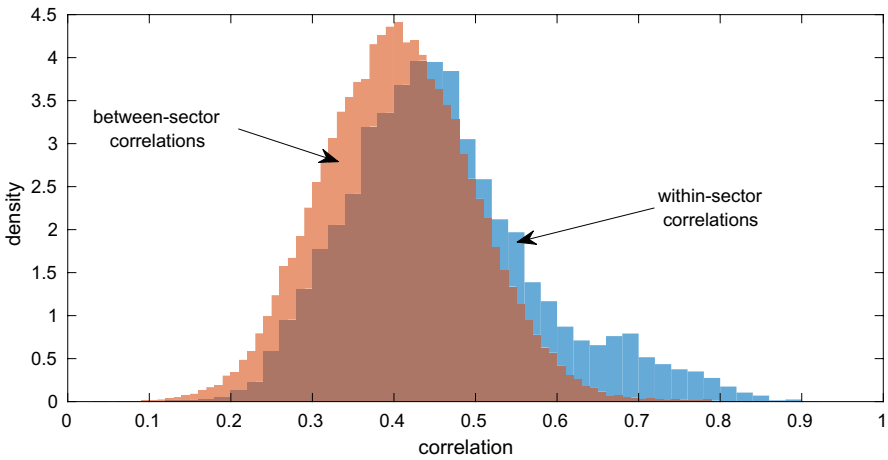


Fig. 2 Histogram of correlation coefficients for the S & P stocks. The two colored areas show the distributions of correlations between stocks from different sectors (between-sector, orange) and within the sectors (blue)

The assessment of the significance of correlations therefore deserves a closer look. It is well known that both the error of the correlation coefficient itself as well as intervals for significance levels that determine whether a correlation coefficient is significantly different from zero can easily be calculated making use of the corresponding t-distribution. The latter is achieved by a parametric test that essentially answers the question what range of correlation coefficients can be expected from time series with a specific length when they are uncorrelated and normally distributed (i.e., when they consist of random variables that are independent and identically distributed). In order to evaluate in how far the violation of this assumption affects the correlations, we can compare correlation coefficients of permuted time series pairs of returns (of which the expected correlation is zero) with that of uncorrelated noise (see also Aste et al. 2010). Figure 4 shows the distributions of correlation coefficients, both for permuted real data as well as for independent normally distributed noise (we draw 500,000 samples for $\Delta t = 20$ and 60 days with random starting points, and we calculate all possible correlation coefficients for $\Delta t = 3775$, which corresponds to the entire sample).

First of all, it is important to realize that the significance of correlation coefficients becomes much more difficult to assess once the length of the time window becomes smaller. We can clearly see this by comparing the distribution of obtained correlation coefficients for the entire sample period with that associated with the time window length used before, $\Delta t = 60$ days (left panel). With regard to our results on the dynamics of correlations presented in Fig. 3, this means that the correlations between most stocks are only significantly positive for some time windows.

A second aspect is the slight inaccuracy when parametric tests are applied to this correlation measure of asset returns. In order to visualize the effect it is necessary to focus on the tails of the distributions (right panel in Fig. 4). When we compare these tails with the corresponding distributions of independent normally distributed noise,

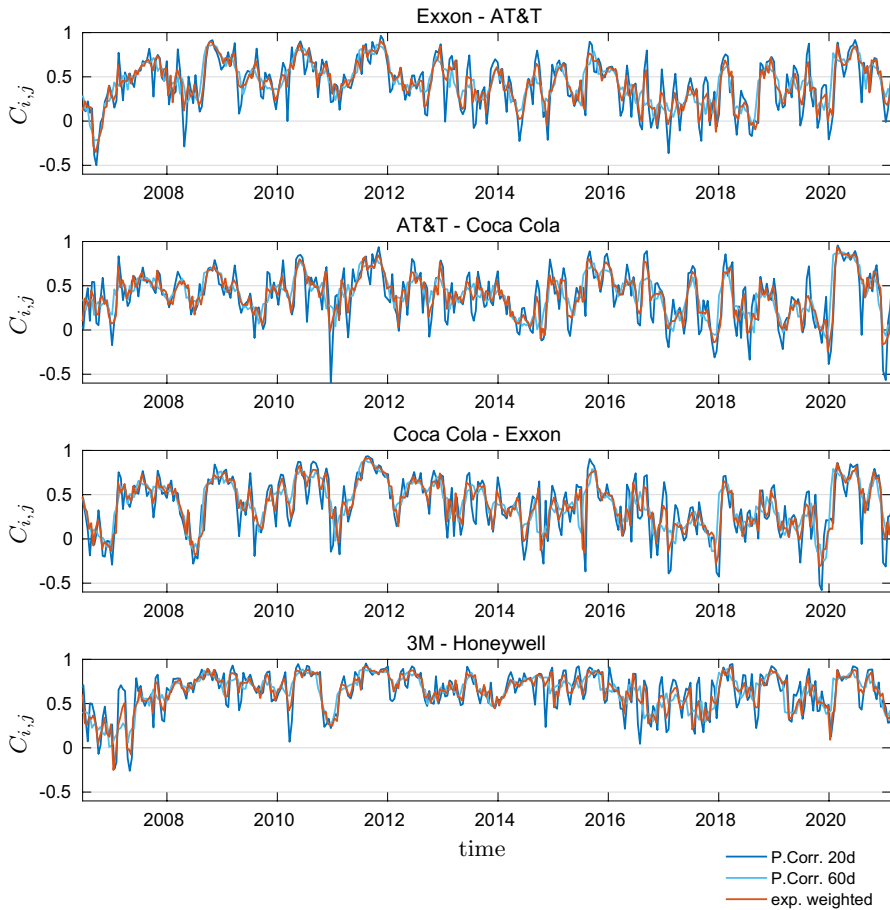


Fig. 3 Stock correlations with different window lengths. These panels show four examples for the development of correlations over time. In each of them the dark blue line shows Pearson's correlations calculated in a window with $\Delta t = 20$ days. The light blue line shows these correlations calculated for $\Delta t = 60$ days. The red line shows exponentially weighted correlations with a time window of $\Delta t = 60$ days (and $\theta = 20$). The top three panels show examples where the stocks come from different sectors. The bottom panel shows an example of within-sector correlations

we observe slightly too many extremes in the correlation coefficients of the permuted returns, especially for small Δt . Hence, relying on a parametric test for the correlation coefficients can in this case lead to an overstatement of their significance (see also Kowalski 1972).

It is therefore of utmost importance to select time windows such that true dependencies can be reliably identified as significant. In some applications, the choice of the window length can be a trade-off, especially when there is the necessity to associate networks with specific economic events. It turns out that an improvement can be achieved by introducing a weighting of the observations within a time window. By putting high weights on the latest returns and successively lower weights on past

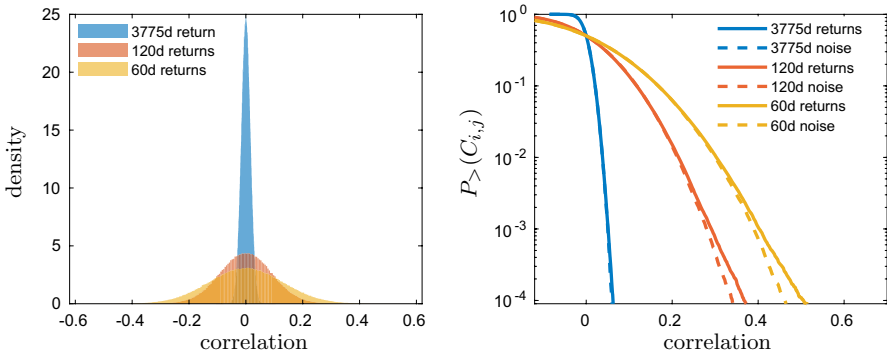


Fig. 4 Correlation coefficients from permuted returns and from noise. The left panel shows probability distribution functions that were obtained by calculating the correlation coefficients of randomly drawn pairs of permuted time series with different length from the S &P returns data ($\Delta t = 60, 120, 3775$ days). The right panel shows the comparison of the tails of these distributions (as a CCDF) with corresponding time series composed of independent and identically distributed noise

observations, the correlations can be smoothed while retaining a relatively high time resolution (Pozzi et al. 2012). In this case a weighting applies to the sample mean, the variance and the covariances in Eq. 7. The weights $\sum_{t=1}^{\Delta t} w_t = 1$ follow

$$w_t = w_0 \exp\left(\frac{t - \Delta t}{\theta}\right), \quad t \in \{1, 2, \dots, \Delta t\}, \quad (8)$$

which describes an exponential weighting scheme.³

It should be noted that several other techniques exist to work around the unreliability of Pearson’s correlations or that try to filter the returns for certain effects.

One technique for such a filtering is the use of partial correlations, which investigates how the correlation between two stocks i and j depends on the correlation of each of the stocks with a third variable, typically a stock index m as the mediating variable. The partial correlation $\rho_{i,j|m}$ can be viewed as the residual correlation between stocks i and j , after subtraction of the contribution of the correlation between each of the stocks with the index.

Some studies have extended this approach to the analysis of the ratio of partial and raw correlations and the metacorrelations between different markets (Shapira et al. 2009; Kenett et al. 2011, 2012).

Another alternative is the use of correlation measures that are nonparametric, like rank correlation, or the tail dependence. The latter is a nonparametric estimator of the tail copula that provides a measure of dependence focused on extreme events (Musmeci et al. 2017). Also a transformation of the returns can improve

³ θ is the weights’ characteristic time, with $\theta > 0$. w_0 is a constant, in the case of Pearson’s correlations equal to $(1 - \exp(-1/\theta))/(1 - \exp(-\Delta t/\theta))$.

measurements if it brings their distribution closer to that assumed by a particular model or test.

5 Properties of correlation matrices

In the previous sections, we have introduced the correlation matrix of returns and have given examples for the empirical results that we typically obtain by calculating correlations coefficients from stock returns. In this section, we want to connect these findings to the most important theoretical concepts about the decomposition of correlation matrices, since these are useful foundations to understand the structure contained in stock markets.

In Eq. 7, we have already shown how to calculate correlation coefficients. An equivalent statement is the definition of the sample correlation $R_{i,j}$.

$$R_{i,j} = \frac{\sum_{t=1}^T (r_{i,t} - \bar{r}_i)(r_{j,t} - \bar{r}_j)}{\sqrt{\sum_{t=1}^T (r_{i,t} - \bar{r}_i)^2} \sqrt{\sum_{t=1}^T (r_{j,t} - \bar{r}_j)^2}} . \tag{9}$$

Note that we use sum notation here instead of the brackets as in Eq. 7; the mean of r is thus denoted by \bar{r} . It is worth to be familiar with both notations since they are both frequently used in the literature.

We also note that some contributions to the literature, especially about the eigenvalue spectrum, assume standardized returns \tilde{r} . In this case, the sample correlation \tilde{R} can be written as follows and $R = \tilde{R}$

$$\tilde{R}_{i,j} = \frac{1}{T} \sum_{t=1}^T \tilde{r}_{i,t} \tilde{r}_{j,t} . \tag{10}$$

We can use the spectral analysis of the correlation matrix to infer a few properties with respect to the level of noise versus structure that it contains. A prominent method concerned with the spectrum of correlation matrices of uncorrelated data is random matrix theory (RMT). Assuming time series of white noise with zero mean and unit variance, Marčenko and Pastur (1967) state that the distribution of eigenvalues λ of its correlation matrix follows

$$p(\lambda) = \frac{Q}{2\pi\sigma_r^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} , \tag{11}$$

where $Q = T/N$ adjusts for the ratio of the length of the time series versus the sample size. σ_r is a further adjustment parameter which is 1 in the case of standardized data. It can also be used to compensate for single eigenvalues outside the predicted spectrum.

The boundaries of the spectrum are given by

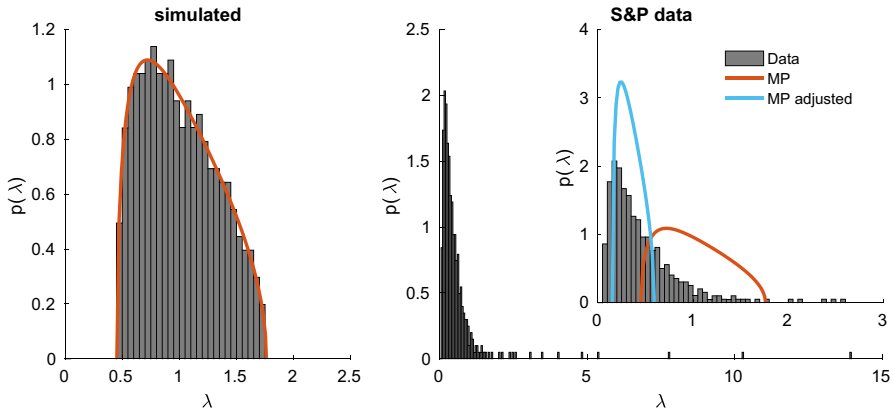


Fig. 5 Histograms of eigenvalues of correlations matrices and MP spectrum. The left panel shows the histogram of the eigenvalues for random time series with the same Q as the real data. The MP spectrum fits the data exactly. The right panel shows the distribution of eigenvalues of the market data. The largest eigenvalue, $\lambda_1 = 170$, is not shown. The insert shows the eigenvalues in the range predicted by MP. Large parts of the market’s spectrum deviate from MP. When one adjusts the MP spectrum for the leading eigenvalue by setting $\sigma_r = 0.58$ some qualitative alignment can be achieved

$$\lambda_{\pm} = \sigma_r^2 \left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}} \right). \tag{12}$$

If $N \rightarrow \infty$ and $T \gg N$ the eigenvalues should all lie in the interval $[\lambda_-, \lambda_+]$ (Laloux et al. 1999; Plerou et al. 2002; Bouchaud and Potters 2009a).

An example is shown in Fig. 5. Artificial time series, i.e., simulated returns drawn from a normal distribution that are uncorrelated, produce a histogram of eigenvalues that agrees with the prediction from Eq. 5 (left panel).

We should clarify that this analysis does not suggest that the MP prediction describes the spectrum of eigenvalues from empirical asset return data perfectly (see also Livan et al. 2011). It merely shows that the number of eigenvalues that are far outside the predicted spectrum is rather small.

It is very common to find one large eigenvalue, which is associated with the market mode. Its corresponding eigenvectors can for example be used to calculate the weights of stocks in a market index or conversely for the construction of certain optimized portfolios. The other eigenvalues outside the MP spectrum generally describe groups of stock that show a systematic behavior that is different from the market. It should therefore be possible to generate good approximations of the correlation matrix by models that decompose it using a number of factors $k = 1 \dots K$ much smaller than N . We will now introduce two approaches for such a decomposition.

The correlation matrix defined above is by construction an $N \times N$ symmetric matrix that can be diagonalized. This is the basis of the well-known principal component analysis (PCA). In fact, for many cases it makes sense to go one step further and to use this decomposition to model the returns themselves. We therefore decompose changes in $r_{i,t}$ into decorrelated contributions of decreasing variance. In terms of the eigenvalues λ_k and eigenvectors v_k , the decomposition reads

$$r_{i,t} = \sum_{k=1}^N \sqrt{\lambda_k} v_{k,i} \varepsilon_{k,t}, \quad (13)$$

where $v_{k,i}$ denotes element i in the k th eigenvector and $\varepsilon_{k,t}$ are uncorrelated random variables with unit variance.

We note that the formulation above is very common in any physics textbook, yet statistics textbooks in economics or business often use a different viewpoint, which we will also briefly sketch out. Let us assume that our sample of returns (our variables) are stored in a matrix X with dimensions $N \times T$. We can then calculate a new set of variables, the principal components F , which are linear combinations of the original variables (for examples for a dimensional reduction of the system). Hence, we search for $F = YX$. The column vectors in Y carry the weights for each new variable, and it can be shown that the solution to this problem amounts to solving for

$$F = VX, \quad (14)$$

where $Y = V$ contains the eigenvectors of the correlation matrix R (typically ordered by descending eigenvalues, see Jobson 1991). The eigenvectors with the highest eigenvalues account for a large amount of the variance in the data, while low eigenvalues stand for eigenvectors and components that contribute very little to the variance and can therefore be neglected to describe X .

The difference between these two approaches is what is explained on the left hand side (note that we can derive the factor loadings by scaling the eigenvectors with the variance $\sqrt{\lambda}$). Equation 13 states the decomposition of the returns as a stochastic model, Eq. 14 states an identity that describes the factors of said decomposition. For a more detailed example the reader might refer to Verma et al. (2019).

6 Correlations and networks

6.1 Networks, filtering, and dimensional reduction

Financial markets are systems with a large number of elements, in our case the number of stocks N . All these stocks can be differently affected by one another. We aim to infer the network of most relevant links by studying the mutual dependencies between stocks based on an analysis of correlation measures. In other words, we search for stocks that behave similarly and we want to link them with edges in the network. Conversely, we do not want to directly connect stocks that behave independently. In many cases, it is further necessary to determine which nodes can be clustered together. This information can either be used to group these stocks, and it can also be used to dimensionally reduce the system and to define a new set of nodes based on these clusters' aggregations.

Hence, there are typically two considerations to be dealt with when it comes to mapping stocks into networks. The first one is which links to include into a network given the full set of stocks. The second one is the dimensional reduction on the part

of the nodes (stocks) itself. The later becomes more important once the size of the network becomes large and can no longer be usefully visualized, or when subsequent analyses of the network are not feasible for the entire sample. The problem of a dimensional reduction can either come into play at the stage of the visualization of the network or it may come into play earlier, namely at the stage of estimating correlations within part of an econometric model.

Let us start with the simplest case, which is the mapping of the correlation matrix into an adjacency matrix by using a threshold rule. The adjacency matrix is a matrix where rows and columns are labeled according to the nodes (stocks) in the network and where the entries in the matrix describe whether certain nodes are connected (adjacent) to each other. In this case, we create an adjacency matrix A with dimensions $N \times N$, identical to our correlation matrix. One constructs this adjacency matrix by setting a threshold for the correlation coefficient and defining that the elements A_{ij} are 1 if the corresponding correlation coefficients are above a certain threshold and zero otherwise. The diagonal elements of A are typically zero, unless the concrete example requires to express a self-loop. The result will be an undirected unweighted network. Weighted networks can be obtained by using the absolute value of the correlation coefficients (or transformations of these) as weights of the edges. Examples are shown in Appendix A.⁴

While these networks can be informative, there are also some drawbacks. Their visualizations can sometimes change drastically depending on the threshold level that is chosen. They often contain a very high number of edges and it can be difficult to assess the significance of these connections. Resulting networks may also contain several components. In other words – they do not necessarily provide a meaningful and robust simplification of the system by means of network topology.⁵

The second consideration in creating networks is the dimensional reduction of the system itself. We will illustrate this by applying a simple clustering technique to our dataset. The output of such a clustering can be used to reduce the number of nodes in the system. Most clustering techniques will require to assess the similarity of stock returns by some distance measures. The correlation matrix contains information about similarity, yet it makes sense to transform it in such a way that it reflects a metric. Hence, the following transformation can be used to derive a distance matrix D (see also Mantegna 1999)

$$D_{ij} = \sqrt{2(1 - C_{ij})}. \quad (15)$$

Hence, the elements of the correlation matrix are transformed in such a way that D has a lower bound of 0 that can only be obtained by perfectly correlated stock returns.

⁴ The network visualization in this review were produced in Gephi (Bastian et al. 2009), mostly using the algorithm by Hu (2005).

⁵ It should be noted that on an only slightly generalized level the problem of finding the ‘relevant’ links in a network connects to an incredibly large amount of literature in mathematics, computer science and statistics, all in some way concerned with finding sparse versions of covariance matrices. A noteworthy example, outside the scope of this review, is, for example, specific convex optimization (Bayesian) methods also known as the Lasso (see, for example, Zou 2006; Bien and Tibshirani 2011).

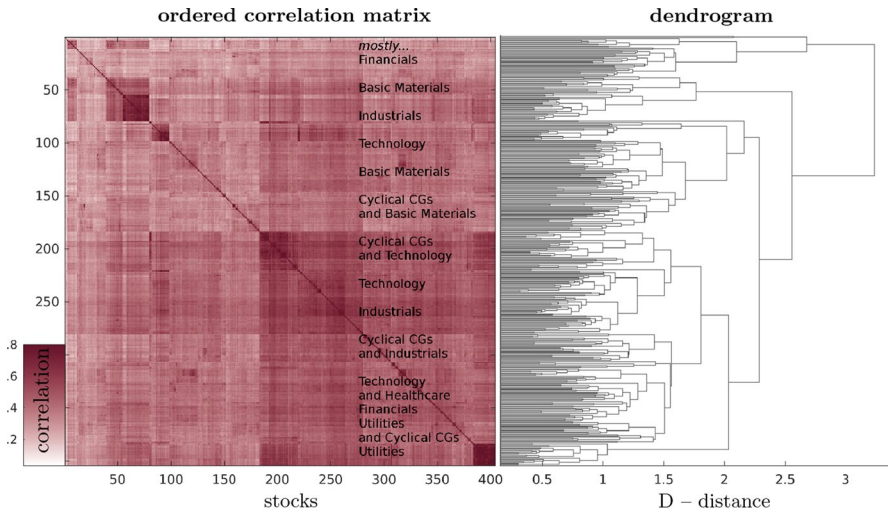


Fig. 6 Color-coded correlation matrix and dendrogram of the hierarchical clustering for the constituents of the S & P 500. The rows and columns in the left panel are ordered according to the dendrogram in the right panel. Sector classifications for the clusters are superimposed based on the most frequent label within a cluster

One can then use the matrix D in an algorithm that joins the stocks together based on their distances. Here we will focus on a hierarchical approach. Assume that at the start all stocks are in a group (or cluster) on their own. We will start by joining the pair of stocks with the lowest distance. Before searching for the next shortest distance between all groups, we have to make a choice about how to define the distance between the now joined first cluster and the remainder of the stocks. Common solutions are to use the average or the minimum. Hence, by repeatedly joining the groups that are closest together and by updating our distance matrix for the groups, we will ultimately connect all stocks in a hierarchical structure – at different levels of the distance measure.

This result can be demonstrated by the dendrogram shown in the right panel of Fig. 6. It illustrates which stocks would be sorted into the same group given a certain threshold for the distance. The left panel shows a visualization of the correlation matrix where the rows and columns are sorted according to the dendrogram. The regularities in the color-coding hint at uncovered structures in the correlation matrix.

While such a grouping can certainly be visualized as a network, we should note that a few more steps would be needed to obtain proper adjacency matrices. First, one would have to set a threshold for a cutoff point in the dendrogram. Second, one would have to set a rule that defines how connections within the identified clusters are dealt with, or if new nodes, based on the identified clusters, would be constructed.

The number of clustering techniques is very large, as are its applications to networks, which leads to the problem of community detection (see also Jobson 1991; Newman 2010). Studies on the use of correlations in this context include MacMahon

and Garlaschelli (2015) and Masuda et al. (2018). An excellent comparison of community detection algorithms can be found in Lancichinetti and Fortunato (2009).

The literature in financial economics has a slightly different approach on the filtering of dependencies. Since the sample sizes discussed in this strand of the literature are typically slightly smaller than in econophysics, many approaches aim at deriving full matrices of weighted dependency measures. The problems that these approaches face with respect to system size are however tightly related to the problems that we have just sketched out. Dimensional reductions are in these models often made in the process of estimating the dependencies in order to circumvent numerical limitations. We will come back to some of these issues in Sects. 7, 8 and 9.

In the following, we will now present approaches that provide an efficient filtering of the system of dependencies and we show how these can be extended to an endogenous clustering.⁶

6.2 Minimum spanning tree

The minimum spanning tree (MST) is a tool to analyze and filter the information contained in the correlation structure of a set of financial assets. We look at the correlation matrix as the adjacency matrix of a network and generate an MST on this network in order to retain the most significant links. The complete network is represented by links which weights are determined by the corresponding correlation coefficients. This complete network has $\frac{N(N-1)}{2}$ links, representing the full information of the system.

The MST approach by itself is a rather old mathematical problem that has been discussed since the 1930 s (and solved, see Sedgewick and Wayne 2011). A minimum spanning tree (MST) consists of the subset of edges (links) of a graph that connects all the vertices (nodes) using the minimal total edge weight. Hence, it can be interpreted as the “least expensive” connected graph. It contains no loops and the number of edges is $N - 1$, given that N is the number of vertices (nodes).⁷

The details of applying MSTs to asset markets were for the first time presented by Mantegna (1999). In order to describe the dependencies of stock returns as a MST, we first have to define distances between our stocks, and we do this again by applying Eq. 15. Several algorithms for the construction of MSTs have been proposed, here we are using hierarchical clustering and single linkage (Gower and Ross 1969; Kruskal 1956; Prim 1957).

To construct the MST, let us assume that at the start our network consists of all the stocks (nodes) in our dataset, yet no links exist between them. We start by sorting all the pairs of stocks by their distances in ascending order. We start by connecting the stocks with the shortest distance. We then proceed with all the remaining pairs in the order of their distances while omitting links that would create a loop.

⁶ We note that further important contributions to the research on correlation-based networks include Song et al. (2011a), Onnela et al. (2003b) and Tumminello et al. (2010).

⁷ MST being a classical math problem, you will find the terminology of graphs, vertices and edges in the related literature instead of networks, nodes and links. The former terms are common in settings when abstract networks are discussed, yet by today they are often used interchangeably.

It can be shown that this algorithm does in fact produce the MST with the lowest overall edge-weight (sum of distances) possible. The proof goes along the lines of showing that at each step of the algorithm we are using the lowest weight edge to connect the single parts of the so-called forest of clusters of nodes into a single tree, hence, there can be no improvement to the total edge weight by swapping edges.

This algorithm will in the early stage typically produce several components of connected stocks that will gradually grow together to form a connected network. Often one will find several star-like regions that are made up of stocks that belong to similar sectors of the economy. A visualization is provided as part of Fig. 7. The MST algorithm provides a hierarchical classification of the stocks (see also Musmeci et al. 2015b). For a further analysis of asset trees see also Onnela et al. (2003a).

6.3 Planar maximally filtered graphs

6.3.1 Construction principles

The MST does by its very nature lead to a relatively drastic reduction in information by describing the dependencies by only those links that are absolutely necessary to produce a connected network. It is therefore natural to ask if further algorithms can be formulated, which allow for more links and thus information to be represented in more detail.

A side effect of the construction of the MST is that it produces networks that can without much effort be displayed in a 2-dimensional plain representation. When we want to increase the amount of information in the network by adding more links we therefore have to define a consistent rule that stops us from adding links that make the network become too complex. The most cited method for this is the planar maximally filtered graph (PMFG) described in Aste et al. (2005) and Tumminello et al. (2005). The idea here is to include all links that produce a graph that is planar. This corresponds to creating a network embedded on a sphere with genus equal to 0. In principal, this method can be applied to surfaces of different genera, yet in practice the most useful one is that of the planar case (Aste et al. 2012).

To construct the PMFG, we again define distances according to Eq. 15 and employ an algorithm similar to that used for the construction of the MST. Hence, we add links in ascending order based on the distance measure. The difference is that now we do not discard links between already connected stocks right away. Before adding links, we confirm that their addition does not produce a non-planar graph.

Checking this feature is unfortunately not entirely trivial. The problem of proving that a graph is planar goes back to work done by Leonard Euler. It turns out that it is relatively easy to show which conditions are necessary for planarity and non-planarity, but that sufficient conditions are relatively tricky. The problem was however solved by Kuratowski (1930) who showed that a graph is non-planar if and only if it contains a sub-graph homeomorphic to two particular graphs that are mostly described as $K_{3,3}$ and K_5 .⁸

⁸ Two graphs are homeomorphic if one is a subdivision of another, or they are both subdivisions of some third graph. $K_{3,3}$ is a bipartite structure with 3 nodes in each set, K_5 in a completely connected set of 5 nodes.

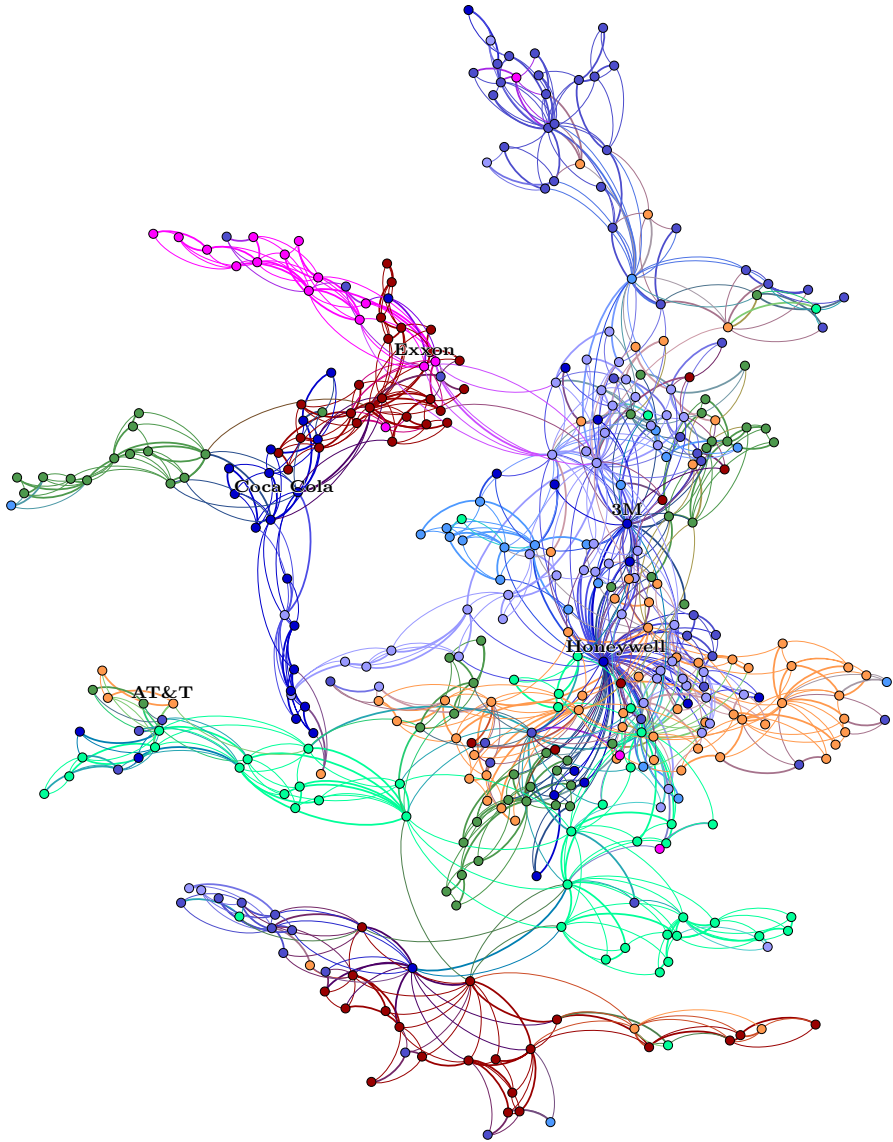


Fig. 7 PMFG for the constituents of the S & P 500. Connections of the corresponding MST are shown by bold edges. Color-coding by sector classification: pink—Energy, bright blue—Basic Materials, bright purple—Industrials, dark purple—Cyclical Consumer Goods, dark blue—Non-cyclical Consumer Goods, bright green—Financials, dark green—Healthcare, orange—Technology, brown—Utilities. This visualization highlights the clustering features of the PMFG and not its planarity—thus some edges cross

We will not go further into the details of this proof; the problem is however rather apparent when one sketches a network with 4 nodes and connects them all with a pen on a piece of paper. Adding another node and connecting it to all other

nodes is impossible. Hence, cliques with more than 4 nodes are not possible in a PMFG, and the number of links in a PMFG is always equal to $3(N - 2)$.

A visualization of a PMFG is shown in Fig. 7. We have labeled the positions of stocks that were mentioned before. *Honeywell* appears as a root in this network, as it comoves closely with many other stocks. Finding one stock as a dominant hub in the PMFG is a common finding in PMFGs of asset markets. The structure of the periphery of the network partly coincides with the sector classifications, while the center of the network is less segmented. For more detailed applications of the PMFG the reader may refer to Tumminello et al. (2007) and Kumar et al. (2020). A computationally more efficient algorithm suitable for large datasets is developed in Massara et al. (2017).

6.3.2 Clustering

When we look at the MST- and PMFG-filtered networks in Fig. 7, we note that some clustering emerges naturally. In the following, we therefore present a clustering that is based on the PMFG, the so-called DBHT method (short for Directed Bubble Hierarchical Tree). This particular algorithm produces a clustering that is deterministic and does not use any thresholding values and/or parameters, and it does not fix the number of clusters a priori.

In short, the DBHT method (Song et al. 2012) can achieve a clustering of the PMFG, because the network can be hierarchically divided by separating 3-cliques of nodes. This results in a set of planar graphs, which are referred to as bubbles. Nodes typically belong to more than one bubble, hence in a first step nodes are assigned to a unique bubble. In a second step the algorithm will then merge some bubbles to larger structures based on different distance measures. For further details on the method, see also Song et al. (2011b) and Musmeci et al. (2015a).

The results are shown as a color-coded graph in Fig. 8. We obtain 12 clusters with varying size. The borders of the groups follow the hierarchical structure of the PMFG.

The grouping is in some relationship with the sector classification, as the bar plot in the bottom of the figure shows. Many clusters are dominated by stocks from only one or two sectors. An exception is the largest cluster (1) which is comprised of a large number of stocks from different sectors which do not show very strong sector-specific behaviors.

7 GARCH models

7.1 Overview

The stylized facts of financial data presented in Sect. 3 have of course not only inspired research from statistical physicists, but are in fact at the root of the development of financial econometrics as we know it today. Modeling the volatility of an asset serves as the foundation for applications not only in asset pricing but also in risk management, derivatives pricing and forecasting.

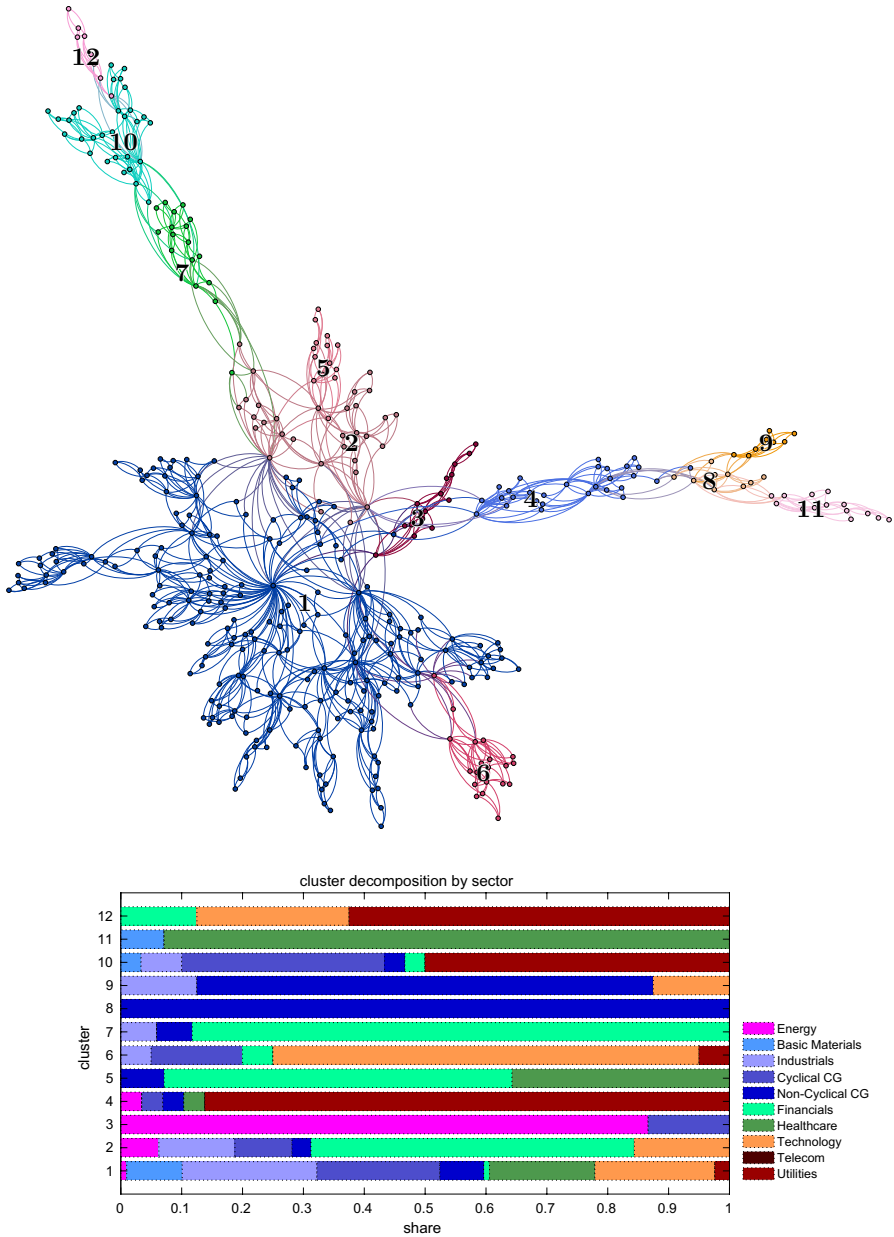


Fig. 8 Clusters in the PMFG. The network on top shows the color-coded clusters in the PMFG that are identified by the DBHT algorithm. The bar plot in the bottom shows the decomposition of each of these 12 clusters into fractions, i.e., the share of stocks from a particular sector

The most popular class of models in this field are the GARCH (generalized autoregressive conditional heteroskedasticity) models. These models are first of all time-series models. They provide a framework to model and estimate the

heteroscedasticity in volatility and at the same time they allow a rudimentary forecast of volatility. They have however originally not been invented with the analysis of correlations in mind, at least not of large samples. Hence, while an estimated dynamic conditional covariance matrix is the natural by-product of any multivariate GARCH model, only recent improvements on the estimation methodology have enabled us to utilize these kind of models in network settings. Independent of this fact, GARCH models, even in simpler form, can be helpful for constructing financial networks, since they allow us to disentangle joint volatility shocks from idiosyncratic ones.

We will first have a look at the univariate GARCH model and some examples of estimated stock volatility. We will then briefly show the basics of a multivariate GARCH and its complications before finally sketching out a network application of such a multivariate GARCH model.

7.2 Univariate GARCH

Let us first have a look at the univariate form of GARCH (Bollerslev 1986). We assume that the returns follow a random process with $\varepsilon_t = v_t \sqrt{h_t}$, where v_t is white noise and

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}. \quad (16)$$

Hence, we assume that the volatility h_t depends on the lagged values of the returns and its own lagged values (α_i and β_i are the estimated parameters for lag i , α_0 is a constant).

It is common to refer to a specific lag specification as a GARCH(p,q)-process. One typically finds that already one lag for q and p are sufficient to describe the volatility by the time series h_t (see also Hansen and Lunde 1998). The parameters for this recursion formula can easily be estimated by maximum likelihood. One of the applications of this model is obviously to relate the current (expected) volatility to past realizations of the returns. Another one is that h_t can be used to normalize (de-garch) the returns and to derive time series with constant unit volatility. Some examples for volatility estimates are shown in Fig. 9. Note that most of the GARCH literature differentiates between the actual returns r_t and the process that is modeled, ε_t , like in the equation above. In the following we will however label any return again as r_t .

7.3 Multivariate GARCH

The GARCH approach has of course been extended to a large number of multivariate versions, the most popular ones are the BEKK (Engle and Kroner 1995) and DCC (Engle 2002) model. The volatility is in this case described by a matrix of conditional covariances H_t (and its mean \bar{H}). The recursion of the K -component BEKK(1,1,K) model can for example be stated as follows:

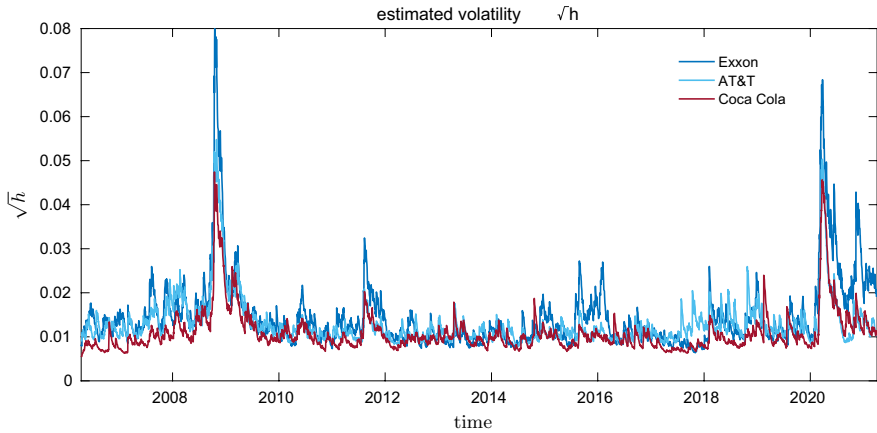


Fig. 9 Estimated volatility from a GARCH(1,1) model for three stocks. These three time series of volatility show typical qualitative features, including pronounced peaks and decay. Nevertheless there are significant differences between the stocks that are relevant for the evaluation of stock comovement

$$H_t = C'C + \sum_{k=1}^K A'_k (r_{t-1} r'_{t-1}) A_k + \sum_{k=1}^K B'_k H_{t-1} B_k. \tag{17}$$

C, A_k and B_k are $N \times N$ matrices (C is upper triangular, we denote the transpose of matrices and vectors by the superscript $'$). This specific form has the advantage that it ensures positivity of H_t even for $K = 1$. An application is shown in Fig. 10. We will not go into the details of multivariate GARCH here, instead the reader is referred to the excellent review by Bauwens et al. (2006), yet we will discuss some of the problems related to large sample sizes (see also Raddant and Wagner 2022) and possible applications to financial networks.

To start with, when the number of stocks N is large, the high number of so-called nuisance parameters in \bar{H} can lead to biases in the maximum likelihood estimation. More important in our case however is the estimation of the dynamics of H_t , which runs into similar problems for large samples. Most models therefore try to limit the number of parameters that describe the covariances of assets. For example, in the VEC model (Bollerslev et al. 1988) the elements of H_t are treated as a $N(N + 1)/2$ vector. In its scalar version (Ding and Engle 2001), the dynamics are reduced to two parameters. The K -factor BEKK model (Engle and Kroner 1995) shown above uses N -dimensional matrix multiplication.

Additionally, a computational problem is that most estimation methods require H_t to be inverted repeatedly during the estimation. This can further limit the application to large samples. This problem can partly be circumvented by the use of covariance targeting (Engle and Mezrich 1996). Here \bar{H} is replaced by the observed time-averaged covariance. This method also allows for covariances and GARCH parameters to be estimated in two separate steps.

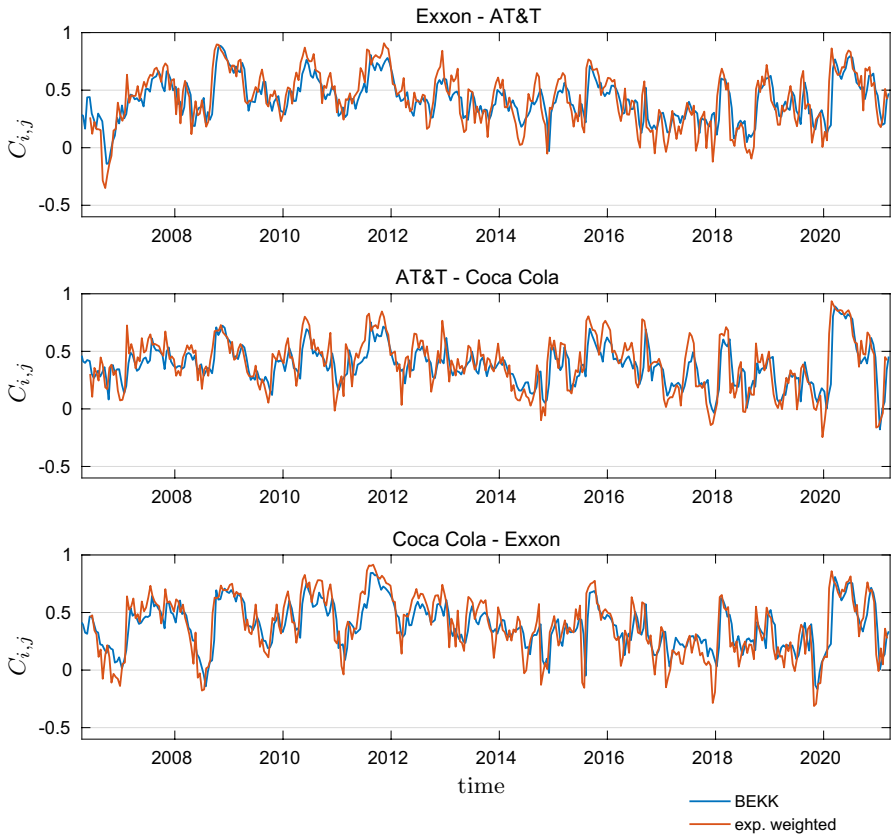


Fig. 10 Stock correlations over time. The plots show comparisons of the correlations estimated with a multivariate GARCH model (using a 10-day average) and those calculated from using the method of exponentially weighted correlations (with a time window of $\Delta t = 60$ days and $\theta = 20$)

Several approaches have worked on the scalability of the estimation of multivariate GARCH models, most notably Aielli (2013) who proposed improvements in the modeling of the conditional covariance matrix and Engle et al. (2021) who developed an approach that uses a composite likelihood and builds on the estimation of pairs of assets. A different solution is the method proposed by Engle et al. (2019) which renders improved estimates by using nonlinear shrinkage.

A different approach to deal with large samples and to reduce the number of parameters is to use factor models. An early example is the OGARCH model by Alexander (2001) which allows to drastically reduce the number of parameters by describing the cross section of asset returns by a number of factors K that can typically be chosen much smaller than N . This reduces the number of parameters that have to be estimated substantially. A more involved variant is the GO-GARCH model (van der Weide 2002). It is applied to the de-correlated returns. The principal components of these are then

rotated by a K -dimensional rotation. The $K(K - 1)/2$ angles are additional parameters in the estimation.

Yet another approach to reduce the number of parameters is to restrict the form for H_t as in the DECO model (Engle and Kelly 2012). It assumes identical yet time-varying correlations between all pairs of stocks and one calculates the covariance matrix based on univariate de-garched returns. An alternative way to restrict H_t is presented by Raddant and Wagner (2022), they employ a dynamic two-factor model and a system of recursion equations to reduce the dynamics.

7.4 Application

An interesting approach to apply the GARCH framework in a network setting has been presented by Isogai (2016). He derived dynamic networks of 50 stocks in the Japanese market using the DCC framework. As mentioned above, this framework is advantageous for large N since it can be estimated in two steps. A drawback is that by construction the conditional correlations all follow the same dynamics.

Hence, in this case we model the vector of returns r_t as

$$r_t = \mu_t + H_t^{1/2} z_t, \tag{18}$$

where the residuals $\varepsilon_t = H_t^{1/2} z_t$, $E(z_t) = 0$ and $Var(z_t) = I_N$. The $N \times N$ matrix H_t corresponds to the conditional covariance of r_t . $H_t^{1/2}$ is positive definite.

The mean μ_t is modeled as an autoregressive moving average model depending on lagged returns and volatility. The volatility part is given by

$$\tilde{H}_t = \Omega + \sum_{i=1}^q S_i \varepsilon_{t-i} \odot \varepsilon_{t-i} + \sum_{j=1}^p T_j \tilde{H}_{t-j}, \tag{19}$$

where \tilde{H}_t is the diagonalized version of H_t . S and T are also diagonal matrices, \odot is the Hadamard (element-wise) product. The actual dynamic correlation R_t of the model is calculated via a proxy variable Q_t (that again depends on the lagged z and Q) to ensure that R_t stays positive definite, which leads to the well-known sandwich structure where

$$R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2}. \tag{20}$$

These dynamic correlations can then be used to describe the relationships between stocks over time. Isogai (2016) shows that these correlations will in many cases differ from correlations calculated in rolling windows applied to r_t .

8 Pairwise estimation of dependencies

8.1 Estimating feedback

A seminal work in the literature on the estimation of dependencies of time-series is the paper by Granger (1969). This paper discusses, in a rather general sense,

approaches to identify influences of co-evolving processes, also called feedback. The idea that became famous however was the two-variable model that describes the influence of two stationary time series x_t and y_t (stationary and with zero mean), which can be modeled by their lagged values, hence

$$x_t = \sum_{j=1}^m a_j x_{t-j} + \sum_{j=1}^m b_j y_{t-j} + \varepsilon_t \quad (21)$$

$$y_t = \sum_{j=1}^m c_j x_{t-j} + \sum_{j=1}^m d_j y_{t-j} + \eta_t, \quad (22)$$

where ε and η are uncorrelated noise.

In this case, the influences from x to y and vice versa are described by the parameters a , b , c and d . Today we would refer to this specification as a vector autoregressive model (an alternative would be to allow for contemporaneous effects). When we restrict the model to the case $j = 1$ we arrive as what is often stated as a model that tests *Granger causality*. This expression should not be taken too literally. It refers to the idea that if the values of one time-series explain the lagged values of the other, then this temporal relationship might also indicate some fundamental relationship between the variables (see also Zellner 1962; Lütkepohl 1991).

8.2 Application of Granger-causality

There have been several applications of Granger-causality to financial time series, yet the most interesting and complete one in the context of financial networks is probably the one by Billio et al. (2012). They analyze Granger-causality for a sample of 100 hedge funds, brokers, banks and insurers for the time before and during the 2008 financial crisis. For the time series of returns i and j they test the significance of their connection by estimating a_i , a_j , b_{ij} and b_{ji} , hence

$$r_{i,t} = a_i r_{i,t-1} + b_{ij} r_{j,t-1} + \varepsilon_{i,t} \quad (23)$$

$$r_{j,t} = a_j r_{j,t-1} + b_{ji} r_{i,t-1} + \varepsilon_{j,t}. \quad (24)$$

If the b -parameters are significantly different from zero, one can speak of a relationship between i and j . This relationship can be uni-directional or bi-directional.

There are two aspects to note for this model. First, it is not at all likely to have significant relationships that explain the return of another asset in the future (the study uses monthly returns). One would expect that arbitrage and a certain level of market efficiency would prohibit this. In fact the study by Billio et al. (2012) shows that at most times the number of connections that is significant at a 5% level is only slightly larger than 5%. This however changes in times of pronounced market stress, in this case at the end of 2007. Hence, while this method could show the effects of

the financial crisis relatively well, applications to other time periods can be more challenging.

The second aspect is of course related to the long memory in asset prices. Similarities in volatility changes of the assets would drastically influence the estimation results for all parameters and would lead to a misrepresentation of connectivity. The analysis is therefore not performed with the raw returns but the authors assume a GARCH(1,1) model. In practice, we can assume that the returns have been de-garched (see Eq. 16) before estimating the parameters for the model in Eq. 23.

8.3 Pairwise estimation with a larger dataset

A related approach to measure financial dependencies has been proposed by Raddant and Kenett (2021). Here the estimation is based on synchronous daily and weekly returns of a much larger sample of stocks from different markets. Also here we assume that the returns follow a random process with $\varepsilon_t = v_t \sqrt{h_t}$, where v_t is white noise and h follows a GARCH(1,1) process.

The conditional variance h_t is then used to produce filtered returns r^f , such that

$$r_{i,t}^f = \frac{r_{i,t}}{\sqrt{h_{i,t}}} . \quad (25)$$

By de-garching the returns, we obtained time series that can be treated in an almost standard regression framework, since performing pairwise regressions of all the filtered returns generates a measure for the comovement. It should however be noted that the distributions of the filtered returns are not perfectly normally distributed. It is therefore appropriate to assume t-distributed errors and apply what is mostly called the robust regression approach (Lange et al. 1989) (the distributions of filtered returns are closer to a Normal distribution for monthly returns).

Hence, to measure dependence we estimate the dependencies for all pairs of stocks i, j as

$$r_{i,t}^f = \beta_{0,ij} + \beta_{1,ij} r_{j,t}^f + \varepsilon_t . \quad (26)$$

The results of the estimation can now be mapped into a network representation. We choose a threshold in terms of the significance level and construct networks in which connections are defined by observing a significantly positive $\beta_{1,ij}$. For constructing weights we can use the corresponding p-values, namely p_{ij} (of the null-hypothesis significance test for $\beta_{1,ij}$), that are obtained in the course of the estimation. The exact values of the link weights A_{ij} correspond to the extend to which p_{ij} exceeds a given threshold γ , i.e.,

$$A_{ij} \propto (\gamma - p_{ij}) . \quad (27)$$

This adjacency matrix has a weighted positive entry if stocks i and j are significantly linked, measured by the estimated conditional correlation.

In this particular case where the assets under consideration come from different markets from all around the globe the time difference in trading hours are an

additional complication. The estimated dependencies based on daily returns may be biased downwards because the time series are asynchronous (see also Martens and Poon (2001); Christensen et al. (2010); Hayashi and Yoshida (2008)). In general, this can be dealt with in two ways: by using tick data and calculation of synchronous pseudo closing prices, or by time aggregation. The approach by Raddant and Kenett (2021) uses an approximate correction for this effect that is applied to the p-values derived from the estimation. The method assumes that the estimation results based on the weekly data are unbiased. The results from the daily data are thus adjusted in such a way that they are consistent with the dependencies obtained from the weekly data.

Since the sample is rather large, it is for most purposes not useful to report results only on the level of stocks. A dimensional reduction can be achieved by hierarchically aggregating the results on the level of sectors and countries (according to GICS and TRBC classification). The resulting network is shown in Fig. 11. The p-values have been averaged on a sector-by-sector and country-by-country level. The visualization shows that many sectors of 'western' markets are globally integrated, while some markets like China, Japan and India remain more segregated.

9 Variance decomposition

In the previous section we have already touched upon the use of vector autoregressive models (VAR) for the analysis of dependencies in financial markets. While many applications of such models to financial markets and macroeconomic performance exist, these have mostly been cases with samples of limited size. One reason is that once N becomes large, it can become difficult to estimate all possible influences on a certain time-series jointly. The other reason is that one is typically interested in obtaining an orthogonalized version of the shocks to calculate impulse responses or decomposed forecast errors, which can become troublesome even for moderate N (see also Koop et al. 1996; Pesaran and Shin 1996, 1998).

A decisive improvement on this issue is the variance decomposition proposed in Diebold and Yilmaz (2009). As in any case, we will first estimate a standard VAR model for a specific sample or time window with a certain number of lags. Then we can evaluate in how far each variable contributes to the generalized forecast error variance decomposition of the other variables. For the application shown in Diebold and Yilmaz (2014), the idea is that the sums of these contributions can then for example be used to calculate contributions to market volatility and are also seen as a proxy for contributions to systemic risk.

We assume that the underlying process x_t that describes, for example, a vector with daily volatility of stocks, follows $x_t = \Theta(L)u_t$, where

$$\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_L^2. \quad (28)$$

Hence, x_t depends on the lagged orthogonal shocks u_t in a way that is described by the parameters contained in the Θ matrices. L is the lag operator. The decisive part of this approach is now to calculate the generalized variance decomposition matrix

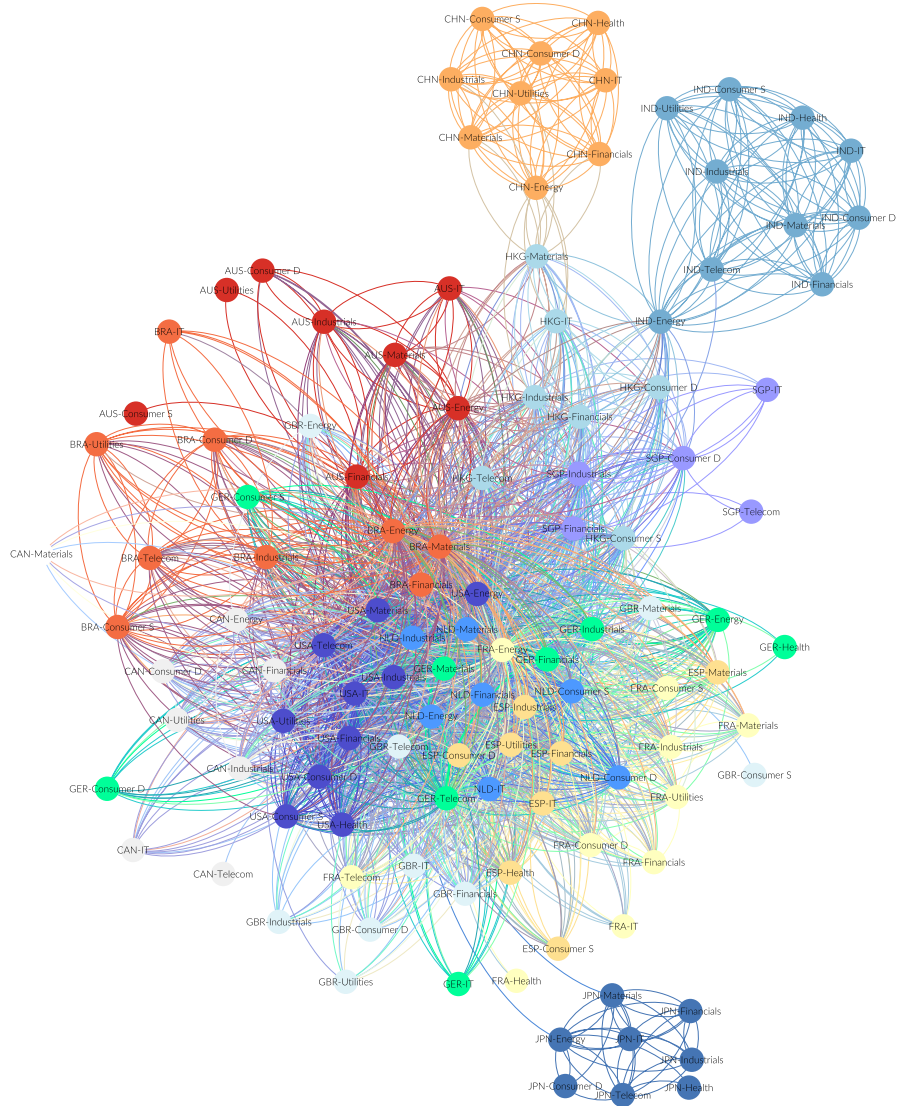


Fig. 11 Visualization of pairwise estimated stock dependencies, averaged on the sector level, color-coding applied by country. Stocks from the materials, energy and financial sector, mainly from the US (purple), Germany (green), the UK (gray), The Netherlands (blue), France (yellow) and Spain (beige), form the densely connected center of this network. The markets of Canada, Brasil, Australia, China, India, Singapore and Hong Kong show in mostly peripheral positions

D^H with entries d_{ij}^H that specify how stock i 's H step ahead forecast error variance is influenced by stock j (where $i \neq j$). These entries are given by

$$d_{ij}^H = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' \Theta_h \Omega e_j)^2}{\sum_{h=0}^{H-1} (e_i' \Theta_h \Omega \Theta_h' e_j)}, \quad (29)$$

where e_j is a vector where the j th element is 1 and the others are zeros. Θ_h is the coefficient matrix multiplying the h -lagged shock vector in the infinite moving-average representation of the non-orthogonalized VAR. Ω is the covariance matrix of the shocks and σ_{jj} is the j th element of the diagonal of Ω .

This variance decomposition matrix can easily be interpreted as an adjacency matrix. The matrix is directed, because the elements d_{ij} and d_{ji} are typically not the same. The sums over the rows of this adjacency matrix sum to 1 since they resemble the shares of received contributions to the variance decomposition. To account for the remainder of these contributions it makes sense to write the diagonal elements as $A_{ii} = \sum_{j:j \neq i} A_{ij}$.

While this approach presents a useful connection from the class of VAR models toward network models, there are also some weaknesses. As Chan-Lau (2017) and Lanne and Nyberg (2016) point out, the generalized forecast error variance decomposition in the form presented in the original paper can lead to inconsistencies when comparing the contributions of firms to systemic risk over time. Improvements in the method can however mostly mitigate this issue. Another issue is the still somewhat limited scalability with respect to the sample size, which is discussed by Greenwood-Nimmo et al. (2021).

A further development in this class of models is the so-called TVP-VAR (time-varying parameter vector autoregressive model). Here the variance-covariance matrix is allowed to vary by applying a Kalman filter estimation with forgetting factors. This means that it is no longer necessary to choose time windows and window sizes. For the details, the reader may refer to Koop and Korobilis (2014); Lütkepohl et al. (2015); Korobilis and Yilmaz (2018); Antonakakis et al. (2018) and Antonakakis et al. (2020).

10 Conclusions

This survey has shown that a variety of viewpoints exist in analyzing financial dependencies. While some approaches focus on topological network representation constraints, others are more focused on a modeling of details of the time-series properties of returns. This has of course implications for the scalability to larger datasets and its applicability to different current issues in financial market research. Nevertheless, this survey has also shown that the field has moved beyond a phase of exploration of asset market data.

Therefore we encourage researchers to take into account results outside of their domains and to apply network-related research to problems that are relevant from a societal, economic or political perspective. We believe that a collaborative effort among different disciplines will be the key to solve many current challenges.

Appendix A Correlation-based graphs

See Fig. 12.

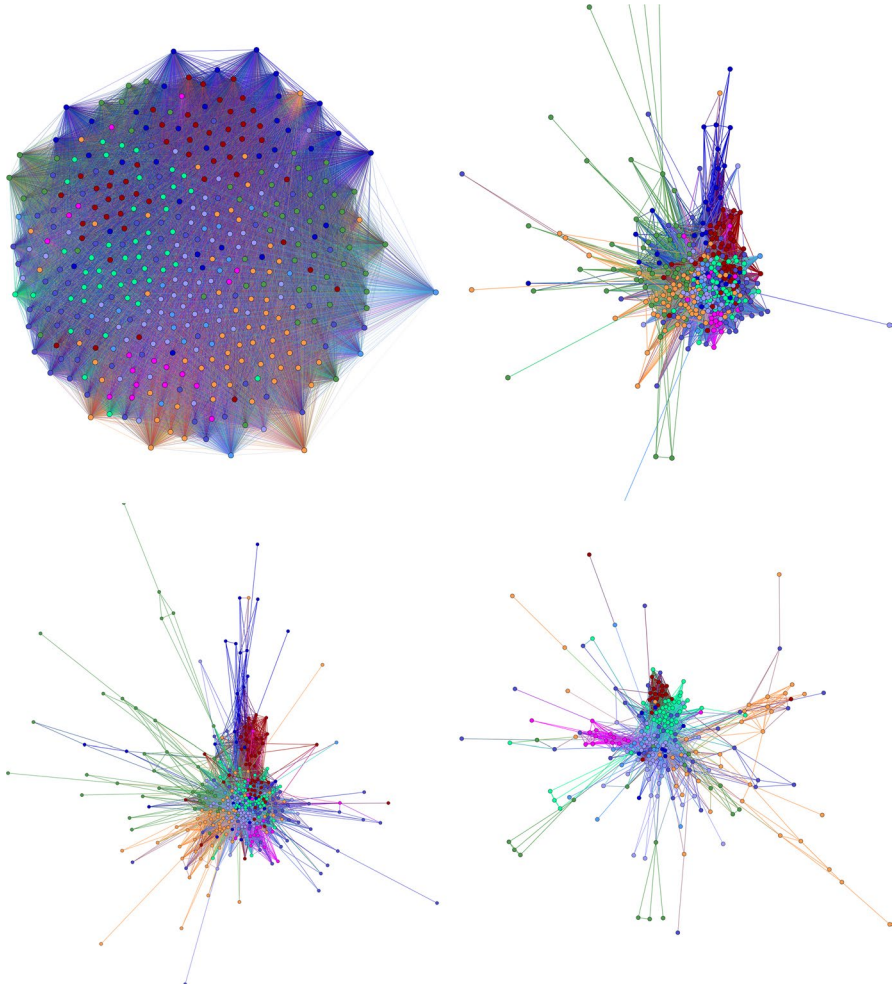


Fig. 12 Complete network and networks with links above the 60%, 80% and 95% interval of the distribution of correlation coefficients. The giant connected components consist of 404, 397, 372, 240 nodes and 81,406, 32,562, 16,279, 3,727 links. Color-coding by sector classification: pink—Energy, bright blue—Basic Materials, bright purple—Industrials, dark purple—Cyclical Consumer Goods, dark blue—Non-cyclical Consumer Goods, bright green—Financials, dark green—Healthcare, orange—Technology, brown—Utilities

Acknowledgements We thank Wladimir Ostrovsky and two anonymous reviewers for their careful reading and comments on this paper.

Funding Open access funding provided by Danube University Krems University for Continuing Education.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acemoglu D, Ozdaglar A, Tahbaz-Salehi A (2015) Systemic risk and stability in financial networks. *Am Econ Rev* 105:564–608
- Aielli GP (2013) Dynamic conditional correlation: on properties and estimation. *J Bus Econ Stat* 31(3):282–299
- Alexander C (2001) *Market models*. Wiley, New York
- Antonakakis N, Gabauer D, Gupta R et al (2018) Dynamic connectedness of uncertainty across developed economies: a time-varying approach. *Econ Lett* 166:63–75
- Antonakakis N, Chatziantoniou I, Gabauer D (2020) Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *J Risk Financ Manag* 13(4):84
- Aste T, Matteo TD, Hyde ST (2005) Complex networks on hyperbolic surfaces. *Phys A* 346:20–26
- Aste T, Shaw W, Matteo TD (2010) Correlation structure and dynamics in volatile markets. *New J Phys* 12(8):085009
- Aste T, Gramatica R, Di Matteo T (2012) Exploring complex networks via topological embedding on surfaces. *Phys Rev E* 86:036109
- Barberis N, Shleifer A, Wurgler J (2005) Comovement. *J Financ Econ* 75(2):283–317
- Bardoscia M, Barucca P, Battiston S et al (2021) Physics of financial networks. *Nat Rev Phys* 3:490–507
- Bastian M, Heymann S, Jacomy M (2009) Gephi: An open source software for exploring and manipulating networks. In: *International AAAI conference on weblogs and social media*
- Battiston S, Farmer JD, Flache A et al (2016) Complexity theory and financial regulation. *Science* 351(6275):818–819
- Bauwens L, Laurent S, Rombouts JV (2006) Multivariate GARCH models: a survey. *J Appl Economet* 21(1):79–109
- Bekaert G, Hodrick RJ, Zhang X (2009) International stock return comovement. *J Financ* 64(6):2591–2626
- Bekaert G, Campbell RH, Lundblad CT et al (2011) What segments equity markets? *Rev Financ Stud* 24(12):3841–3890
- Bien J, Tibshirani RJ (2011) Sparse estimation of a covariance matrix. *Biometrika* 4(98):807–820
- Billio M, Getmansky M, Lo AW et al (2012) Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J Financ Econ* 104(3):535–559
- Bodurtha J, Mark N (1991) Testing the CAPM with time-varying risk and returns. *J Financ* 46:1485–1505
- Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *J Econ* 31:307–327
- Bollerslev T, Engle RF, Wooldridge JM (1988) A capital asset pricing model with time-varying covariances. *J Polit Econ* 96(1):116–131

- Bouchaud JP, Potters M (2009a) Financial applications of random matrix theory: a short review. [arXiv:0910.1205](https://arxiv.org/abs/0910.1205)
- Bouchaud JP, Potters M (2009b) Theory of financial risk and derivative pricing: from statistical physics to risk management. Cambridge UP
- Bracker K, Docking DS, Koch PD (1999) Economic determinants of evolution in international stock market integration. *J Empir Financ* 6:1–27
- Buldyrev SV, Parshani R, Paul G et al (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 6:1025–1028
- Caldarelli G (2012) Scale-free networks: complex webs in nature and technology. Oxford University Press, Oxford
- Caldarelli G, Wolf S, Moreno Y (2018) Physics of humans, physics for society. *Nat Phys* 14(870):1745–2481
- Chan-Lau JA (2017) Variance decomposition networks: Potential pitfalls and a simple solution. IMF working paper
- Christensen K, Kinnebrock S, Podolskij M (2010) Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *J Econom* 159(1):116–133
- Cohen L, Frazzini A (2008) Economic links and predictable returns. *J Financ* 63:1977–2011
- Cont R (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantit Financ* 1:223–236
- Dacorogna MM, Gencay R, Mueller UA et al (2001) An introduction to high-frequency finance. Academic Press, Cambridge
- Diebold FX, Yilmaz K (2009) Measuring financial asset return and volatility spillovers, with application to global equity markets. *Econ J* 119:158–171
- Diebold FX, Yilmaz K (2014) On the network topology of variance decompositions: measuring the connectedness of financial firms. *J Econom* 182(1):119–134
- Ding Z, Engle RF (2001) Large scale conditional covariance matrix modeling, estimation and testing. working paper FLN-01-029, NYU Stern School of Business
- Ding Z, Granger CW, Engle RF (1993) A long memory property of stock market returns and a new model. *J Empir Financ* 1:83–106
- Dutt P, Mihov I (2013) Stock market comovement and industrial structure. *J Money, Credit, Bank* 45(5):891–911
- Elliott M, Golub B, Jackson MO (2014) Financial networks and contagion. *Am Econ Rev* 104:3115–53
- Engle RF (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J Bus Econ Stat* 20(3):339–350
- Engle RF, Kelly B (2012) Dynamic equicorrelation. *J Bus Econ Stat* 30:212–228
- Engle RF, Kroner KF (1995) Multivariate simultaneous generalized ARCH. *Econ Theor* 11:122–150
- Engle RF, Mezrich J (1996) GARCH for groups. *Risk* 9:36–40
- Engle RF, Ledoit O, Wolf M (2019) Large dynamic covariance matrices. *J Bus Econ Stat* 37:363–375
- Engle RF, Paker C, Shephard KK et al (2021) Fitting vast dimensional time-varying covariance models. *J Bus Econ Stat* 39(3):652–668
- Fama E, French K (1992) The cross-section of expected stock returns. *J Financ* 47(2):427–465
- Forbes K, Rigobon R (2002) No contagion, only dependence: measuring stock market comovements. *J Financ LVI* 5(5):2223–2261
- Forbes KJ, Chinn MD (2004) A decomposition of global linkages in financial markets over time. *Rev Econ Stat* 86(3):705–722
- Fry R, Martin VL, Tang C (2010) A new class of tests of contagion with applications. *J Bus Econ Stat* 28(3):423–437
- Gower JC, Ross GJS (1969) Minimum spanning trees and single linkage cluster analysis. *J R Stat Soc Ser C* 18(1):54–64
- Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3):424–438
- Green TC, Hwang BH (2009) Price-based return comovement. *J Financ Econ* 93(1):37–50
- Greenwood-Nimmo M, Nguyen VH, Shin Y (2021) Measuring the connectedness of the global economy. *Int J Forecast* 37:899–919
- Gualdi S, Cimini G, Primicerio K et al (2016) Statistically validated network of portfolio overlaps and systemic risk. *Sci Rep* 6:39467
- Hansen PR, Lunde A (1998) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *J Appl Economet* 20:873–889

- Hayashi T, Yoshida N (2008) Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Ann Inst Stat Math* 60(2):367–406
- Hu Y (2005) Efficient and high quality force-directed graph drawing. *Math J* 10(1):37–71
- Isogai T (2016) Building a dynamic correlation network for fat-tailed financial asset returns. *Appl Netw Sci* 1(7):1–24
- Jobson DJ (1991) *Appl Multivar Data Anal*. Springer, New York
- Kenett D, Shapira Y, Madi A et al (2011) Index cohesive force analysis of NY market reveal phase transition into stiff market state. *PlosOne* 6(4):e19378
- Kenett DY, Raddant M, Lux T et al (2012) Evolvement of uniformity and volatility in the stressed global financial village. *PLoS ONE* 7(2):e31144
- Koop G, Korobilis D (2014) A new index of financial conditions. *Eur Econ Rev* 71:101–116
- Koop G, Pesaran HH et al (1996) Impulse response analysis in nonlinear multivariate models. *J Econom* 74:119–147
- Korobilis D, Yilmaz K (2018) Measuring dynamic connectedness with large bayesian VAR models. *SSRN* 3099725
- Kowalski CJ (1972) On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *J R Stat Soc Ser C* 21(1):1–12
- Kruskal JB (1956) On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc Amer Math Soc* 7:48–50
- Kumar S, Di Matteo T, Chakrabarti AS (2020) Disentangling shock diffusion on complex networks: identification through graph planarity. *J Complex Netw* 8(3):cnaa023
- Kuratowski K (1930) Sur le problème des courbes gauches en topologie. *Fund Math* 15:271–283
- Kutner R, Ausloos M, Grech D et al (2019) Econophysics and sociophysics: their milestones challenges. *Phys A* 516(9):240–253
- Laloux L, Cizeau P, Bouchaud J et al (1999) Noise dressing of financial correlation matrices. *PhysRev-Lett* 83:1467
- Lancichinetti A, Fortunato S (2009) Community detection algorithms: a comparative analysis. *Phys Rev E* 80:056117
- Lange KL, Little RJ, Taylor JM (1989) Robust statistical modeling using the t distribution. *J Am Stat Assoc* 84(408):881–896
- Lanne M, Nyberg H (2016) Generalized forecast error variance decomposition for linear and nonlinear multivariate models. *Oxford Bull Econ Stat* 78(4):595–603
- Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev Econ Stat* 47:13–37
- Livan G, Alfarano S, Scalas E (2011) Fine structure of spectral properties for random correlation matrices: an application to financial markets. *Phys Rev E* 84:016113
- Lütkepohl H (1991) *Introduction to multiple time series analysis*. Springer, Berlin
- Lütkepohl H, Staszewska-Bystrova A, Winker P (2015) Comparison of methods for constructing joint confidence bands for impulse response functions. *J Stat Mech* 31:782–798
- MacMahon M, Garlaschelli D (2015) Community detection for correlation matrices. *Phys Rev X* 5:021006
- Mantegna RN (1999) Hierarchical structure in financial markets. *Eur Phys J B* 11:193–197
- Mantegna RN, Stanley HE (1999) *Introduction to econophysics: correlations and complexity in finance*. Cambridge University Press, Cambridge
- Marčenko V, Pastur L (1967) Distribution of eigenvalues for some sets of random matrices. *MathUSSR-Sb* 1(457)
- Martens M, Poon SH (2001) Returns synchronization and daily correlation dynamics between international stock markets. *J Bank Financ* 25(10):1805–1827
- Massara GP, Di Matteo T, Aste T (2017) Network filtering for big data: triangulated maximally filtered graph. *J Complex Netw* 5(2):161–178
- Masuda N, Kojaku S, Sano Y (2018) Configuration model for correlation matrices preserving the node strength. *Phys Rev E* 98:012312
- Musmeci N, Aste T, Di Matteo T (2015) Relation between financial market structure and the real economy: comparison between clustering methods. *PLoS ONE* 10(3):e0116201
- Musmeci N, Aste T, Di Matteo T (2015) Risk diversification: a study of persistence with a filtered correlation-network approach. *J Netw Theory Financ* 1(1):1–22
- Musmeci N, Nicosia V, Aste T, et al (2017) The multiplex dependency structure of financial markets. *Complexity* 9586064

- Newman MEJ (2010) *An introduction networks*. Oxford University Press, New York
- Onnela JP, Chakraborti A, Kaski K et al (2003) Dynamic asset trees and black monday. *Phys A* 324:247–252
- Onnela JP, Chakraborti A, Kaski K et al (2003) Dynamics of market correlations: taxonomy and portfolio analysis. *Phys Rev E* 68:056110
- Pentland A (2014) *Social physics: How good ideas spread—the lessons from a new science*. Penguin
- Pesaran M, Shin Y (1996) Cointegration and speed of convergence to equilibrium. *J Econome* 71:117–43
- Pesaran M, Shin Y (1998) Generalized impulse response analysis in linear multivariate models. *Econ Lett* 58(1):17–29
- Plerou V, Gopikrishnan P, Amaral L et al (1999) Scaling of the distribution of price fluctuations of individual companies. *Phys Rev E* 60:6519–29
- Plerou V, Gopikrishnan P, Rosenow B et al (2002) Random matrix approach to cross correlations in financial data. *Phys Rev E* 65:066126
- Pozzi F, Di Matteo T, Aste T (2012) Exponential smoothing weighted correlations. *Eur Phys J B* 175:85
- Prim RC (1957) Shortest connection networks and some generalizations. *Bell Syst Tech J* 36(6):1389–1401
- Raddant M, Kenett DY (2021) Interconnectedness in the global financial market. *J Int Money Financ* 110:102280
- Raddant M, Wagner F (2017) Transitions in the stock markets of the US, UK and Germany. *Quant Financ* 17(2):289–297
- Raddant M, Wagner F (2022) Multivariate GARCH with dynamic beta. *Eur J Financ* 28(13–15):1324–1343
- Rigobon R (2003) On the measurement of the international propagation of shocks: Is the transmission stable? *J Int Econ* 61:261–283
- Sedgewick R, Wayne K (2011) *Algorithms*. Addison-Wesley, Boston
- Shapira Y, Kenett Y, Ben-Jacob E (2009) The index cohesive effect on stock market correlations. *Eur Phys J B* 72(4):657–669
- Sharpe W (1963) A simplified model for portfolio analysis. *Manage Sci* 9:277–293
- Sharpe WF (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *J Financ* 19:425–442
- Song D, Tumminello M, Zhou W et al (2011) Evolution of worldwide stock markets, correlation structure and correlation based graphs. *Phys Rev E* 84(2):026108
- Song WM, Di Matteo T, Aste T (2011) Nested hierarchy in planar graphs. *Discret Appl Math* 159:2135–2146
- Song WM, Di Matteo T, Aste T (2012) Hierarchical information clustering by means of topologically embedded graphs. *PLoS ONE* 7(3):e31929
- Steinbacher M, Raddant M, Karimi F et al (2021) Advances in the agent-based modeling of economic and social behavior. *SN Bus Econ* 1:99
- Summer M (2013) Financial contagion and network analysis. *Annu Rev Financ Econ* 5(1):277–297
- Tsay RS (2010) *Analysis of financial time series*. Wiley, New York
- Tumminello M, Aste T, Di Matteo T et al (2005) A tool for filtering information in complex systems. *PNAS* 102(30):10421–10426
- Tumminello M, Aste T, Di Matteo T et al (2007) Correlation based networks of equity returns sampled at different time horizons. *Eur Phys J B* 55:209–217
- Tumminello M, Lillo F, Mantegna RN (2010) Correlation, hierarchies, and networks in financial markets. *J Econ Behav Organ* 75:40–58
- Tumminello M, Miccichè S, Lillo F et al (2011) Statistically validated networks in bipartite complex systems. *PLoS ONE* 6(3):1–11
- Verma A, Vivo P, Di Matteo T (2019) A memory-based method to select the number of relevant components in principal component analysis. *J Stat Mech* 9:093408
- van der Weide R (2002) GO-GARCH: A multivariate generalized orthogonal garch model. *J Appl Economet* 17:549–564
- Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J Am Stat Assoc* 57(298):348–368
- Zou H (2006) The adaptive lasso and its oracle properties. *J Am Stat Assoc* 476(101):1418–1429