

# A network-based analysis for evaluating conditional covariance estimates

Andrea Scozzari (✉ [andrea.scozzari@unicusano.it](mailto:andrea.scozzari@unicusano.it))

Università degli Studi Niccolò Cusano <https://orcid.org/0000-0003-3038-3957>

Carlo Drago

Università degli Studi Niccolò Cusano

---

## Research Article

**Keywords:** Network analysis, Mining financial data, Maximal cliques, Multivariate GARCH models, BEKK, DCC, Targeting

**Posted Date:** March 15th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1449824/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# A network-based analysis for evaluating conditional covariance estimates.

Carlo Drago<sup>1</sup>, Andrea Scozzari<sup>1,2</sup>

<sup>1</sup> Università degli Studi Niccolò Cusano Roma, carlo.drago@unicusano.it

<sup>2</sup> CESDE, Centro Studi per l'Analisi delle Dinamiche Economiche  
Università Niccolò Cusano, andrea.scozzari@unicusano.it

March 14, 2022

## Abstract

Modeling and forecasting of dynamically varying covariances have received much attention in the literature. The two most widely used conditional covariances and correlations models are the BEKK and the DCC ones. In this paper, we advance a new method based on Network analysis and a new targeting approach for both the two above models with the aim of better estimate covariance matrices associated with financial time series. Our approach is based on specific groups of highly correlated assets in a financial market, and assuming that these relationships remain unaltered at least in the long run. Based on the estimated parameters, we evaluate our targeting method on simulated series by referring to two well-known loss functions introduced in the literature. Furthermore we find and analyze all the maximal cliques in correlation graphs to evaluate the effectiveness of our method. Results from an empirical case study are encouraging, mainly when the number of assets is not large.

**Keywords:** Network analysis, Mining financial data, Maximal cliques, Multivariate GARCH models, BEKK, DCC, Targeting.

## 1 Introduction

Modeling financial volatility is a crucial issue in empirical finance. *Volatility clustering* is a peculiar feature of a financial time series, that is, the series is characterized by periods more

volatile than others due to unpredictable outer events, or, in other words, the series suffers from *heteroscedasticity*. The existence of heteroscedasticity is a major concern in a univariate approach aiming at estimating and forecasting a given (financial) phenomenon. In 1982 Robert Engle [22] introduced formal methods of analyzing economic time series with time-varying volatility named the Auto Regressive Conditional Heteroscedasticity (ARCH) model. This model was further generalized by Bollerslev in [8] and renamed as the Generalized ARCH (GARCH) model. Modeling financial volatility becomes a more difficult task if we consider  $N$  time series (multivariate approach) and we are interested in understanding the co-movements of financial returns. It is therefore important to extend the analysis to multivariate GARCH (MGARCH) models. In fact, these models play an important role in different fields of finance like, for instance, in portfolio optimization [24], option pricing [39], analysis of contagion and volatility spillovers [7, 16]. In the literature MGARCH models are classified into different categories (see, e.g., [5]): (1) direct generalizations of the univariate GARCH models; (2) linear combinations of univariate GARCH models, such as, generalized orthogonal models and latent factor models; (3) nonlinear combinations of univariate GARCH models; (4) nonparametric and semiparametric models that constitute an alternative to the parametric estimation of the financial volatility that do not impose a particular structure on the data. In this paper we will focus on categories (1) and (3).

In finance, the main aim of multivariate GARCH models is to predict the future values of the variance and covariance matrix of asset returns. The literature has received much attention for modeling and forecasting dynamically varying covariances, with numerous multivariate ARCH, GARCH, and stochastic volatility specifications proposed for this problem. Research on these models has been lively ever since 1986. Readers may refer to the very recent paper [46] and the references for a comprehensive review of the existing literature on MGARCH approaches.

Modeling a covariance matrix is difficult because of the likely high dimensionality of the problem and the constraint that a covariance matrix must be positive definite. The first developed MGARCH models of category (1) were the *VEC* model introduced in [11], and the *BEKK* one proposed in [26]. Both approaches deal with the direct estimation of the covariance matrix between the assets returns. One disadvantage of the *VEC* model is that there is no guarantee of a positive semi-definite estimated matrix. On the other hand, the *BEKK* model has the attractive property that the conditional covariance matrices are positive definite by construction. However, estimation of a *BEKK* model involves somewhat heavy computations due to it contains a large number of parameters. To overcome this problem, one approach was to decompose the conditional covariance matrix into conditional standard deviations and a conditional correlation matrix (models of category (3)). In the first model proposed in [9] the conditional correlation is assumed to be constant over time, and only the conditional standard deviation is time-varying,

hence the name Constant Conditional Correlation (*CCC-GARCH*) model. The restriction of constant correlations over time was not considered always reasonable and was later removed. In [27] and [23] the authors defined the *Dynamic* Conditional Correlation model (*DCC-GARCH*), which is an extension of the *CCC-GARCH* one, for which the conditional correlation matrix is designed to vary over time. This approach has one important advantage, that is, the number of parameters to be simultaneously estimated is reduced since a complex optimization problem is disaggregated into simpler ones.

Several authors provided a number of different modifications of conditional covariance MGARCH models, some of them are mentioned in [46]. Other meaningful approaches explicitly incorporates the effect of measurement errors and time-varying attenuation biases into the covariance estimation and forecasts (see, e.g., [12]). Engle and Kelly [25] studied a special case of the *DCC-GARCH* model, the *dynamic equicorrelation* model, in which all correlations are identical. This can be a useful model in situations where the number of parameters to estimate are large.

All the MGARCH models presented in the literature estimate the conditional covariance or correlation matrices via the optimization of an appropriate log-likelihood function. Once having obtained the parameters, the following step consists in an estimation evaluation phase, which is performed with several methods. In particular, when comparing different approaches, the estimations are usually compared w.r.t. some measures like the Mean Absolute Error (MAE) and/or the Mean Square Error (MSE) (see, e.g., [35]). Alternatively, the performance of the estimated conditional covariance and correlation matrices is measured by using some *loss* functions that measure the distance from the fitted covariance matrix from the ex-post realized one. Among others we mention the Frobenius distance, the quasi-likelihood measure, and the Kullback-Leibler divergence [12, 30]. Finally, in order to determine the goodness of fit of the residuals, univariate and/or multivariate approaches consisting of the application of a number of different tests are performed, such as, for example the Ljung-Box test.

In this paper we advance a novel approach based on network analysis for evaluating the estimates of the time-varying correlation matrices in financial markets. First, we consider the two most widely used MGARCH models, that is, the BEKK and DCC ones, and propose a variant of these two models by suitably modifying the log-likelihood function to maximize. We call *modified BEKK* and *modified DCC* the two resulting models, respectively. This modification consists in introducing a term in the function incorporating a loss measure based on the difference between time-varying covariance matrices and the covariance matrix estimated w.r.t. the whole in-sample period. In fact, as observed in [47], a financial market is characterized by some stylized facts. In particular, there are often specific groups of assets that are highly correlated in

such a way that, for instance, positive price changes of one asset in the group determine positive price changes of all the other assets in the group, and these relationships remain unaltered over time. Hence, the idea beyond our modification is that (extremely) high/low values of correlations observed between pairs of assets do not change too much during time, more precisely, any two highly positive or negative correlated assets in the market remain highly (positive/negative) correlated during the given observed time period.

The correlation among assets in a financial market can be pointed out via graph theory in which, given a time period, assets are identified by vertices of a complete graph  $G$  and distances (weights) assigned to pairs of assets (edges) incorporate the dependency structure of returns. Mantegna [33] was one of the first to construct asset graphs based on stocks prices correlations in order to detect the hierarchical organization inside a stock market. In order to highlight only groups of assets whose correlation is above or below a given threshold  $\delta$ , one can extract a (sub)graph  $G(\delta)$  of  $G$  that contains only a subset of the connections in  $G$  between pairs of assets. In the new graph  $G(\delta)$  two stocks are linked if and only if they exhibit high/low values of correlation. In other words, the *observed* subgraph  $G(\delta)$  reveals the strong relationships between groups of assets in a financial market during a given time period  $T$ . Examples of such interconnected groups are, for instance, firms belonging to a specific industrial sector.

In this paper, the additional term introduced in the log-likelihood function takes into account that the time-varying covariance or correlation matrices must not alter the clusters formed by highly correlated assets observed during  $T$ . Hence, with our method we estimate a modified MGARCH model, simulate  $T$  realizations of returns of each of the  $N$  assets, compute the correlation matrix w.r.t. the whole simulated in-sample period and, given the threshold  $\delta$ , we obtain the corresponding *simulated* graph  $G_S(\delta)$ . Then, we compare the observed and simulated graphs  $G(\delta)$  and  $G_S(\delta)$ , respectively. In particular, we compare all the *maximal Cliques* of the two graphs. Given a general graph  $G$ , a *Clique* is a set of interconnected stocks that forms a complete subgraph of  $G$ . Hence, Cliques in  $G(\delta)$  and in  $G_S(\delta)$  represent highly (positive/negative) correlated assets, that is, any stock that belongs to a clique is highly correlated with all other stocks in the clique. The comparisons allow us verifying if a modified MGARCH model has been able to correctly reproduce the volatility and all the strong relationships among assets of the given financial market, and if a modified model is able to outperform the corresponding original MGARCH model.

## 1.1 Motivation

A similar modification of the two above considered models through the introduction of a *target* matrix was provided in [15]. In that paper, the authors directly modified the “structure” of the BEKK and DCC models with the target matrix and compare their modified models with the aim of improving estimate of the conditional covariance and correlation matrices. The authors consider the long-run covariance matrix as a target matrix, which is consistently estimated by the corresponding sample estimator, and consider different versions of the two MGARCH models. Finally, they showed that, from a theoretical perspective, the optimal model for estimating conditional covariances was the scalar BEKK one, regardless of whether targeting was used (see [15]). In particular, they point out that, other than for the scalar BEKK model, imposing positive definiteness and covariance stationarity for other considered versions of BEKK is extremely complicated when there is more than one asset. For the DCC models, targeting can be useful to reduce the number of parameters to estimate. However, in [3] the authors proved that the estimation of DCC models with targeting may be inconsistent. Therefore, the suggestion is that one has to reconsider all models with targeting.

In any case, from a theoretical viewpoint, the paper [15] is worth of interesting since the authors study the availability of analytical forms for the sufficient conditions for consistency and asymptotic normality of the appropriate estimators and computational tractability for large problem dimensions. From an empirical point of view, in [15] no empirical tests are presented, but the authors conclude that it would not seem to be possible to provide an appropriate judgment a priori as to which of the different MGARCH models is to be preferred.

In light of the above considerations, the contribution of the present paper is threefold. First, we propose a new alternative targeting method that does not alter the models’ structure. In fact, we modify only the log-likelihood objective function by introducing a suitable loss or distance measure between the long-run sample covariance or correlation matrix and the corresponding conditional covariance or conditional correlation matrices, respectively. Hence, we do not have to impose additional constraints, for instance, for the covariance stationarity or for guaranteeing the matrices to be definite positive. Second, we provide an empirical analysis for evaluating our modified models. We do this with a recommendation that it is not our goal to definitely decide which of the different models is the best one. Rather, we seek to understand whether our modified models allow to better capture the strong relationships between assets in a stock market. At the very least, we hope that the modified models can be more accurate than the standard ones. Third, we advance a new method for evaluating the effectiveness of the estimated models through some specific tools commonly used in network analysis.

The paper is organized as follows: Section 2 reports notation and some basic definitions. Sections 2.1 and 2.2 resume the two MGARCH models considered in the paper along with the new log-likelihood functions used in the optimization/estimation phase. Section 3 reports the structure of the empirical financial dataset we employed in the experimental phase described in Section 3.1. Finally, some conclusions and further research are depicted in Section 4

## 2 Notation, definitions

Consider a financial market formed by  $N$  assets. Let  $P_{jt}$  be the daily closing price of asset  $j$  at time  $t$ ,  $t = 0, \dots, T$ , and  $r_{jt} = \log(\frac{P_{jt}}{P_{jt-1}})$  be the corresponding log-return of asset  $j$ . In the following, Pearson correlation coefficients are used to detect dependencies between assets returns. A general MGARCH model is defined:

$$\begin{cases} \mathbf{r}_t &= \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t \\ \boldsymbol{\epsilon}_t &= \mathbf{H}_t^{\frac{1}{2}} \boldsymbol{\eta}_t, \end{cases} \quad (1)$$

where  $\mathbf{r}_t$  is the  $N \times 1$  vector of log-returns at time  $t$ ,  $\boldsymbol{\epsilon}_t$  is the  $N \times 1$  vector of mean-corrected returns of  $N$  assets at time  $t$ , with  $E[\boldsymbol{\epsilon}_t]=0$  and  $\text{Cov}[\boldsymbol{\epsilon}_t] = \mathbf{H}_t$ . The vector  $\boldsymbol{\mu}_t$  represents the expected value of  $\mathbf{r}_t$ . Observe that  $\boldsymbol{\mu}_t$  may be modeled as a constant vector or as a time series model. In this paper we assume  $\boldsymbol{\mu}_t$  constant.

$\mathbf{H}_t$  is the  $N \times N$  matrix of conditional variances and covariances of the unpredictable component  $\boldsymbol{\epsilon}_t$  at time  $t$ . Finally,  $\boldsymbol{\eta}_t$  is the  $N \times 1$  vector of i.i.d errors such that  $E[\boldsymbol{\eta}_t]=0$  and  $E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t'] = \mathbf{I}$ .

In a general MGARCH model, we have to estimate the conditional covariance matrix  $\mathbf{H}_t$ , which, in addition, has to be positive definite for all  $t$ . Depending on the possible specifications for  $\mathbf{H}_t$ , we have different MGARCH models each belonging to one of the four categories mentioned before.

In order to estimate the conditional covariance matrix  $\mathbf{H}_t$ , a common issue using multivariate data is to resort to the maximization of an appropriate log-likelihood function  $L(\theta)$ , where  $\theta$  denotes the vector of all the parameters to estimate. Depending on the MGARCH model, the function  $L(\theta)$  assumes a different appearance. It is well-known that the quality of the Maximum Likelihood (ML) estimation relies also on the assumed data distribution. In general, when dealing with models with conditional heteroscedasticity, the estimates are known to be asymptotically normal [10]. In our approach, we consider a multivariate Gaussian distribution for the standardized error  $\boldsymbol{\eta}_t$ , even if, in principle, our method can be applied assuming different

distributions.

Given a stock market, consider  $\mathbf{H}_t$  the conditional variance and covariance matrix of the returns of the  $N$  assets at time  $t$ , and let  $\mathbf{R} = [\rho_{ij}]$ , with  $\rho_{ii} = 1, i = 1, \dots, N$ , be the *global* correlation matrix w.r.t. the sample period. The idea is that if asset  $i$  is highly positively/negatively correlated to asset  $j$ , then this kind of (high positive/negative) correlation does not change during the observed period. Note that we assume that the relationship based on the correlation between two assets  $i$  and  $j$  does not change during time while the value of the correlation coefficient  $\rho_{ij}$ , can (obviously) vary.

Let  $\delta > 0$  be a threshold, we construct the correlation graph  $G(\delta) = (V, E)$ , where  $V$  is the set of vertices each representing an asset, and  $E$  is the set of edges of  $G(\delta)$ , that is, the set of connections between pairs of assets. We assume that there is an edge  $(i, j)$  between assets  $i$  and  $j$  if and only if the (global) correlation coefficient  $|\rho_{ij}| > \delta$ . Let  $A(\delta)$  be the adjacency matrix of  $G(\delta)$ . Observe that the generic element  $a_{ij}(\delta)$  can be 0 or 1 if we assume an unweighted graph, otherwise  $a_{ij}(\delta) = w_{ij}$  if we assume that the graph  $G(\delta)$  is weighted with weights  $w_{ij}$  assigned to each edge  $(i, j)$  and obtained as a function of the correlation coefficients  $\rho_{ij}$ , that is,  $w_{ij} = f(\rho_{ij})$ . Given a general graph  $G = (V, E)$ , the following definitions hold [13, 17]:

**Definition 1** A subset  $C \subseteq V$  is called a *clique* of  $G$  if any two vertices in  $C$  are connected by an edge. The order  $q$  of a clique is the cardinality of  $C$ .

**Definition 2** A subset  $C \subseteq V$  is called a *maximal clique* of order  $q$  of  $G$  if  $C$  is not included in any other clique  $C'$  of order  $q + 1$ .

Other concepts not defined in the paper can be found in the book [13].

In a modified MGARCH model, the *modified* likelihood function to maximize is:

$$L(\theta) = \sum_{t=1}^T [\text{Dist}(\mathbf{A}(\delta), \mathbf{A}_t(\delta))], \quad (2)$$

where  $\text{Dist}(\cdot)$  is (any) distance or divergence measure between the adjacency matrix  $A(\delta)$  referred to the global correlation matrix  $\mathbf{R}$  and the adjacency matrix  $A_t(\delta)$  related to the conditional correlation matrix  $\mathbf{R}_t$ . Since  $A(\delta)$  is the correlation matrix computed w.r.t. the sample period, it can be considered as a target matrix.

In our framework, as a measure we can consider any distance or divergence measure between two  $N \times N$  positive definite matrices  $P$  and  $Q$ . In this paper, we consider the well-known Kullback-Leibler (KL) distance (or divergence) [29] between  $P$  and  $Q$ , namely  $KL(P, Q)$ . As

a statistical measure, it is a measure of the distance between two probability densities  $P$  and  $Q$ . In the case of Gaussian multivariate distributions this distance is completely defined by the correlation matrices of the whole system. Thus, it can be interpreted as how a multivariate probability distribution represented by the matrix  $Q$  is different from a reference multivariate probability distribution represented by the matrix  $P$ . It is also well-known that it is not a metric since this measure is not symmetric and does not satisfy the triangle inequality. Given matrices  $P$  and  $Q$ , the  $KL(P, Q)$  measure is:

$$KL(P, Q) = \frac{1}{2} \left[ \log\left(\frac{|Q|}{|P|}\right) + Tr(Q^{-1}P) - N \right], \quad (3)$$

where the operator  $|\cdot|$  computes the determinant of a matrix and  $Tr(\cdot)$  is the trace of a square matrix, which is the sum of its diagonal elements. In our approach  $Dist(A(\delta), A_t(\delta)) = KL(A(\delta), A_t(\delta))$ .

Note that assuming the identity function as the function of the graph's weights, i.e.  $w_{ij} = \rho_{ij}$ ,  $i \neq j$ , the matrix  $A(\delta)$  corresponds, in fact, to the unconditional correlation matrix where, given  $\delta$ ,  $a_{ij}(\delta) = \rho_{ij} \neq 0$ ,  $i \neq j$ , if and only if  $|\rho_{ij}| > \delta$ ,  $a_{ij}(\delta) = 0$ ,  $i \neq j$ , otherwise (i.e.,  $|\rho_{ij}| \leq \delta$ ), and further imposing  $a_{ij}(\delta) = 1$ ,  $i = 1, \dots, N$ . Denote by  $\widehat{Z}$  the matrix  $A(\delta)$  defined so far. Under the above hypothesis on the graph's weights, also  $A_t(\delta)$  can be considered a conditional correlation matrix at time  $t$ . Hence, in this special case, in order to preserve the strong relationships among specific groups of assets, in the maximization of the modified likelihood function (2) we are, in fact, requiring that  $a_{ij}(\delta)$  and  $a_{ij,t}(\delta)$  be as close as possible, that is, the corresponding correlation values  $|\rho_{ij}|$  and  $|\rho_{ij,t}|$  be as close as possible. In the rest of the paper we assume  $w_{ij} = \rho_{ij}$ ,  $i \neq j$ .

## 2.1 The BEKK model

The BEKK class of multivariate GARCH models was introduced in [26]. In order to make the model easier to solve, Ding and Engle [18] proposed a variant of the original one namely the *diagonal* BEKK model. The general model is:

$$\mathbf{H}_t = \mathbf{C} \cdot \mathbf{C}' + \sum_{k=1}^K \sum_{i=1}^p \mathbf{A}_{ki} \cdot (\epsilon_{t-i} \epsilon'_{t-i}) \cdot \mathbf{A}'_{ki} + \sum_{k=1}^K \sum_{j=1}^q \mathbf{B}_{kj} \cdot \mathbf{H}_{t-j} \cdot \mathbf{B}'_{kj}, \quad (4)$$

where  $A_{ki}$  and  $B_{kj}$  are parameter matrices,  $C$  is a lower triangular matrix,  $p$  and  $q$  represent the lagged error term and the number of conditional covariance lags, respectively.  $K$  determines the generality of the process. We assume that  $p = q = 1$  and  $K = 1$ , so that the diagonal BEKK model in a compact form can be written:

$$\mathbf{H}_t = \mathbf{C} \cdot \mathbf{C}' + \mathbf{A} \cdot (\epsilon_{t-1} \epsilon_{t-1}') \cdot \mathbf{A}' + \mathbf{B} \cdot \mathbf{H}_{t-1} \cdot \mathbf{B}', \quad (5)$$

with  $A$  and  $B$  diagonal matrices. Positive definiteness of conditional covariance matrices is guaranteed, by construction (see [18]). The procedure used in estimating the parameters of the model is the maximization of a likelihood function constructed under the assumption of an i.i.d. of the errors  $\eta_t$ . Under the further assumption of a conditional normality, the set of all the parameters  $\theta$  of the multivariate diagonal BEKK model can be estimated by maximizing the following sample log-likelihood function:

$$L(\theta) = -\frac{TN}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T (\log |H_t| + \epsilon_t' H_t^{-1} \epsilon_t), \quad (6)$$

with  $T$  the number of returns observations. Note that in the BEKK model equation (5) refers to the conditional covariance matrices. Thus, in the following modified log-likelihood function we have to use covariance matrices in place of the correlation matrices. Hence, under the assumption  $w_{ij} = \rho_{ij}$ ,  $i \neq j$ , as a special case for the graph's weights, given  $\delta$ , we first compute matrix  $\widehat{Z}$  and then the corresponding unconditional covariance matrix  $\widehat{\Sigma}$  w.r.t. the whole sample period  $T$ , that is,  $\widehat{\Sigma} = \Gamma \widehat{Z} \Gamma$ , with  $\Gamma$  the diagonal matrix of standard deviations w.r.t. the whole sample period. Finally, in the Kullback-Leibler divergence measure of formula (7) we compute the difference between matrices  $\widehat{\Sigma}$  and  $H_t$ . The modified log-likelihood function is:

$$\begin{aligned} L(\theta) &= -\frac{TN}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T (\log |H_t| + \epsilon_t' H_t^{-1} \epsilon_t) - KL(\widehat{\Sigma}, H_t) \\ &= -\frac{TN}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T (\log |H_t| + \epsilon_t' H_t^{-1} \epsilon_t) - \frac{1}{2} \sum_{t=1}^T (\log(\frac{|H_t|}{|\widehat{\Sigma}|}) + Tr(H_t^{-1} \widehat{\Sigma}) - N) \\ &= -\frac{1}{2} [N(T \log(2\pi)) + \sum_{t=1}^T (2 \log(|H_t|) + \epsilon_t' H_t^{-1} \epsilon_t - \log(|\widehat{\Sigma}|) + Tr(H_t^{-1} \widehat{\Sigma}))]. \end{aligned} \quad (7)$$

Observe that the minimization of the distance between the target unconditional covariance matrix  $\widehat{\Sigma}$  and  $H_t$  forces the values  $\sigma_{ij,t}$  to be as close as possible to the values  $\sigma_{ij}$ , that is, forces the values  $\rho_{ij,t}$ ,  $i \neq j$ , to be as close as possible to the values  $\rho_{ij}$ .

## 2.2 The DCC model

The idea of this model, introduced and analyzed in [27], is that the conditional covariance matrix  $\mathbf{H}_t$  can be decomposed into the conditional standard deviations  $\mathbf{D}_t$  of each of the  $N$  series and

a conditional correlation matrix of the returns  $\mathbf{R}_t$ . The dynamic of the model is described by the equations (1) and:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t. \quad (8)$$

The matrix  $\mathbf{D}_t$  is a diagonal matrix and consists of the  $N$  univariate GARCH models. Since it is a diagonal non negative matrix with all diagonal elements positive,  $\mathbf{D}_t$  is positive definite. To ensure  $\mathbf{H}_t$  to be positive definite, it is, therefore, necessary that the matrix  $\mathbf{R}_t$  be positive definite with the additional constraint that all its elements have to be equal to or less than 1 by definition. The dynamic of correlation matrix  $\mathbf{R}_t$  is, in turn, derived from another matrix  $\mathbf{Q}_t$  of the form:

$$\mathbf{R}_t = \bar{\mathbf{Q}}_t^{-1} \mathbf{Q}_t \bar{\mathbf{Q}}_t^{-1}, \quad (9)$$

where  $\bar{\mathbf{Q}}_t = \text{diag}(\mathbf{Q}_t)$  with  $\text{diag}(\cdot)$  the diagonal of a square matrix. The form of  $\mathbf{Q}_t$  determines the dynamic of the model and its complexity (see, e.g., [23, 48]). For example, following [23], the dynamics is:

$$\mathbf{Q}_t = (\mathbf{1} - \theta_1 - \theta_2) \hat{\mathbf{Q}}_t + \theta_1 (\epsilon_{t-1} \epsilon'_{t-1}) + \theta_2 \mathbf{Q}_{t-1}, \quad (10)$$

with  $\hat{\mathbf{Q}}_t = \text{Cov}[\epsilon_t \epsilon'_t] = E[\epsilon_t \epsilon'_t]$ . In order to ensure  $\mathbf{R}_t$  be positive definite, the parameters  $\theta_1$  and  $\theta_2$  must satisfy:

$$\theta_1 \geq 0, \theta_2 \geq 0, \theta_1 + \theta_2 < 1.$$

The parameter estimation phase is rather difficult, and hence for the DCC model a two stage estimation procedure is provided. In the first stage the parameters of the univariate GARCH models are estimated for each asset series. In the second stage, a second set of parameters are estimated given the parameters found in the previous phase. Referring to the dynamics described in (10) and assuming multivariate Gaussian distributed errors, after the first step only the parameters  $\theta_1$  and  $\theta_2$  are unknown so they are estimated in the second stage. In this second phase the log-likelihood function is:

$$L(\theta) = -\frac{1}{2} \sum_{t=1}^T (\log(|R_t|) + \epsilon'_t R_t^{-1} \epsilon_t). \quad (11)$$

In our approach we are interested in the second stage of the process where the log-likelihood function takes into account the correlation matrix of the assets returns at time  $t$ . Hence, under the assumption  $w_{ij} = \rho_{ij}$ ,  $i \neq j$ , given  $\delta$ , we consider the matrix  $\hat{\mathbf{Z}}$  and the matrix  $\mathbf{R}_t$ . The

corresponding modified log-likelihood function is:

$$\begin{aligned}
L(\theta) &= -\frac{1}{2} \sum_{t=1}^T (\log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t) - KL(\widehat{Z}, R_t) \\
&= -\frac{1}{2} \left[ \sum_{t=1}^T (2 \log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t - \log(|\widehat{Z}|) + Tr(R_t^{-1} \widehat{Z}) - N) \right].
\end{aligned} \tag{12}$$

As for the modified BEKK model, to highlight the strong relationships among groups of assets in the market, the minimization of the distance between the (target) matrix  $\widehat{Z}$  and  $R_t$  forces the values  $\rho_{ij,t}$ ,  $i \neq j$ , to be as close as possible to the values  $\rho_{ij}$ .

To conclude this section, we observe that if we set  $\delta = 0$  and under the assumption that  $w_{ij} = \rho_{ij}$ ,  $i \neq j$ , the correlation matrix  $\widehat{Z}$ , and consequently the unconditional covariance matrix  $\widehat{\Sigma}$ , correspond exactly to the long-run correlation and covariance matrices used as target matrices in [15]. Hence, on the one hand, our method can be precisely considered as an alternative targeting method w.r.t. the one proposed in [15]. On the other hand, assuming  $\delta > 0$ , with our method we are, in fact, “forcing” our modified MGARCH models in highlighting clusters of highly correlated assets in the simulated series.

### 3 Data analysis and test

In order to show the effectiveness of our new estimation approach, which aims at correctly identifying highly (positive or negative) correlated groups of assets, given a financial dataset we have to perform a first pre-processing phase by a method of clustering financial time series. In a broad context, cluster analysis is about assembling a collection of objects, and each of these objects is called a cluster, in which everything should be similar but still as different as possible from everything else in that cluster (see [40]). Time series clustering creates groupings based on similarities between time series without using pre-existing categorization labels [38]. Various approaches were proposed in time series clustering, and various authors attempted to synthesize the literature on the field (see, e.g., [1] [31] [34]).

According to [34], there is a wide range of applications for time series clustering that can be found in a variety of domains (e.g., economic or medical time series analysis, visual processing, anomalous activity detection). Different clusters of time series can be formed based on evolving patterns or real-world interactions (group membership can be interpreted as temporal context). In this case, the prediction may also differ from the results of cluster analysis [32].

Time series clustering is particularly relevant in finance. There is a particular interest in discovering typical dynamics of financial markets and the impact of different shocks on time series

and portfolio allocation groups. In this respect, according to [36] clustering of financial time series is essential to determine how much wealth should be allocated to financial assets and opportunities. Therefore, financial time series need to be clustered to select an appropriate portfolio and analyze an economic system. When studying highly complicated phenomena such as financial time series, one has to deal with substantial heterogeneity and peculiar characteristics and features (see [42] and [44]). Therefore, robustness measures should address this challenge [21]. Moreover, studying their behavior over time in a dynamic framework is also essential, as these systems are associated with uncertainty and for this reason, other approaches representing this uncertainty were proposed in [19].

In this section we provide a statistical framework for classifying time series and an example of applying the proposed method to a group of time series. Thus, our whole approach is based on two stages: In the first stage, we explicitly define the different groups of time series using a clustering procedure, which is helpful to identify the different partitions (groups). In the second stage we apply our new estimation method to correctly recognize the groups obtained so far. In our example, we perform the first stage on the following set of stocks: Facebook, Apple, Google, Boeing, Microsoft, Amazon, General Motors, Goldman Sachs, JPMorgan Chase, Intel, Verizon Communications, Visa, Cisco, Coca-Cola Company, and Salesforce. They were selected without any particular pattern from the stocks listed on the NYSE, the New York Stock Exchange. The period considered is from 2020-01-01 to 2021-01-01, and the resulting 253 observations refer to the daily closing prices of each financial time series.

Following [34], we have to clarify what is the claim in what two data are similar, and in what way we can obtain a good similarity and dissimilarity measure between two different observations (in this case, two different financial time series). These are the central questions of cluster analysis. In this regard, we have explicitly considered a hierarchical clustering algorithm that uses a Pearson correlation-based distance and the *complete* method (see [28]) as a helpful approach to distinguish the different groups that we can identify. Concerning financial time series, Pearson's correlation coefficient is widely used in the related financial literature to quantify the degree of similarity or dissimilarity between two time series [36]. If we denote  $y_t$  and  $x_t$  as two different time series, we can consider their correlations [34]:

$$\text{COR}(x_t, y_t) = 1 - \frac{\sum_{t=1}^T (y_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sqrt{\sum_{t=1}^T (x_t - \bar{x}_t)^2} \sqrt{\sum_{t=1}^T (y_t - \bar{y}_t)^2}}, \quad (13)$$

So we can compute the following distance, and we have:

$$d_{\text{Cor}}(x_t, y_t) = 1 - \text{COR}(x_t, y_t) \quad (14)$$

We apply the distance and cluster approach to develop the dendrogram and interpret the different clusters (see Figure 1). Then, we follow an exploratory approach to identify the groups of different stocks that are included in each cluster. We use the dendrogram to collect information about the variability of a cluster structure and the overall groups structure.

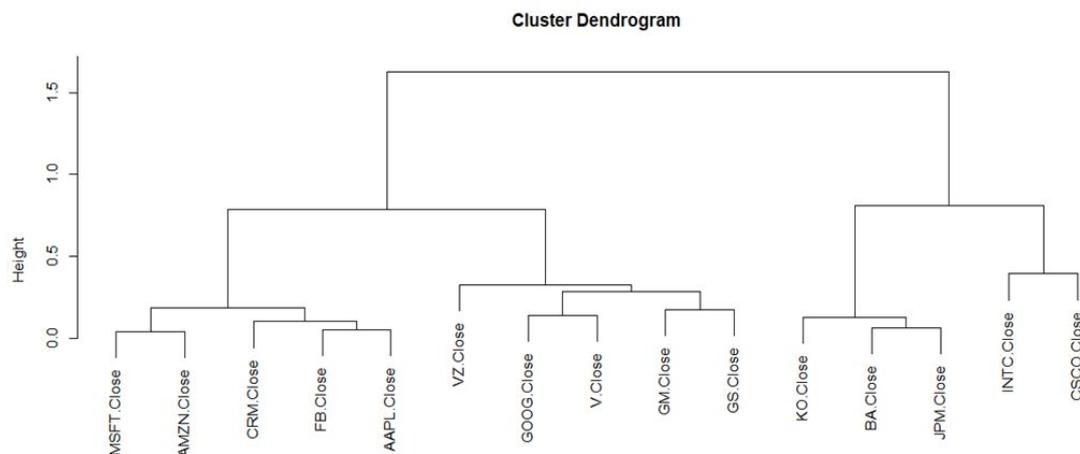


Figure 1: Clusters obtained in the first phase.

No prior knowledge is required for this approach, and it allows understanding patterns in financial time series data without the need for additional information [2]. This approach can be summarized as follows: use similarity measures to identify subgroups (see [49]); determine which generic grouping each individual belongs to. Through the above clustering procedure, we can identify four general clusters (see Figure 1). First, we can identify Microsoft's most significant joint movement (MSFT) and Amazon (AMZN). At the same time, Facebook (FB) and Apple (AAPL), and finally, Salesforce (CRM) belong to the same cluster. This cluster is clearly based on the technological innovations that can have an overarching impact on different sectors and companies (see also [19]). Finally, at the same time, Coca Cola (KO) and, more importantly, Boeing (BA) and JPMorgan Chase (JPM) are in a different cluster. Furthermore, Intel (INTC) and Cisco (CSCO) can also be considered part of the same cluster. We can further observe the critical role of economic similarities for these two stocks in this case. In this respect, it is possible to suppose relevant economic shocks that affect both financial time series. Finally, in the third cluster, we can observe a slightly different behavior for Verizon Communications (VZ), but also daily solid joint movements for Google (GOOG) and Visa (V), as well as for General Motors (GM) and Goldman Sachs (GS). Overall, through observation and visual analysis of the dendrogram, the distinct clusters obtained will be used as a reference for the evaluation of the effectiveness of our new estimation approach (second stage) described in the following.

### 3.1 Estimation phase

In this section we present some examples of different sizes pointing out the differences between the simulated series obtained by the BEKK model and its modified version, as well as, the DCC model and its modified version. All MGARCH models have been implemented in Matlab R2018 version using the MFE Toolbox code repository by Kevin Sheppard [43]. In the resulting correlation graphs we detected all the maximal cliques by the Bron-Kerbosch algorithm [14]. The experiments were carried out on a PC equipped with an Intel Core i7-3632MQ processor with 2.20Ghz.

#### Example 1

In this first example we consider only the group of technological innovation assets formed by MSFT, AMZN, CRM, FB and AAPL (see Figure 1). We compute the 252 log-return values and the corresponding unconditional correlation matrix, which is (see Table 1):

	<b>MSFT</b>	<b>AMZN</b>	<b>CRM</b>	<b>FB</b>	<b>AAPL</b>
<b>MSFT</b>	1	0.7712	0.7690	0.6855	0.6914
<b>AMZN</b>	0.7712	1	0.8435	0.7028	0.6458
<b>CRM</b>	0.7690	0.8435	1	0.7413	0.7508
<b>FB</b>	0.6855	0.7028	0.7413	1	0.6121
<b>AAPL</b>	0.6914	0.6458	0.7508	0.6121	1

Table 1: The correlation matrix for  $N = 5$  assets.

Let us now consider the (sub)graph  $G(\delta)$  of the complete correlation graph  $G$  whose vertex set contains only the subset of pairs of assets such that  $|\rho_{ij}| > \delta$ , with  $\delta = 0.5$  (see Figure 2). In the figure, the numbers associated with the vertices correspond to the following labeling: 1-MSFT, 2-AMZN, 3-CRM, 4-FB, 5-AAPL.

Note that the graph  $G(\delta)$  corresponds to the complete graph  $G$  since for all pairs of assets  $i$  and  $j$  we have  $\rho_{ij} > 0.5$ .

Consider the whole original dataset consisting of  $T = 252$  return observations, we first estimates the parameters of the BEKK and DCC models and then the parameters of each of the modified versions of the two models using the log-likelihood functions (7) and (12), respectively. Then, using these sets of parameters we simulate (new)  $T = 252$  log-return observations of the five assets. Our aim is to verify whether our modified models allow to better capture the strong relationships between assets in the stock market. In other words, we want to compare the graph  $G(\delta)$  with the simulated graphs  $G_S(\delta)$ .

Comparing the five graphs, only the original BEKK model was not able to correctly reproduce

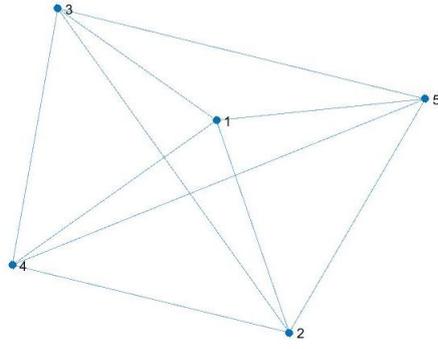


Figure 2: The graph  $G(\delta)$  with  $\delta = 0.5$ .

the original series. In fact, vertex 5, corresponding to asset AAPL, is not connected to all the other vertices despite its correlation value with other assets is greater than 0.5. In addition to the graphs  $G_S(\delta)$ , in Table 2 we report the Frobenius distance between the fitted covariance matrices  $H_t$  with respect to the unconditional covariance matrix  $\hat{\Sigma}$ , and the Kullback-Leibler divergence between the matrix  $\hat{\Sigma}$  and the unconditional covariance matrix of the new simulated series of log-returns, denoted by  $\hat{\Sigma}_S$ , w.r.t. all the considered models. The Frobenius norm is:

$$F = \sqrt{\sum_{t=1}^T \text{Tr}[(H_t - \hat{\Sigma})'(H_t - \hat{\Sigma})]}, \quad (15)$$

and the Kullback-Leibler divergence is:

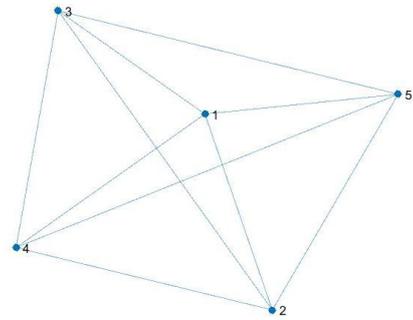
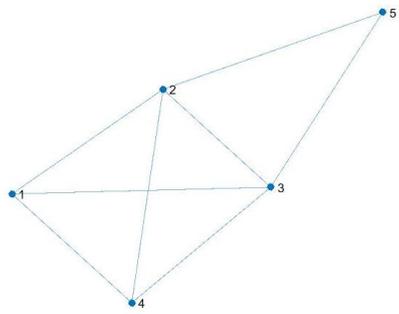
$$KL = \frac{1}{2} \left[ \log\left(\frac{|\hat{\Sigma}_S|}{|\hat{\Sigma}|}\right) + \text{Tr}(\hat{\Sigma}_S^{-1}\hat{\Sigma}) - N \right]. \quad (16)$$

	<b>BEKK</b>	<b>modified BEKK</b>	<b>DCC</b>	<b>modified DCC</b>
F	0.0669	0.0555	0.0881	0.0870
KL	0.0855	0.0229	0.0162	0.0154

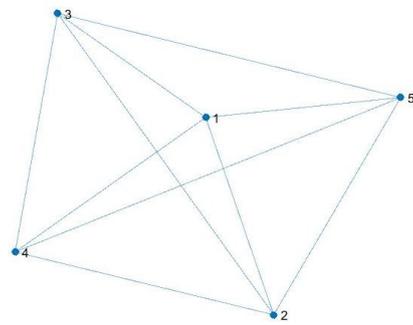
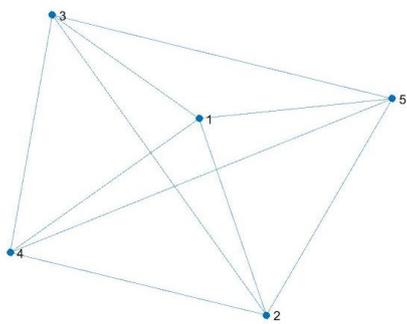
Table 2: Values of the Frobenius and the Kullback-Leibler loss functions: case  $N = 5$  and  $\delta = 0.5$

Both functions  $F$  and  $KL$  measure loss, so that lower values are preferable. We note that the values of the modified MGARCH models are better than the corresponding original models. Here we are not interested in comparing the values of the two loss functions among all the models since, again, it is not our goal to find a single winner.

Assume now we are interested in detecting groups of assets with a higher value of correlation.



(a) The graph  $G_S(\delta)$  resulting from the BEKK model      (b) The graph  $G_S(\delta)$  resulting from the modified BEKK model



(c) The graph  $G_S(\delta)$  resulting from the DCC model      (d) The graph  $G_S(\delta)$  resulting from the modified DCC model

Figure 3: The simulated graphs  $G_S(\delta)$  with  $\delta = 0.5$ .

On the basis of the values in Table 1, we set  $\delta = 0.71$ . The corresponding graph  $G(\delta)$  is:

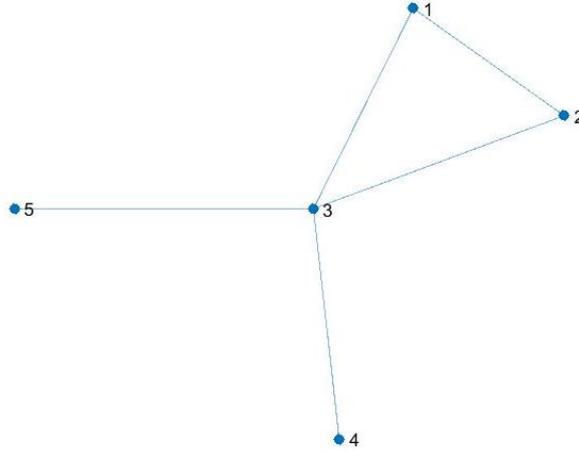


Figure 4: The graph  $G(\delta)$  with  $\delta = 0.71$ .

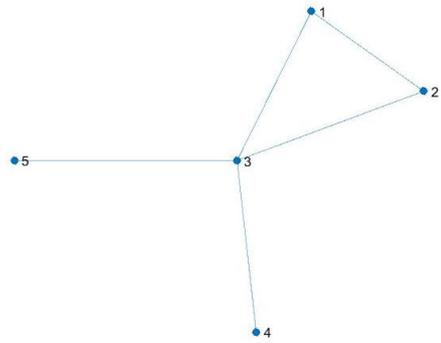
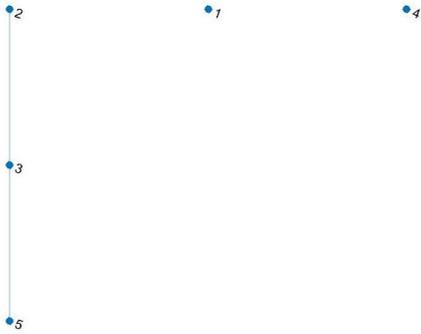
We observe that with this value of  $\delta$ , only assets  $\{1, 2, 3\}$  form a clique. The goal is to find simulated graphs  $G_S(\delta)$ , with  $\delta = 0.71$ , as similar as possible to the subgraph  $G(\delta)$  of Figure 4. The results are reported in the following Figure 5 along with the corresponding table (see Table 3) reporting the values of the two loss functions considered.

The simulated graph of the modified BEKK model exactly replicates  $G(\delta)$ . The two graphs associated to the DCC and modified DCC models correctly find the clique  $C = \{1, 2, 3\}$ , but the DCC model determines a new clique  $C' = \{1, 2, 3, 4\}$  that contains  $C$ , while the modified DCC model introduces the additional edge  $(4, 2)$ . Thus, these two models “overestimate” the correlation between assets pointing out correlations that do not really exist. The original BEKK model has the lowest ability to correctly identify groups of highly correlated assets. On the other hand, the values of the loss functions reported in Table 3 are in line with the values in Table 2.

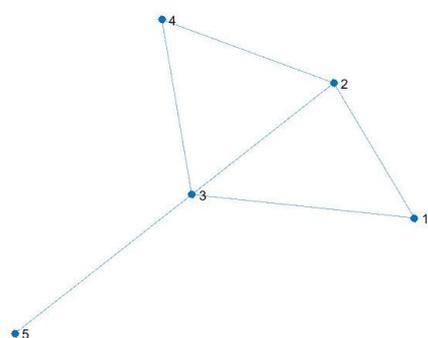
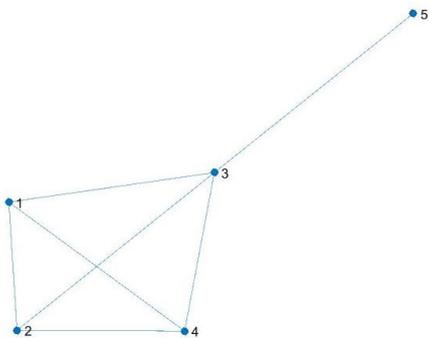
	<b>BEKK</b>	<b>modified BEKK</b>	<b>DCC</b>	<b>modified DCC</b>
F	0.0669	0.0575	0.0881	0.0840
KL	0.0976	0.0269	0.0306	0.0262

Table 3: Values of the Frobenius and the Kullback-Leibler loss functions: case  $N = 5$  and  $\delta = 0.71$

## Example 2



(a) The graph  $G_S(\delta)$  resulting from the BEKK model      (b) The graph  $G_S(\delta)$  resulting from the modified BEKK model



(c) The graph  $G_S(\delta)$  resulting from the DCC model      (d) The graph  $G_S(\delta)$  resulting from the modified DCC model

Figure 5: The simulated graphs  $G_S(\delta)$  with  $\delta = 0.71$ .

In this example we consider  $N = 8$  assets. The set of stocks is formed by the five previously considered assets and the group formed by KO, BA and JPM (see Figure 1). In the corresponding graphs the numbers associated with the vertices now correspond to the following labeling: 1-MSFT, 2-AMZN, 3-CRM, 4-FB, 5-AAPL, 6-KO, 7-BA, 8-JPM. We choose this new group of three stocks since the two clusters are adequately separated from each other, so that we expect finding two well separated groups of assets in the following correlation graphs. The unconditional correlation matrix is:

	MSFT	AMZN	CRM	FB	AAPL	KO	BA	JPM
MSFT	1	0.7712	0.7690	0.6855	0.6914	0.4038	0.4834	0.4387
AMZN	0.7712	1	0.8435	0.7028	0.6458	0.4810	0.5412	0.5580
CRM	0.7690	0.8435	1	0.7413	0.7508	0.4798	0.5912	0.5838
FB	0.6855	0.7028	0.7413	1	0.6121	0.2451	0.2697	0.3162
AAPL	0.6914	0.6458	0.7508	0.6121	1	0.3995	0.4562	0.4472
KO	0.4038	0.4810	0.4798	0.2451	0.3995	1	0.7211	0.6461
BA	0.4834	0.5412	0.5912	0.2697	0.4562	0.7211	1	0.7304
JPM	0.4387	0.5580	0.5838	0.3162	0.4472	0.6461	0.7304	1

Table 4: The correlation matrix for  $N = 8$  assets.

Consider  $\delta = 0.5$ , the graph  $G(\delta)$  is:

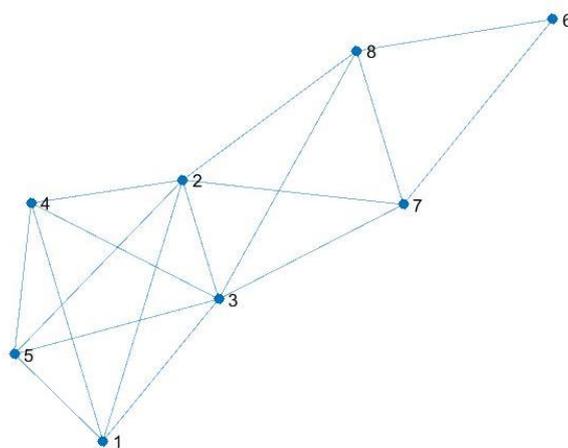
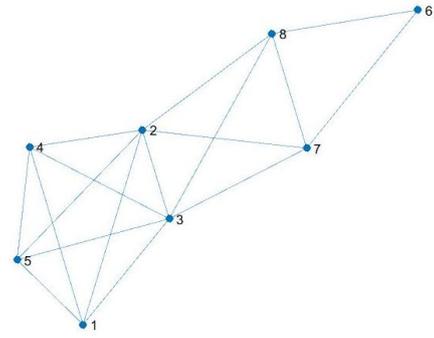
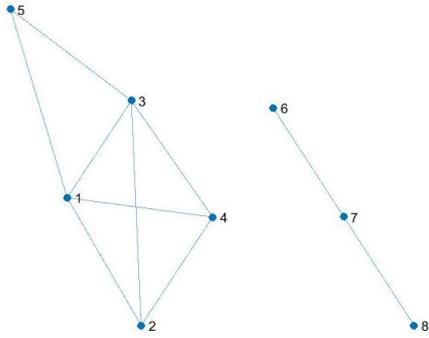


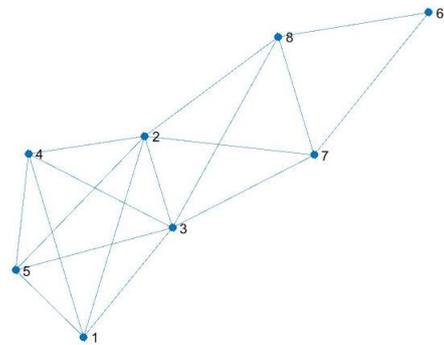
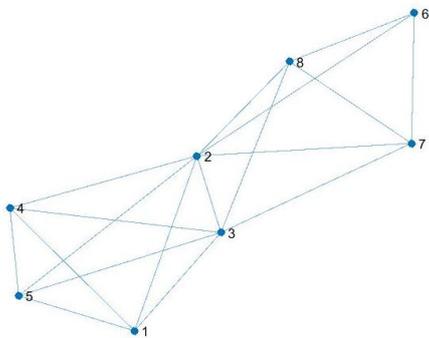
Figure 6: The graph  $G(\delta)$  with  $\delta = 0.5$  for  $N = 8$  assets.

In  $G(\delta)$  we clearly detect the three maximal cliques  $C_1 = \{1, 2, 3, 4, 5\}$ ,  $C_2 = \{6, 7, 8\}$  and  $C_3 = \{2, 3, 7, 8\}$ . In particular, cliques  $C_1$  and  $C_2$  refer to the two clusters of assets in Figure 1. The simulated graphs  $G_S(\delta)$  are in Figure 7:



(a) The graph  $G_S(\delta)$  resulting from the BEKK model

(b) The graph  $G_S(\delta)$  resulting from the modified BEKK model



(c) The graph  $G_S(\delta)$  resulting from the DCC model

(d) The graph  $G_S(\delta)$  resulting from the modified DCC model

Figure 7: The simulated graphs  $G_S(\delta)$  with  $\delta = 0.5$  for  $N = 8$  assets.

In the following table we report the values of the loss functions  $F$  and  $KL$  and the additional information on the number of maximal cliques in each of the simulated graphs  $G_S(\delta)$ .

	<b>BEKK</b>	<b>modified BEKK</b>	<b>DCC</b>	<b>modified DCC</b>
F	0.0669	0.0575	0.0881	0.0840
KL	0.0976	0.0269	0.0306	0.0262
Maximal Cliques	{1,2,3,4}, {1,3,5}, {6,7}, {7,8}	{1,2,3,4,5}, {2,3,7,8}, {6,7,8}	{1,2,3,4,5}, {2,3,7,8}, {2,6,7,8}	{1,2,3,4,5}, {2,3,7,8}, {6,7,8}

Table 5: Values of the Frobenius, the Kullback-Leibler loss functions and maximal cliques: case  $N = 8$  and  $\delta = 0.5$

The graphs  $G_S(\delta)$  related to the two modified MGARCH models are exactly the same as the graph  $G(\delta)$ . It is worth noticing that this does not mean that the correlation matrices of the simulated returns series with the modified BEKK and DCC models are equal to the correlation matrix of the original series, but that both the modified models are able to detect the same (original) clusters of highly correlated assets. Actually, assuming highly (positive) correlated assets remain unaltered over time, by forecasting future values of the series using the results provided by the modified BEKK and/or DCC models, one might expect to simulate returns series that do not alter too much the correlation structure among assets, as well as, the strong relationships among stocks. This can be useful in a portfolio selection problem from a diversification viewpoint. It is well-known that correlation represents the degree of relationship between the price movements of different assets included in the portfolio for diversified portfolios. Thus, choosing pairs of assets less correlated decreases the portfolio's overall risk. Consider, for example, the graph  $G_S(\delta)$  in Figure 7-c referred to the original DCC model, which is very similar to the graph  $G(\delta)$  (in fact there is just one additional edge in  $G_S(\delta)$ , that is, edge (2, 6)). It includes the maximal clique  $C = \{2, 6, 7, 8\}$  showing that asset 2 shows now a correlation value with asset 6 greater than 0.5, and precisely equal to 0.536, while from Table 4 we have  $\rho_{26} = 0.4810$ . Hence, the original DCC model inserts a strong correlation between these two assets, which is not truthful, and, without reason, this might prevent choosing assets 2 and 6 in an optimal portfolio. Finally, note that these considerations are difficult to obtain looking only at the values of the loss functions that show that the modified models behave better than the original ones. Finally, observe that the graph  $G_S(\delta)$ , related to the original BEKK model (see Figure 7-a), reports a very distorted relationship between the assets in the market. Hence, on the basis of the two above examples, we can state that it seems that the BEKK model benefits more effectively from the introduction of the modified likelihood function.

### Example 3

In this example we consider all the  $N = 15$  assets described in Section 3. In the graphs, the new labels associated to each asset are (see Figure 1): 1-MSFT, 2-AMZN, 3-CRM, 4-FB, 5-AAPL, 6-VZ, 7-GOOG, 8-V, 9-GM, 10-GS, 11-KO, 12-BA, 13-JPM, 14-INTC, 15-CSCO. The corresponding correlation matrix is in Table 6:

The correlation graph for  $\delta = 0.5$  is depicted in Figure 8.

The maximal cliques in  $G(\delta)$  with  $\delta = 0.5$  are:  $C_1 = \{1, 2, 3, 4, 5, 6, 14\}$ ,  $C_2 = \{1, 2, 3, 5, 6, 8, 9, 14, 15\}$ ,  $C_3 = \{2, 3, 6, 8, 9, 10, 12, 13, 14, 15\}$  and  $C_4 = \{6, 7, 8, 9, 11, 12, 13, 15\}$ . These cliques show the most interconnected groups of assets at a correlation level greater than 0.5. In the following we report the table about the values of the loss functions  $F$  and  $KL$  and the information on the maximal cliques in each of the simulated graphs  $G_S(\delta)$ . More precisely, in Table 7 we report in bold the cliques of the simulated graphs that coincide to the ones of the original graph  $G(\delta)$ . The modified BEKK model seems to be the best performing one both w.r.t the loss function values and the number of the same maximal cliques of  $G(\delta)$ . We also note that the first clique in the graph referring to the modified BEKK model has the same number of nodes of clique  $C_1$  but vertex 15 in place of vertex 14. We observe that the correlation values of the first six assets in Table 6 w.r.t. vertices 14 and 15 are very close to each other. Thus, the difference in those two cliques might be due to numerical reasons in the estimation and simulation phase of the modified BEKK model. In any case, all models introduce correlation values, and then a larger number of quite different cliques, that do not show up in the graph  $G(\delta)$ .

To evaluate the effectiveness of our method in finding groups of strongly correlated assets, which should be also distinct enough from each other, we consider a higher value for the threshold, that is,  $\delta = 0.65$ . The corresponding correlation graph  $G(\delta)$  is in Figure 9:

As expected, the graph  $G(\delta)$ , with  $\delta = 0.65$ , is sparser than the graph  $G(\delta)$ , with  $\delta = 0.5$ . In the new graph we identify a larger number of cliques, but of cardinality less than the cardinality of the maximal cliques when  $\delta = 0.5$ .

In Table 8 we report in bold the cliques of the simulated graphs that coincide to the ones in  $G(\delta)$ , with  $\delta = 0.65$ . It is evident that when  $N$  grows, the results are less clear-cut w.r.t. the two previous examples. On the one hand, this can be due that the log-likelihood functions to optimize are highly nonlinear functions and it is hard to find provably optimal solutions. Local solutions may be problematic and this creates difficulties in the estimation of the models. Additionally, since in the optimization phase matrices have to be inverted in each iteration, it makes the overall computation demanding unless  $N$  is small. In any case, Tables 7 and 8 still highlight that the modified MGARCH models performs slightly better than the original ones,

	<b>MSFT</b>	<b>AMZN</b>	<b>CRM</b>	<b>FB</b>	<b>AAPL</b>	<b>VZ</b>	<b>GOOG</b>	<b>V</b>	<b>GM</b>	<b>GS</b>	<b>KO</b>	<b>BA</b>	<b>JPM</b>	<b>INTC</b>	<b>CSCO</b>
<b>MSFT</b>	1	0.7712	0.7690	0.6855	0.6914	0.8067	0.4671	0.5619	0.6656	0.4566	0.4038	0.4834	0.4387	0.5493	0.5777
<b>AMZN</b>	0.7712	1	0.8435	0.7028	0.6458	0.7581	0.4539	0.6225	0.7123	0.5523	0.4810	0.5412	0.5580	0.6333	0.6596
<b>CRM</b>	0.7690	0.8435	1	0.7413	0.7508	0.8558	0.4849	0.6484	0.7830	0.6080	0.4798	0.5912	0.5838	0.7187	0.7171
<b>FB</b>	0.6855	0.7028	0.7413	1	0.6121	0.6837	0.2288	0.3825	0.4519	0.3955	0.2451	0.2697	0.3162	0.5065	0.4991
<b>AAPL</b>	0.6914	0.6458	0.7508	0.6121	1	0.6846	0.4191	0.5118	0.6353	0.4143	0.3995	0.4562	0.4472	0.5170	0.5358
<b>VZ</b>	0.8067	0.7581	0.8558	0.6837	0.6846	1	0.5379	0.6644	0.7903	0.5877	0.5163	0.6020	0.6172	0.6535	0.6838
<b>GOOG</b>	0.4671	0.4539	0.4849	0.2288	0.4191	0.5379	1	0.7488	0.7014	0.4135	0.6972	0.7590	0.5881	0.4913	0.5131
<b>V</b>	0.5619	0.6225	0.6484	0.3825	0.5118	0.6644	0.7488	1	0.7604	0.5899	0.6963	0.8907	0.7001	0.5981	0.6526
<b>GM</b>	0.6656	0.7123	0.7830	0.4519	0.6353	0.7903	0.7014	0.7604	1	0.6280	0.6541	0.7773	0.7363	0.6404	0.7295
<b>GS</b>	0.4566	0.5523	0.6080	0.3955	0.4143	0.5877	0.4135	0.5899	0.6280	1	0.4366	0.5975	0.7133	0.5698	0.6544
<b>KO</b>	0.4038	0.4810	0.4798	0.2451	0.3995	0.5163	0.6972	0.6963	0.6541	0.4366	1	0.7211	0.6461	0.4908	0.5061
<b>BA</b>	0.4834	0.5412	0.5912	0.2697	0.4562	0.6020	0.7590	0.8907	0.7773	0.5975	0.7211	1	0.7304	0.5663	0.6145
<b>JPM</b>	0.4387	0.5580	0.5838	0.3162	0.4472	0.6172	0.5881	0.7001	0.7363	0.7133	0.6461	0.7304	1	0.5542	0.5984
<b>INTC</b>	0.5493	0.6333	0.7187	0.5065	0.5170	0.6535	0.4913	0.5981	0.6404	0.5698	0.4908	0.5663	0.5542	1	0.6700
<b>CSCO</b>	0.5777	0.6596	0.7171	0.4991	0.5358	0.6838	0.5131	0.6526	0.7295	0.6544	0.5061	0.6145	0.5984	0.6700	1

Table 6: The correlation matrix for  $N = 15$  assets.

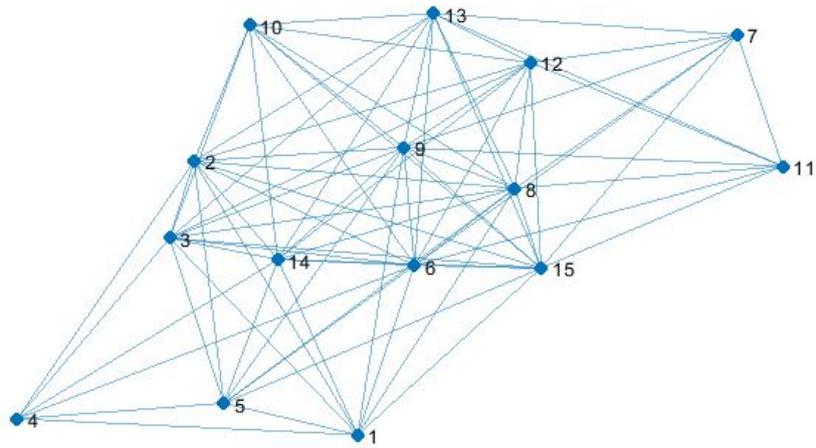


Figure 8: The graph  $G(\delta)$  with  $\delta = 0.5$  for  $N = 15$  assets.

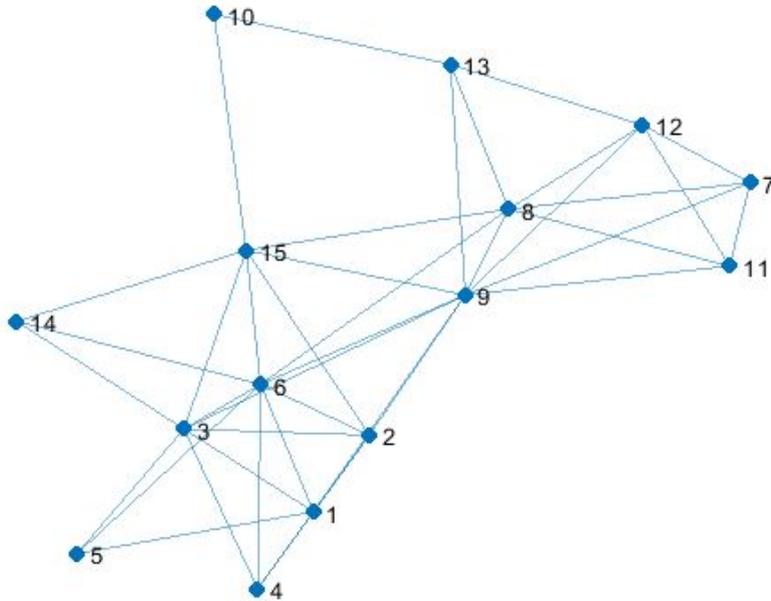


Figure 9: The graph  $G(\delta)$  with  $\delta = 0.65$  for  $N = 15$  assets.

	<b>BEKK</b>	<b>modified BEKK</b>	<b>DCC</b>	<b>modified DCC</b>
F	0.1708	0.1682	0.1995	0.2028
KL	0.4035	0.3002	0.3309	0.3091
Maximal Cliques	{1,2,3,4,5,6}, {7,8,9,11,12,13}, <b>{1,2,3,5,6,8,9,14,15}</b> , {1,2,3,6,8,9,10,14,15}, <b>{2,3,6,8,9,10,12,13,14,15}</b> , {7,8,9,12,13,14,15}	{1,2,3,4,5,6,15}, <b>{1,2,3,5,6,8,9,14,15}</b> , <b>{2,3,6,8,9,10,12,13,14,15}</b> , <b>{6,7,8,9,11,12,13,15}</b> , {6,8,9,11,12,13,14,15}	{1,2,3,4,5,6}, {1,2,3,5,6,8,9}, {2,3,6,8,9,10,15}, {3,6,8,9,10,13,15}, {3,8,9,10,12,13,15}, {3,8,9,10,12,14,15}, {7,8,9,11,12,13}, {2,8,9,11,14,15}, {8,9,11,12,13,15}, {8,9,11,12,14,15}	{1,2,3,4,6}, {1,3,4,5,6}, {1,3,5,6,9}, {1,2,3,6,8,9,10,13,14,15}, {1,3,6,8,9,10,12,13,14,15}, {1,6,7,8,9,12,13}, {7,8,9,11,12,13}

Table 7: Values of the Frobenius, the Kullback-Leibler loss functions and maximal cliques: case  $N = 15$  and  $\delta = 0.5$ .

	<b>BEKK</b>	<b>modified BEKK</b>	<b>DCC</b>	<b>modified DCC</b>
F	0.1708	0.1659	0.1995	0.2038
KL	0.3721	0.2982	0.2627	0.2433
Maximal Cliques	<b>{1,2,3,4,6}</b> , {1,2,3,5,6}, {2,3,5,6,9}, {2,3,6,9,14}, {3,6,9,14,15}, <b>{7,8,9,11,12}</b> , {8,9,11,12,13}, {6,8,9}, {9,10,13}	<b>{1,2,3,4,6}</b> , <b>{1,2,3,6,9}</b> , {1,2,3,5,6}, {2,3,6,9,14,15}, {3,6,8,9,14,15}, <b>{7,8,9,11,12}</b> , {8,9,11,12,13}, {8,9,12,14}, {9,10,12,13}, {9,10,15}	<b>{1,2,3,4,6}</b> , {2,3,6,9}, {3,5}, {3,14}, {7,8,9,12}, {7,8,11,12}, <b>{8,9,12,13}</b> , <b>{10,13}</b>	<b>{1,2,3,4,6}</b> , {1,2,3,5,6}, <b>{1,2,3,6,9}</b> , {7,8,9,12}, <b>{8,9,12,13}</b> , <b>{10,13}</b> , {7,8,11,12}, {3,15}

Table 8: Values of the Frobenius, the Kullback-Leibler loss functions and maximal cliques: case  $N = 15$  and  $\delta = 0.65$ .

in particular, when  $\delta$  increase. For example, for  $\delta = 0.65$  the simulated series obtained with the modified BEKK and DCC models are still able to better capture the strong relationships among assets in the market than the corresponding original models.

## 4 Discussion and conclusions

In this paper, we advance a method for improving the estimation phase of financial time series to improve the analysis and evaluation of large portfolios of financial assets whose performance strictly depends on the correlation among assets. Several different estimation models were proposed in the literature; among others, the family of models referred to as multivariate GARCH models are the most widely used. These models take into account that financial time series suffer from heteroscedasticity and that correlation matrices are time-varying. This paper considers two such models, namely the BEKK and DCC models. Furthermore, we modified the log-likelihood objective function based on the idea that in a financial market, there are often specific groups of assets that are highly correlated, and these relationships remain unaltered over time. Hence, in the likelihood function, we introduced a term referring to a loss measure computed on the difference between the time-varying covariance/correlation matrices and the covariance/correlation matrix estimated w.r.t. the whole in-sample period. Given the set of estimated parameters, we use them for simulating new time series to evaluate the effectiveness of the estimation phase. We also propose a new approach for the evaluation phase based on network analysis and, more precisely, on detecting maximal cliques in correlation graphs.

Results from an empirical case study are encouraging, particularly when the number of assets is not large, and we are likely able to find (global) optimal solutions in the maximization of the log-likelihood objective functions. In fact, in these cases, we observed a significant improvement in the ability of the modified MGARCH models to replicate the correlation structure of the assets in the market compared with the original models. When the number of assets increases, the estimation of the models involves somewhat heavy computations because they contain a large number of parameters, and there is no guarantee of finding provably global optimal solutions but only locally optimal solutions. This undermines the ability of the estimated models to replicate the correlation structure of the assets correctly. Despite this, even when the number of assets is large, the modified MGARCH models seem to perform better than the original ones. Overall, our experiments show that the BEKK model is the one that benefits most from the modification of the log-likelihood function.

To conclude, we cannot state whether the modified models always prevail over the original ones, and much more experiments are needed. This, therefore, leaves plenty of room for further

research. On the one hand, one can experiment with other loss functions introduced in the log-likelihood objective functions to improve the estimates. In fact, we observe that our approach is extremely flexible. Different loss functions can be considered without imposing additional constraints on covariance or correlation matrices. On the other hand, much attention should be paid in order to improve the optimization phase. This involves developing ad hoc strategies for finding global optimal solutions or solutions close to the optimal ones. In addition, the development of Metaheuristic procedures could be a further line of research to improve the optimization phase. Finally, finding new network indexes that better highlight other peculiar aspects of a financial market is worth investigating.

## **Funding and Competing interests**

The research of the second author was partially supported by the research program supported by Junta de Andalucía, Consejería de Economía y Conocimiento n. US-1256951. The authors have no financial or proprietary interests in any material discussed in this article.

## **References**

- [1] Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering—a decade review. *Information Systems*, 53, 16-38.
- [2] Ahn, J., & Lee, J. H. (2018, July). Clustering Method for Financial Time Series with Co-Movement Relationship. In 2018 3rd International Conference on Computational Intelligence and Applications (ICCIA) (pp. 260-264). IEEE.
- [3] Aielli, G.P. (2009). Dynamic conditional correlations: on properties and estimation. Available at SSRN: <http://ssrn.com/abstract=1507743>
- [4] Bautin, G. A., Kalyagin, V. A., Koldanov, A. P., Koldanov, P. A., & Pardalos, P. M. (2013). Simple measure of similarity for the market graph construction. *Computational Management Science*, 10, 105-124.
- [5] Bauwens, L., Laurent, S., & Rombouts, J. V. (2006). Multivariate GARCH models: a survey. *Journal of applied econometrics*, 21, 79-109.
- [6] Bhatia, R., Jain, T., & Lim, Y. (2019). On the Bures-Wasserstein distance between positive definite matrices. *Expositiones Mathematicae*, 37, 165-191.

- [7] Billio, M., & Caporin, M. (2010). Market linkages, variance spillovers, and correlation stability: Empirical evidence of financial contagion. *Computational statistics & data analysis*, 54, 2443-2458.
- [8] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- [9] Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72, 498-505.
- [10] Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11, 143-172.
- [11] Bollerslev, T., R. F. Engle, & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, 96, 116-131.
- [12] Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2018). Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. *Journal of Econometrics*, 207, 71-91.
- [13] Bondy, A., Murty, U.S.R. (2008). *Graph Theory*, Springer-Verlag, New York.
- [14] Bron, C., & Kerbosch, J. (1973). Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM*, 16, 575-577.
- [15] Caporin, M., & McAleer, M. (2012). Do we really need both BEKK and DCC? A tale of two multivariate GARCH models. *Journal of Economic Surveys*, 26, 736-751.
- [16] Chang, C. L., Liu, C. P., & McAleer, M. (2019). Volatility spillovers for spot, futures, and ETF prices in agriculture and energy. *Energy Economics*, 81, 779-792.
- [17] Christofides, N. (1975). *Graph theory: An algorithmic approach*. Academic Press, London.
- [18] Ding, Z. & Engle, R. (2001). Large scale conditional covariance modelling, estimation and testing. *Academia Economic Papers*, 29, 157-184.
- [19] Drago C., Lauro C., & Scepti G. (2013) Visualization and Analysis of Large Datasets by Beanplot PCA. In book: *Advances in Latent Variables* Publisher: Vita e Pensiero Editors: Brentari E., Carpita M.

- [20] Drago, C. (2017) Measuring the Interest and the Attraction for the Heritage Over Time Using Social Big Data: The Case of Florence. Available at SSRN: <https://ssrn.com/abstract=3093991> or <http://dx.doi.org/10.2139/ssrn.3093991>
- [21] D’Urso, P., De Giovanni, L., & Massari, R. (2016). GARCH-based robust clustering of time series. *Fuzzy Sets and Systems*, 305, 1-28.
- [22] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50, 987-1008.
- [23] Engle, R.F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20, 339-350 .
- [24] Engle, R., & Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24, 238-253.
- [25] Engle, R., & Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30, 212-228.
- [26] Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11, 122-150.
- [27] Engle, R. F., & Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Papers, No. 8554.
- [28] Everitt B.S., Landau S., & Leese, M. (2001). *Cluster Analysis*. London: Arnold. ISBN 0-340-76119-9.
- [29] Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22, 79-86.
- [30] Laurent, S., Rombouts, J. V., & Violante, F. (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics*, 173, 1-10.
- [31] Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern recognition*, 38(11), 1857-1874.
- [32] Maharaj, E. A., & Inder, B. A. (1999). Forecasting time series from clusters. Working Paper Monash University

- [33] Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11, 193-197.
- [34] Montero, P., & Vilar, J. A. (2014). TSclust: An R Package for Time Series Clustering. *Journal of Statistical Software*, 62(1), 1-43. <https://doi.org/10.18637/jss.v062.i01>
- [35] Patton, A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246-256.
- [36] Piccardi, C., Calatroni, L., & Bertoni, F. (2011). Clustering financial time series by network community analysis. *International Journal of Modern Physics C*, 22(01), 35-50.
- [37] Orskaug, E. (2009). Multivariate DCC-GARCH model with various error distributions, Norwegian Computing Center. Master's thesis, Institutt for matematiske fag.
- [38] Radhimeenakshi, D.S., & Latha, K. (2017). Similarity Measure Selection for Clustering Stock Market Time Series Databases.
- [39] Rombouts, J., Stentoft, L., & Violante, F. (2014). The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options. *International Journal of Forecasting*, 30, 78-98.
- [40] Sarda-Espinosa, A. (2017). Comparing time-series clustering algorithms in r using the dtwclust package. *R package vignette*, 12, 41.
- [41] Serra, A. P., & Zarate, L. E. (2015). Characterization of time series for analyzing of the evolution of time series clusters. *Expert systems with applications*, 42(1), 596-611.
- [42] Sewell, M. (2011). Characterization of financial time series. *Rn*, 11(01), 01. Working Paper
- [43] Sheppard, K. (2013). MFE toolbox. Available online: [https://www.kevinshppard.com/MFE\\_Toolbox](https://www.kevinshppard.com/MFE_Toolbox) (accessed on 16 August 2018).
- [44] Shi, Y., Li, B., Du, G., & Dai, W. (2021). Clustering framework based on multi-scale analysis of intraday financial time series. *Physica A: Statistical Mechanics and its Applications*, 567, 125728.
- [45] Silvennoinen, A., & Terasvirta, T. (2008). Multivariate GARCH models, in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds. *Handbook of Financial Time Series*, New York, Springer.

- [46] Silvennoinen, A., & Terasvirta, T. (2021). Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model. To appear in *Econometrics and Statistics*, <https://doi.org/10.1016/j.ecosta.2021.07.008>.
- [47] Stoyanov, S.V., Rachev, S.T., Racheva-Yotova, B., & Fabozzi, F.J. (2011). Fat-tailed models for risk estimation. *The Journal of Portfolio Management*, 37, 107-117.
- [48] Tse, Y. K., & Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20, 351-362.
- [49] Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data mining and knowledge Discovery*, 13(3), 335-364.