



Research Paper

## **Risk diversification: a study of persistence with a filtered correlation-network approach**

**Nicoló Musmeci<sup>1</sup>, Tomaso Aste<sup>2,3</sup> and T. Di Matteo<sup>1</sup>**

<sup>1</sup>Department of Mathematics, King's College London, Strand, London WC2R 2LS, UK; emails: nicolo.musmeci@kcl.ac.uk, tiziana.di\_matteo@kcl.ac.uk

<sup>2</sup>Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK; email: t.aste@ucl.ac.uk

<sup>3</sup>Systemic Risk Centre, London School of Economics and Political Sciences, Houghton Street, London WC2A 2AE, UK

(Received October 20, 2014; revised December 17, 2014; accepted January 9, 2015)

### **ABSTRACT**

In this paper, we study the evolution over time of the correlation structure of equity returns by means of a filtered-network approach and use this to investigate persistency and recurrences and their implications for risk-diversification strategies. We build dynamically planar maximally filtered graphs from the correlation structure over a rolling window and study the persistence of the associated directed bubble hierarchical tree (DBHT) clustering structure. We observe that the DBHT clustering structure is quite stable during the early 2000s, becoming gradually less persistent before the unfolding of the 2007–8 crisis. The correlation structure eventually recovers persistence in the aftermath of the crisis, setting up a new phase, distinct from the precrisis structure, in which the market structure is less related to industrial sector activity. We observe that the correlation structure is again losing persistence, which indicates the building up of another, different phase. Such dynamical changes in persistency and its occurrence at the unfolding of financial crises raise concerns about the effectiveness of correlation-based portfolio management tools for risk diversification.

---

Corresponding author: T. Di Matteo

Print ISSN 2055-7795 | Online ISSN 2055-7809  
Copyright © 2015 Incisive Risk Information (IP) Limited

**Keywords:** econophysics; correlation-based filtered networks; information filtering networks; systemic risk; risk diversification; community tracking.

## 1 INTRODUCTION

One way to reduce financial risk is by diversifying investments. This involves taking positions in assets that are historically anticorrelated or uncorrelated and in this way reducing the probability that all assets lose value at the same time. This is, for instance, the basis of the capital asset pricing model (Fama and French 2004). However, the applicability of these approaches relies on the implicit assumption that the relevant features of the correlation structure observed in the past have persistent significance into the future. This is not always the case.

In order to characterize the correlation structure and quantify its persistence, we use a network-filtering approach in which the correlation matrix is mapped into a sparse graph that retains only the relevant elements. To this purpose, we use the correlation-filtered networks known as planar maximally filtered graphs (PMFGs) (Tumminello *et al* 2005) and their associated clustering structure, the directed bubble hierarchical tree (DBHT) (Song *et al* 2012). A PMFG is a maximal planar graph that retains only the largest correlations. The DBHT is a hierarchical clustering that is constructed by making use of the separating properties of 3-cliques in planar graphs (Song *et al* 2011).

Since the seminal work of Mantegna (1999), network analysis on asset correlation has provided interesting insights into risk management and portfolio optimization. It has been observed that the structure of such networks not only is significantly related to the industrial sectors' classifications, but also conveys important independent information (Mantegna 1999; Musmeci *et al* 2014). It was shown in Borghesi *et al* (2007) that this network structure can be very robust against changes in the time horizon at which the asset returns are sampled (when the market mode dynamics is removed from the original correlations). This has been interpreted as an indication that "correlations on short time scales might be used as a proxy for correlations on longer time horizons" (Borghesi *et al* 2007). This, however, requires some degree of stationarity in the correlation structure.

Network-filtering procedures have been found to sensibly improve the performance of portfolio optimization methods. For instance, it was shown in Tola *et al* (2008) that Markowitz optimization gives better results on network-filtered correlation matrixes than on unfiltered ones. In Pozzi *et al* (2013), it was reported that the peripheral position of nodes in PMFGs can be a criterion for selecting a well-diversified portfolio. This finding is consistent with that for the minimum spanning tree (MST) in Onnela

*et al* (2003b): the stocks selected using the Markowitz method tend to be the “leaves” of the MST.

Network-filtered correlations carry both local and global information in their structures and the analysis of their temporal evolution may allow us to better understand financial market evolution. For instance, Di Matteo *et al* (2010) observed that stocks belonging to the same industrial sector tend to have similar values of centrality in the network topology, and that this differentiation is persistent over time. In particular, they observed that finance, basic materials and capital goods industrial sectors (Forbes classification) tend to be located mostly in the central region of the network, whereas energy, utilities and health care are located more in the peripheral region. The preeminent role of the financial sector is even stronger when correlation networks based on partial correlations are analyzed (Kenett *et al* 2010). Despite this overall robustness, a certain degree of nonstationarity has also been observed. For instance, the financial sector appears to have lost centrality over the first decade of the 2000s (Aste *et al* 2010). Buccheri *et al* (2013) found both a slow and a fast dynamics in correlation networks topology: while the slow dynamics shows persistence over periods of at least five years, the time scale of the fast dynamics is a few months and linked to special exogenous and endogenous events, such as financial crises. For instance, Onnela *et al* (2003a) showed that sharp structural changes occurred in the graph topology during Black Monday 1987. Similar phenomena have been observed for correlations on foreign exchange (FX) data (Jang *et al* 2010). McDonald *et al* (2008) demonstrated that structural changes on FX correlation data display different features depending on the type of event affecting the market. News that concerns economic matters can trigger a prompt destabilizing reaction, whereas there are periods of “collective discovery” in which dynamics appears to synchronize (McDonald *et al* 2008).

In this paper, we investigate the nonstationarity of correlation, quantifying how much, and in what way, the correlation structure changes over time. This is a particularly relevant topic because most portfolio optimization tools rely on some stationarity, or at least persistence, in the joint distribution of asset returns. It is generally accepted in the literature that financial correlations are nonstationary. For instance, in Livan *et al* (2012) it was shown, by means of a local Kolmogorov–Smirnov test on correlation pairs, how nonstationarity can sensitively affect the effectiveness of portfolio optimization tools. In this paper, we discuss the degree of nonstationarity in the correlations at a nonlocal level by using PMFG networks and the associated DBHT clustering and looking at changes in the hierarchical and clustering structures. In this context, persistence translates into a measure of similarity among communities in a network, for which network-theoretic tools should be used. The PMFG–DBHT method has recently been applied to the study of financial data (Musmeci *et al* 2014), showing that it is a powerful clustering tool that can outperform other traditional clustering methods, such as linkage and  $k$ -medoids, in retrieving economic information.

Moreover, the dynamical analyses have shown that the clustering structure reveals peculiar patterns over the financial crisis, for instance, the increasingly dominant role of the market mode over the period 1997–2012. This implies an increase in nondiversifiable risk in the market. In this paper, we take these analyses a step further by looking at the dynamics of this clustering and its persistence.

The rest of the paper is organized as follows. In Section 2, we summarize the main theoretical concepts underlying the correlation network tools. In Section 3, we describe the analyses we have performed and discuss the results. In Section 4, we draw our conclusions and discuss future perspectives.

## 2 CORRELATION-BASED NETWORKS: AN OVERVIEW

Over the last fifteen years, correlation-based networks have been used extensively in the econophysics literature as tools to filter and analyze financial market data (Aste *et al* 2005; Bartolozzi *et al* 2007; Di Matteo and Aste 2002; Di Matteo *et al* 2004, 2005; Mantegna 1999; Onnela *et al* 2003c; Tumminello *et al* 2005).

The seminal work of Mantegna (1999) exploited for the first time a tool from network theory, the MST (see, for example, West 1996), to analyze and filter from noise the correlation structure of a set of financial assets. Mantegna's idea was to look at a correlation matrix as the adjacency matrix of a network and generate an MST on this network in order to retain the most significant links/entries. Moreover, after mapping the correlation into a suitable metric distance, the MST algorithm provides a hierarchical classification of the stocks.

In the following years, other correlation-based networks were studied in the literature. Onnela *et al* (2003c) introduced the dynamic asset graph. Unlike the MST, which filters the correlation matrix according to a topological constraint (the tree-like structure of the MST), the dynamic asset graph retains all the links, such that the associated correlation (distance) is above (below) a given threshold. In this way, it is less affected by the insignificant, low correlations that are often kept by the MST. As a result, the dynamic asset graph is more robust against time (Onnela *et al* 2003c). On the other hand, the MST, by retaining both high and low correlations, is better equipped to uncover global, multiscale structures of interaction. Indeed, in financial and complex systems in general, several length scales coexist, and thresholding at a given value artificially introduces a characteristic size that might hide effects occurring at other scales.

The tree structure exploited in the MST tool is not the only topological constraint that can be used to filter information. In particular, if we replace the request of absence of loops with the planarity condition, we obtain the PMFG (Aste *et al* 2005). The PMFG can be seen as a generalization of the MST that is able to retain a greater amount of information (Aste 2012; Tumminello *et al* 2005), as it has a less strict

topology constraint that allows it to keep a larger number of links. It can be shown that the hierarchical properties of the MST are preserved in the PMFG.

We can take this concept a step further and generalize the PMFG to a broader class of networks by means of the concept of “genus” (Aste *et al* 2005). The genus,  $g$ , of a surface is the largest number of nonintersecting simple closed cuts that can be made on the surface without disconnecting a portion (equal to the number of handles in the surface). Requiring a network to be planar, as for the PMFG, is equivalent to requiring that the network be embedded on a surface with  $g = 0$  (ie, no handles, a topological sphere). The natural generalization of the PMFG is therefore a network embedded on surfaces with genus greater than zero. The greater the genus, the more handles are in the surface and the more links we can retain from the original correlation matrix. More links retained means more information and network complexity, but it also means more noise. When  $g = \lceil (N - 3)(N - 4)/12 \rceil$  (where  $N$  is the number of nodes and  $\lceil x \rceil$  is the ceiling function that returns the smallest integer greater than or equal to  $x$ ), the original, fully connected, complete graph associated with the correlation matrix can be recovered. The concept of embedding on surfaces therefore provides a quantitative way of tuning the degree of information filtering by means of a single parameter,  $g$ , linking correlation-based networks to algebraic geometry (Aste *et al* 2012).

Correlation-filtered networks are associated with clustering methods. Indeed, the MST is strictly related (Tumminello *et al* 2010) to a hierarchical clustering algorithm, namely the single linkage (SL) (Anderberg 1973). MST can indeed be seen as a network representation of the hierarchy generated by the SL. Recently, it has been shown that a hierarchical clustering can be derived from the PMFG as well (Aste 2014; Song *et al* 2012). This new method is the DBHT. However, the approach is different from the agglomerative one adopted in the linkage methods: the idea of the DBHT is to use the hierarchy hidden in the topology of a PMFG, due to its being made of 3-cliques (Song *et al* 2011, 2012). The DBHT hierarchical clustering was applied to synthetic and biological data in Song *et al* (2012) and financial data in Musmeci *et al* (2014), showing that it can outperform several other clustering methods, including  $k$ -means++ (Arthur and Vassilvitskii 2007),  $k$ -medoids (Kaufman and Rousseeuw 1987), linkage, spectral clustering via normalized cuts on  $k$ -nearest neighbor graphs ( $k$ NN-spectral) (Shi and Malik 2000), the self-organizing map (SOM) (Kohonen *et al* 2001) and the Qcut (Ruan *et al* 2010).

Since the DBHT exploits the topology of the correlation network, it can be viewed as an example of community-detection algorithms in graphs (Fortunato 2010). The implicit assumption underlying these algorithms is that a community is somehow related to the density of edges inside and outside the community itself, unlike strict data-clustering methods (such as the aforementioned linkage algorithms), which only

use the information contained in the similarity/distance matrix. Several different community detection algorithms in graphs have been suggested in the literature; many of them search for the community partition that maximizes modularity, a function that compares the density of links in each community with the one expected in a (null) random graph model (Guimerà *et al* 2004; Newman 2004; Newman and Girvan 2004). Other approaches include spectral analysis on the adjacency matrix or related matrixes, eg, Laplacian (Donetti and Muñoz 2004; Mitrović and Tadić 2009), random walks on networks (Hu *et al* 2008; Zhou 2003) and methods based on statistical inference (Reichardt and White 2007; Rosvall and Bergstrom 2007).

However, in this paper we focus on the DBHT, as it is tailored to planar graphs and is therefore the natural tool to use with PMFGs.

### 3 PERSISTENCE AND TRANSITIONS: DYNAMICAL ANALYSIS OF THE DIRECTED BUBBLE HIERARCHICAL TREE

We studied the dynamical evolution of DBHT clustering on a system of  $N = 342$  US stocks during the time period January 1997–December 2012. We selected a set of  $n = 100$  overlapping time windows,  $T_k$ , with  $k = 1, \dots, n$  (each one of length  $L = 1000$  trading days with a shift of 30 trading days between adjacent time windows), and computed the distance matrix

$$D_{ij}(T_k) = \sqrt{2(1 - \rho_{ij}(T_k))},$$

where  $\rho_{ij}$  is the Pearson correlation coefficient

$$\rho_{ij}(T_k) = \frac{\langle c_i(t)c_j(t) \rangle_{T_k}}{\sqrt{[\langle c_i^2(t) \rangle_{T_k} - \langle c_i(t) \rangle_{T_k}^2][\langle c_j^2(t) \rangle_{T_k} - \langle c_j(t) \rangle_{T_k}^2]}}, \quad (3.1)$$

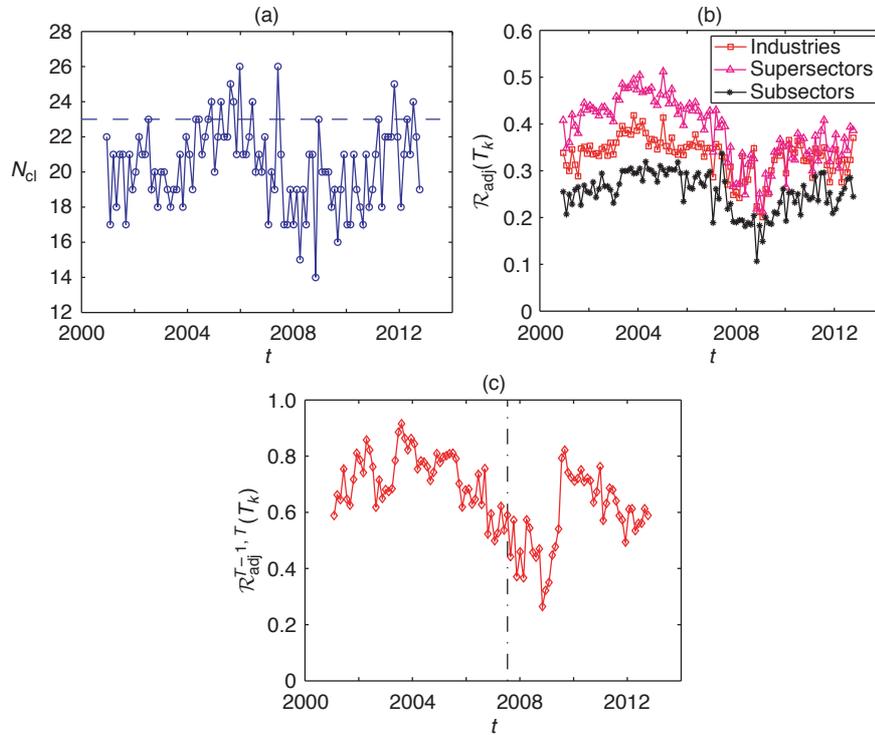
where  $\langle \cdot \rangle_{T_k}$  represents the average over the time window  $T_k$ , and  $c_i(t)$ ,  $c_j(t)$  are the daily log returns of stocks  $i$  and  $j$  detrended of the average market return factor. Following Borghesi *et al* (2007), we computed  $c_i(t)$  for each stock  $i$ , assuming the following one-factor model for the stock log return  $r_i(t)$ :

$$r_i(t) = \alpha_i + \beta_i I(t) + c_i(t), \quad (3.2)$$

where the common market factor  $I(t)$  is the market average return,

$$I(t) = \sum_{\gamma=1}^N r_{\gamma}(t).$$

The coefficients  $\alpha_i$ ,  $\beta_i$  are computed by means of a linear regression and  $c_i(t)$  is the residual. In agreement with Borghesi *et al* (2007), we verified that correlations

**FIGURE 1** Dynamical evolution of the DBHT clustering.

Each plot refers to 100 moving time windows ( $T_k$ ) of length 1000 trading days and shift 30 days. (a) Number of DBHT clusters,  $N_{cl}$ , with the dashed horizontal line representing the  $N_{cl}$  value obtained by taking the entire time window of 4026 trading days (covering years 1997–2012). Overall, we can observe a drop in correspondence with the 2007–2008 financial crisis. (b) Clustering similarity with industrial classification benchmark (ICB) classifications. This graph shows the amount of economic information retrieved by DBHT clustering in terms of similarity between clustering and ICB partitioning, calculated using the adjusted Rand index,  $\mathcal{R}_{adj}$ . Again, a drop at the outbreak of the crisis appears. Over the postcrisis years, there is less economic information than in the precrisis period, and differences among different ICB levels are less evident. (c) Similarity between consecutive clustering, showing the persistence of the DBHT clustering over time, measured as the adjusted Rand index between two adjacent clusterings. The financial crisis is characterized by very low levels of persistence.

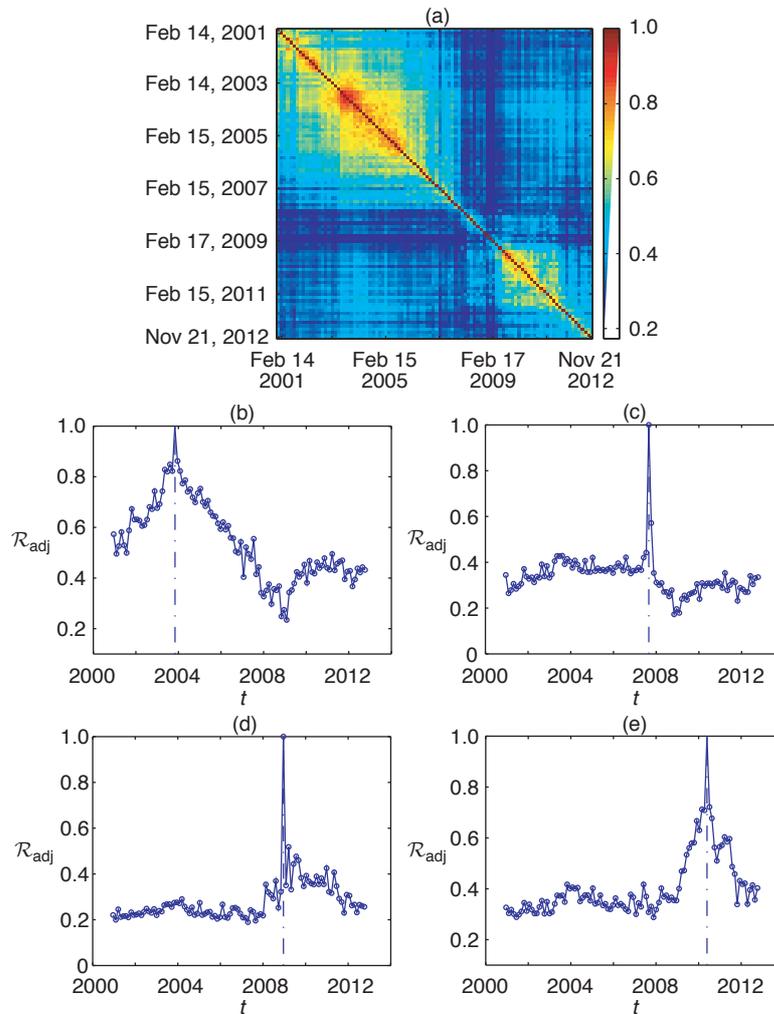
on detrended log returns provide a richer and more robust clustering that can carry information not evident in the original correlation matrix (Borghesi *et al* 2007). We also used a weighted version of the Pearson estimator (Pozzi *et al* 2012) in order to mitigate (exponentially) excessive sensitiveness to outliers in remote observations. The DBHT clustering is calculated on each distance matrix  $D(T_k)$ .

In part (a) of Figure 1, we show the number of DBHT clusters obtained for each time window. The number of clusters ranges from 14 to 26. The dashed line is the value

(23) corresponding to the clustering obtained using the entire period 1997–2012 as our time window. As we can observe, the lowest values are associated with the period around the 2007–8 financial crisis.

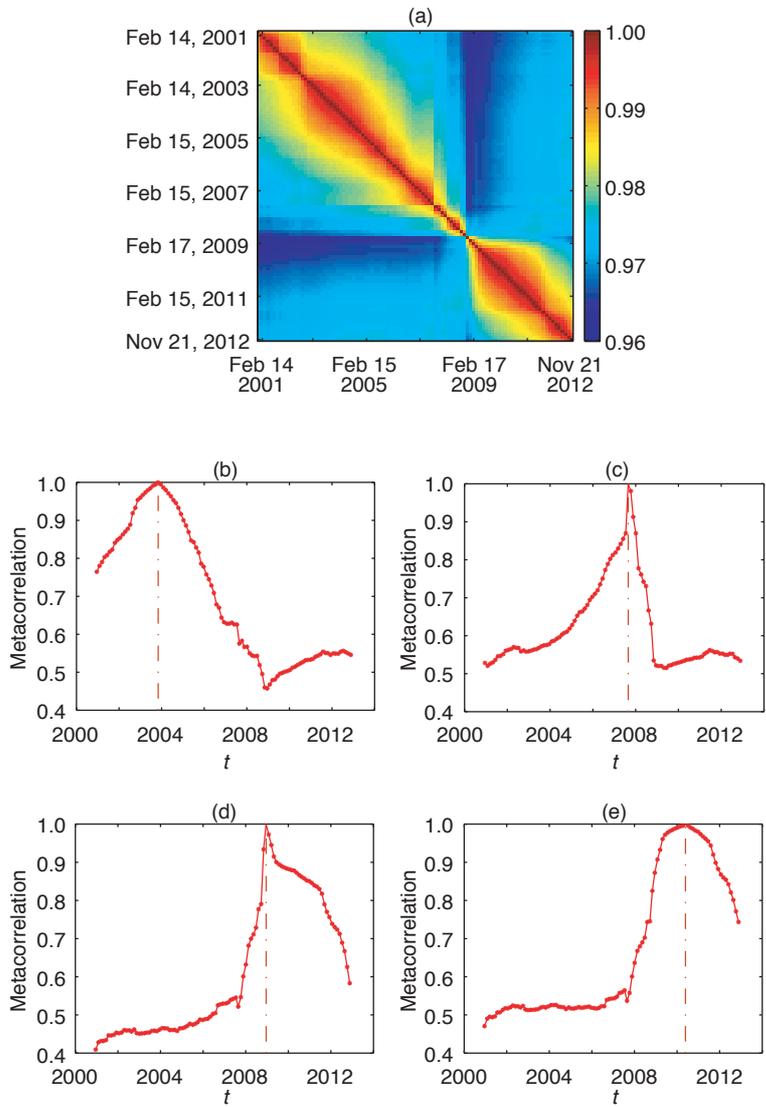
In order to analyze the amount of economic information expressed by the clustering (Coronnello *et al* 2011; Mantegna 1999), we measured the adjusted Rand index,  $\mathcal{R}_{\text{adj}}$  (Hubert and Arabie 1985), between the DBHT clustering at time window  $T_k$  and the community partition generated by the industrial sector classification of stocks.  $\mathcal{R}_{\text{adj}}$  is an index that measures the similarity between two different partitions on the same set of objects (stocks in this case) and ranges from 0 (no similarity) to 1 (complete identity). We provide a formal definition of this index in Appendix A.  $\mathcal{R}_{\text{adj}}$  therefore provides a measure of the industrial information contained in the correlation-based clustering. We use the industrial classification benchmark (ICB), which is a categorization that divides the stocks into four hierarchical levels: namely, 114 subsectors, 41 sectors and 19 different supersectors (which, in turn, are gathered in ten different industries). In order to take all of these levels into account, we measured  $\mathcal{R}_{\text{adj}}(T_k)$  between each of the hierarchical levels and DBHT clustering. In part (b) of Figure 1 on the preceding page, we plot the evolution over time of  $\mathcal{R}_{\text{adj}}(T_k)$  between the DBHT clusters and ICB industries, supersectors and subsectors (for simplicity, we do not plot sector data that is very close to supersectors' values). We can see how the ICB information shows a remarkable drop during the 2007–8 financial crisis, which partially recovers from 2010 onward. Interestingly, before the crisis, the industry, supersector and subsector lines were distinct (with ICB supersectors showing the highest similarity with the DBHT, followed by industries and subsectors), whereas in the crisis and postcrisis periods they display much closer values. Therefore, from the crisis onward correlation clustering is no longer able to distinguish between different levels of ICB. This might indicate that this industrial classification is becoming a less reliable benchmark to diversify risk. These results are confirmed by other industrial partitions, including the Yahoo classification.

The adjusted Rand index can also be used as a tool for analyzing the persistence of the DBHT clustering by measuring the index between two clusterings at two adjacent time windows (we denote by  $\mathcal{R}_{\text{adj}}^{T-1, T}(T_k)$  such a quantity). This gives a measure of local persistence: a drop in the index value indicates decreasing similarity between adjacent clusterings, and therefore less persistence. In part (c) of Figure 1 on the preceding page, we plot  $\mathcal{R}_{\text{adj}}^{T-1, T}(T_k)$  against time. We observe that the clustering persistence changes remarkably over time, dropping in particular with the outbreak of the financial crisis and recovering in 2010. Note that the drop during the crisis starts earlier than the actual outbreak of it (August 2007, the dashed vertical line). This could highlight a possible use of clustering persistence as a tool to forecast systemic risk. Notably, in 2010–12 we again observe a steadily decreasing trend. Interestingly, the pattern of persistence appears to be related to the similarity between clustering and

**FIGURE 2** Persistence analysis based on clustering.

(a) Similarity matrix  $s$  showing the temporal evolution of the correlation-based DBHT clustering. Each entry  $s(T_a, T_b)$  is the adjusted Rand index between clustering  $X_a$  and  $X_b$  at time windows  $T_a$  and  $T_b$  respectively (3.3). Higher values indicate greater similarity. The matrix displays two main blocks of high intrasimilarity: one precrisis and the other postcrisis. The years 2007–2008 fall between these two blocks and display very low similarity with any other time window, revealing an extremely changeable structure. Parts (b)–(e) show the patterns of similarity for four sample time windows (ie, four sample rows of the similarity matrix): (b) September–October 2003, (c) July–August 2007, (d) November–December 2008 and (e) April–May 2010. During the crisis, similarity decays much faster than in the precrisis and postcrisis periods.

**FIGURE 3** Persistence analysis based on metacorrelation.



(a) Similarity matrix  $z$  showing the temporal evolution of correlation matrixes. Each entry  $z(T_a, T_b)$  is calculated as a correlation among correlation matrixes at time windows  $T_a$  and  $T_b$  (3.4). Higher values indicate higher similarity. Parts (b)–(e) show the patterns of similarity for four sample time windows: (b) September–October 2003, (c) July–August 2007, (d) November–December 2008 and (e) April–May 2010. The decay during the crisis years is much less steep than in the corresponding plot in Figure 2 on the preceding page.

ICB, with periods of higher persistence characterized by higher amounts of economic information.

However, the drawback of  $\mathcal{R}_{\text{adj}}^{T-1,T}(T_k)$  as a measure of persistence is that, at any one time, it only provides information on the persistence with respect to the previous, adjacent time window. It tells us nothing about the long-term robustness of each clustering. To investigate this aspect, in Section 3.1 we discuss a set of analyses that evaluate the persistence of each clustering at each time, therefore providing a more complete picture.

### 3.1 A map of structural changes

To investigate the long-term persistence of each clustering, we calculated the adjusted Rand index for each time window between the corresponding clustering and the clustering at any other time. The result is summarized in the (symmetric) similarity matrix  $s$ :

$$s(T_a, T_b) = \mathcal{R}_{\text{adj}}(X_a, X_b), \quad (3.3)$$

where  $X_a$  and  $X_b$  are the DBHT clusterings at time windows  $T_a$  and  $T_b$  respectively. The matrix  $s$  for our data set is shown in part (a) of Figure 2 on page 9. We observe two main blocks, the first precrisis and the other postcrisis, within which high similarity among clusterings may be found. The two blocks show very low mutual similarity (upper right corner/lower left corner of the matrix). The first block begins losing its compactness in 2007, and the second block quite quickly does the same at the beginning of 2011. Between these two times, the outbreak of the financial crisis displays a series of extremely changeable clusterings that do not show similarity with any other time window.

To better highlight these changes of regime, we plot in parts (b)–(e) of Figure 2 on page 9 four time rows from matrix  $s$ , taken as examples of persistence behavior during the precrisis (part (b) September–October 2003), crisis (part (c) July–August 2007), the outbreak of the crisis, and part (d) November–December 2008, the aftermath of Lehman Brothers' default) and postcrisis periods (part (e) April–May 2010). The vertical dashed lines show the end position of the time window whose clustering is taken as a reference. Each point in the plot is the adjusted Rand index between that clustering and all the other clusterings at each other time window, in both the past and the future. In the precrisis period (b), the similarity displays a quite slow decay both forward and backward in time; the original clustering still has a 60% similarity with the seventeenth time window forward/backward in time. The decreasing trend is, however, evident and becomes steeper during the crisis. Taking time windows during the financial crisis, (c) and (d), the pattern changes drastically: the similarity drops by 70–80% in a few months both backward and forward in time. The two stages of crisis also reveal some differences. While in the early crisis period (c) the similarity with

precrisis clusterings is higher than with postcrisis ones, in the post-Lehman Brothers period (d) the situation is reversed. Finally, the postcrisis period (e) shows a partially recovered persistence, although not at the same levels as the 2003 pattern.

We may wonder whether the structural changes highlighted by the clustering analyses can be detected directly by studying the original, unfiltered correlation matrixes. To check this, we introduce an alternative measure of similarity among different time windows that does not make any use of clustering: namely, the correlation between the coefficients of two correlation matrixes (metacorrelation). This measure is

$$z(T_a, T_b) = \frac{\langle \rho_{ij}(T_a) \rho_{ij}(T_b) \rangle_{ij}}{\sqrt{[\langle \rho_{ij}^2(T_a) \rangle_{ij} - \langle \rho_{ij}(T_a) \rangle_{ij}^2][\langle \rho_{ij}^2(T_b) \rangle_{ij} - \langle \rho_{ij}(T_b) \rangle_{ij}^2]}}, \quad (3.4)$$

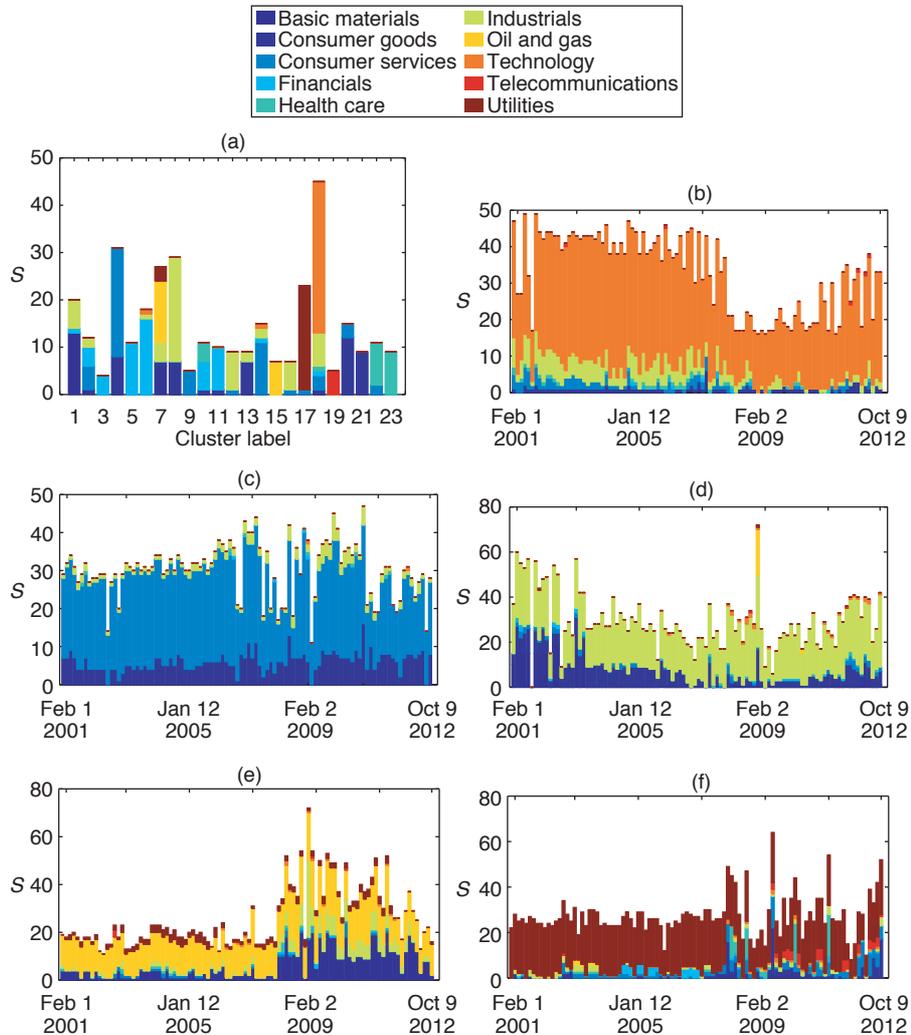
where  $\rho_{ij}(T_a)$  is the correlation between stocks  $i$  and  $j$  at time window  $T_a$  and  $\langle \cdot \rangle_{ij}$  is the average over all couples of stocks  $i, j$ . Munnix *et al* (2012) introduced an alternative measure to identify the possible states of a financial market. In Figure 3 on page 10, we report the matrix  $z(T_a, T_b)$  and four representative time rows, which correspond to the same four time windows chosen in Figure 2 on page 9. We observe that metacorrelation is indeed able to identify the two precrisis and postcrisis time blocks. However, it also shows a smaller, intermediate block during the 2007–8 crisis with a relatively high intrasimilarity. This is different than what we have observed in the clustering-based matrix  $s$ , where the time windows during the crisis were quite dissimilar. Moreover, the precrisis and postcrisis blocks in  $z$  display higher intrasimilarity than  $s$ , especially over the postcrisis years. All these differences can be appreciated when looking at the four  $z$  time rows in parts (b)–(e) of Figure 3. Even if in the crisis time windows (c) and (d) we observe a faster decay of similarity, it is much less steep than the corresponding clustering plot (parts (c) and (d) of Figure 2). Moreover, the postcrisis window in part (e) of Figure 3 recovers completely the high precrisis level of persistence, unlike the clustering case in part (e) of Figure 2.

Therefore, it seems that metacorrelation and clustering analyses depict different dynamics of market correlation structure. In particular, the clustering-based matrix  $s$  reveals higher nonstationarity during the crisis and postcrisis periods. The instability of correlation during crises has recently been observed by Chetalova *et al* (2014); however, their result relies on a specific choice for the multivariate distribution of returns, whereas our analyses are model independent.

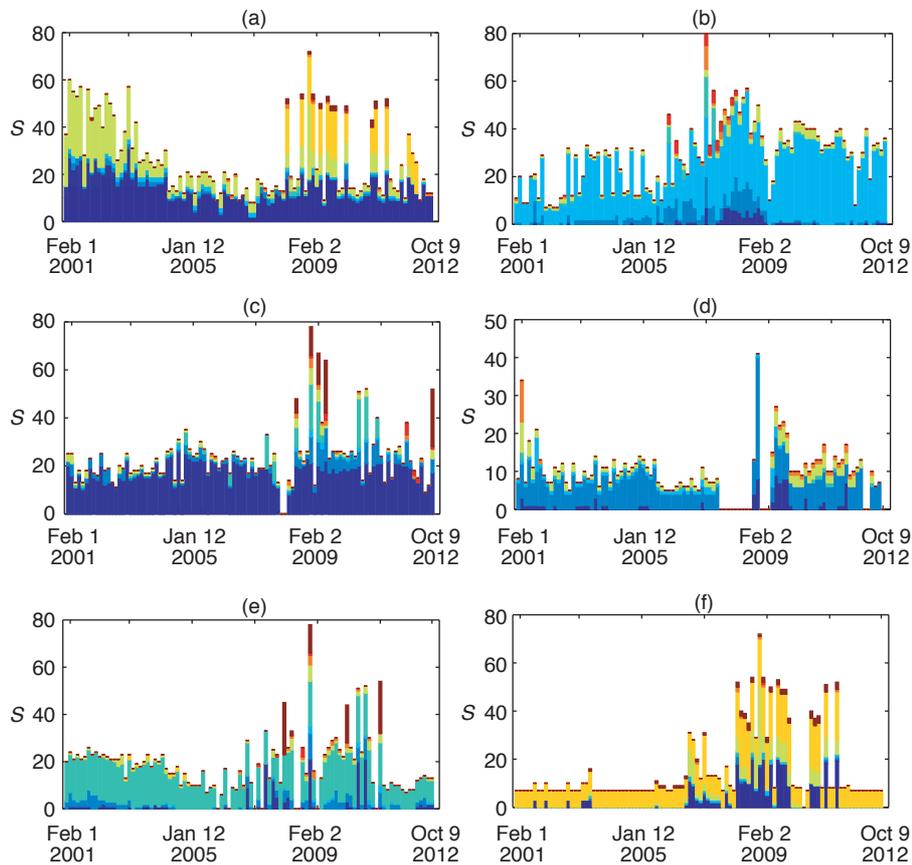
### 3.2 Clusters composition evolution

So far, we have described the persistence of clusters from a global perspective, looking at the clustering as a whole. Let us here focus on the evolution of each cluster, following how their composition changes over time. It is not straightforward to analyze such an evolution, with the main problem being the changeable nature of dynamical

**FIGURE 4** Clusters: dynamical composition (part 1).



(a) Clusters composition of the DBHT clusters obtained by calculating detrended log returns over the entire time window 1997–2012. The number of stocks in each cluster is shown on the  $y$ -axis, with different colors for different ICB industries. (b) For cluster 18 in (a), we have detected at each time window the corresponding “similar” (according to the hypergeometric test) cluster and plotted the composition in time. Zero size corresponds to no “similar” cluster having been found. When more than one “similar” cluster is found, only data for the largest cluster is plotted. Parts (c)–(f) show the same plots as in part (b), but for persistence clusters 4, 8, 7 and 17 respectively.

**FIGURE 5** Clusters: dynamical composition (part 2).

(a) For cluster 1 in part (a) of Figure 4 on the preceding page, we detected at each time window the corresponding “similar” (according to the hypergeometric test) cluster and plotted the composition in time. Zero size corresponds to no “similar” cluster having been found. When more than one “similar” cluster is found, only data of the largest cluster is plotted. Parts (b)–(f) show the same plots as in part (a), but for clusters 6, 20, 14, 22 and 15 respectively. Colors refer to the legend in Figure 4 on the preceding page.

clusters, which makes it difficult to identify the successor for each cluster. Many different approaches can be adopted to address this community tracking problem (Fenn *et al* 2012). Here, we use hypothesis statistical tests based on the hypergeometric distribution (Feller 2008; Tumminello *et al* 2011a) to assess similarity between clusters at different times. In particular, if the number of stocks in common between two clusters is high enough to reject the null hypothesis of the test, we label the two clusters as “similar”. Moreover, we take the DBHT clustering calculated over the entire

time window (1997–2012) as a benchmark clustering through which we can track the evolution of the dynamical clusters obtained with the moving time windows. Let us here describe this idea in more detail.

Let us call  $X$  the clustering obtained on the entire time window and  $Y_i$  a cluster belonging to  $X$ , with  $i = 1, \dots, N_{cl}$ . For each cluster  $Y_i$ , and for each time window  $T_k$  ( $k = 1, \dots, n$ ), we have taken the clustering at time  $T_k$ ,  $X_{T_k}$ , and identified the cluster belonging to  $X_{T_k}$  that is “similar” to  $Y_i$  (if any). We label a cluster as “similar” to  $Y_i$  if the number of stocks in common with  $Y_i$  is high enough to reject the null hypothesis of the hypergeometric test (Musmeci *et al* 2014; Tumminello *et al* 2011b). This test considers a random overlapping between the two clusters (a detailed description of the test can be found in Appendix B). If more than one cluster turns out to be similar, we take the largest cluster. Eventually, we end up with one cluster for each  $Y_i$  for each time window  $T_k$ . All of them have a high degree of similarity with  $Y_i$  in common. We can therefore follow their evolution in terms of the number of stocks and corresponding ICB industrial sectors. The threshold for the tests was chosen equal to 0.01, together with the conservative Bonferroni correction (Feller 2008).

In part (a) of Figure 4 on page 13, the composition of the DBHT clustering computed over the time window 1997–2012 is shown. For each cluster, the  $y$ -axis displays its cardinality  $S$  (ie, the number of stocks belonging to the cluster), with different colors showing stocks belonging to different ICB industries. For the eleven biggest clusters in  $X$ , we plot in parts (b)–(f) of Figure 4 and parts (a)–(f) of Figure 5 on the facing page the number of stocks  $S$  for their similar clusters in time, together with their composition in terms of ICB industries. When no similar clusters can be found for a time window, we have just left the correspondent window empty. The clusters analyzed are the numbers 18, 4, 8, 7, 17, 1, 6, 20, 14, 22 and 15. We summarize the main findings below.

- Overall, all the clusters in  $X$  have a high persistence over time, showing a corresponding “similar” cluster at almost every time window. This result is remarkable, as the persistence has been assessed in quite a conservative way, ie, using the hypergeometric test with a Bonferroni correction. A few clusters display a limited number of gaps in their evolution (eg, clusters 14, 15, 20 and 22), mostly in correspondence with the financial crisis.
- A few clusters show a persistence in terms of industrial composition as well (this is the case with clusters 4 and, to a lesser extent, 8), but most show a clear evolution. In particular, we can quite clearly distinguish a precrisis state and a postcrisis state; the latter is characterized by a higher degree of mixing of different industries. If, over the precrisis period, we find clusters dominated by

one or two industries (technology and industrials in cluster 18, oil and gas in 4 and 15, utilities in 17, consumer services and goods in 14 and 20, financials in 6, health care in 22), in the crisis and postcrisis years the industries tend to mix together much more, forming combinations that were not present earlier (oil and gas with basic materials and industrials in clusters 1 and 7, utilities with telecommunications and consumer services in 17, financials with consumer goods and services in 6, health care with utilities and consumer goods in 20). This again shows that the years since the crisis have seen a drop in the reliability of industries as benchmarks to diversify risk.

- Apart from the precrisis and postcrisis dichotomy, in some cases the 2007–8 crisis years show their own features as well. As stated above, some clusters “disappear” during the peak of the crisis (clusters 14, 20 and 22). Many others instead show several peaks in their sizes, together with a sudden increase in the number of industries. This is probably related to the merging of many clusters into fewer, larger clusters during the crisis.
- The cluster containing financial stocks (cluster 6) is worth analyzing further, since it seems to play a role in the outbreak of the financial crisis. Indeed, it shows a clear change in 2007, becoming larger and including an increasing number of different industries (especially health care, technology and consumer services). This pattern is probably connected to the rising importance of the financial industry as a driving factor over the outbreak of the crisis. Interestingly, at the end of 2008, when Lehman Brothers went bankrupt, this cluster suddenly drops to a much lower size (although still higher than precrisis values) and a less mixed composition. This suggests that the financial industry ends up playing a major role in the correlation structure from 2009 onward.

#### 4 DISCUSSION

In this paper, we have investigated the dynamical evolution and nonstationarity of market-correlation structure by means of filtered-correlation networks. In particular, we have focused on PMFGs and the clustering that its topology naturally provides by means of the DBHT method. We have measured the persistence of correlation structure by calculating similarity among clusterings in different time windows, using the adjusted Rand index to quantify the similarity.

Our analyses reveal that the outbreak of the 2007–8 financial crisis marked a transition from relatively high levels of persistence to a much more unstable and changeable structure. The minimum persistence was reached at the end of 2008 when the crisis had fully unfolded. But the decay in persistence had already started in late 2006, well

before other warning signs were detectable. Correlation structure persistence eventually recovered in the second half of 2009 with relatively high values until the end of 2011. However, such a persistent structure had distinct features from the precrisis structure, including lower relations with the industrial sector's activities. Notably, since the end of 2011 we have been observing a new decay in persistence, which is signalling another unfolding change in the market structure. This also points out that since 2007 correlation matrixes from historical data, both filtered and unfiltered, have become more unstable and therefore less reliable instruments for risk diversification. Moreover, the decrease in the similarity between correlation-based clustering and the industrial sector implies that portfolio-diversification strategies based on economic activity considerations are expected to become less effective. Furthermore, the analysis of the evolving industrial sector composition of each single cluster reveals that most of them display a clear change with the crisis, which makes them more heterogeneous overall in terms of industrial sectors. In particular, we observed that one cluster, mainly made of financial stocks, experienced a sharp rise in its size and heterogeneity that likely reflected the breakdown of the late-2007 financial crisis. This could give interesting insights in terms of early warning signals that we plan to investigate further in future work.

We also plan to carry out the analyses discussed in this work by using alternative community detection methods on graphs (Fortunato 2010). The comparison of different algorithms is a hot topic in network theory (Aldecoa and Marín 2013; Lancichinetti and Fortunato 2009), and these analyses could give other insights into this issue from the perspective of financial data.

## APPENDIX A. ADJUSTED RAND INDEX

Following the notation of Wagner and Wagner (2007), let us call  $X$  the set of  $N$  objects.  $Y$  is a partition into communities of  $X$  or simply a clustering: that is, "a set  $Y = \{Y_1, \dots, Y_k\}$  of nonempty disjoint subsets of  $X$  such that their union equals  $X$ " (Wagner and Wagner 2007). Let us also say we have another clustering  $Y'$ . We call the matrix  $M = \{m_{ij}\}$  the "contingency table", where

$$m_{ij} \equiv |Y_i \cap Y'_j|, \quad (\text{A.1})$$

ie, the number of objects in the intersection of clusters  $Y_i$  and  $Y'_j$ . Let  $a$  be the number of pairs of objects that are in the same cluster in both  $Y$  and  $Y'$ , and let  $b$  be the number of pairs that are in two different clusters in both  $Y$  and  $Y'$ . Then, the Rand index is defined as the sum of  $a$  and  $b$ , normalized by the total number of pairs in  $X$ :

$$\mathcal{R}(Y, Y') \equiv \frac{2(a + b)}{N(N - 1)} = \sum_{i=1}^k \sum_{j=1}^l \binom{m_{ij}}{2}. \quad (\text{A.2})$$

We can assume a generalized hypergeometric distribution to be the null hypothesis associated with two independent clusterings; we describe this in detail in Appendix B. The adjusted Rand index is defined as the difference between the measured Rand index and its mean value under the null hypothesis, normalized by the maximum that this difference can reach:

$$\mathcal{R}_{\text{adj}}(Y, Y') \equiv \frac{\sum_{i=1}^k \sum_{j=1}^l \binom{m_{ij}}{2} - t_3}{\frac{1}{2}(t_1 + t_2) - t_3}, \quad (\text{A.3})$$

where

$$t_1 = \sum_i^k \binom{|Y_i|}{2}, \quad t_2 = \sum_j^l \binom{|Y'_j|}{2}, \quad t_3 = \frac{2t_1 t_2}{N(N-1)}. \quad (\text{A.4})$$

It turns out that  $\mathcal{R}_{\text{adj}} \in [-1, 1]$ , with 1 corresponding to the case of identical clusterings and 0 to two completely uncorrelated clusterings. Negative values instead show anticorrelation between  $Y$  and  $Y'$  (that is, the number of pairs classified in the same way by  $Y$  and  $Y'$  is even less than was expected assuming a random overlapping between the two clusterings).

## APPENDIX B. HYPERGEOMETRIC TEST

Following the notation used in Section 3.2, let us call  $Y_i$  a generic cluster belonging to the clustering calculated over the entire time window. Let  $Y'_j$  be the cluster from clustering  $X_{T_k}$  in time window  $T_k$  with which we want to compare  $Y_i$  in order to find if the number of stocks belonging to both  $Y_i$  and  $Y'_j$  is sensitively higher than was expected by a random overlapping. This can be translated into a statistical one-tail hypothesis test, in which the null hypothesis is the hypergeometric distribution. Say  $k$  is the number of stocks  $Y_i$  and  $Y'_j$  have in common, whereas  $|Y_i|$ ,  $|Y'_j|$  are the cardinalities of the two clusters; then, the hypergeometric distribution reads (Feller 2008)

$$P(X = k) = \frac{\binom{|Y'_j|}{k} \binom{N-|Y'_j|}{|Y_i|-k}}{\binom{N}{|Y_i|}}. \quad (\text{B.1})$$

This distribution is consistent with a scenario in which the overlapping between the two clusters is due purely to chance. For this reason, it is a suitable null hypothesis for testing the similarity between clusters. If  $P(X = k)$  so-calculated is less than the significance level, then the test is rejected, and we conclude that the cluster  $Y_i$  overexpresses the cluster  $Y'_j$ , and they are therefore similar. The significance level of each test performed is 1%, together with the Bonferroni correction for multiple tests (Feller 2008).

## DECLARATION OF INTEREST

T. Di Matteo thanks the COST Action TD1210 for partly supporting this work. T. Aste acknowledges the support of the UK Economic and Social Research Council (ESRC) in funding the Systemic Risk Centre (ES/K002309/1).

## ACKNOWLEDGEMENTS

The authors thank Bloomberg for providing the data, and the referees for their useful suggestions.

## REFERENCES

- Aldecoa, R., and Marín, I. (2013). Exploring the limits of community detection strategies in complex networks. *Scientific Reports* **3**, 2216.
- Anderberg, M. R. (1973). *Cluster Analysis for Applications*. Academic Press.
- Arthur, D., and Vassilvitskii, S. (2007). *k*-means++: the advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 1027–1035. SIAM, Philadelphia, PA.
- Aste, T. (2012). An algorithm to compute Planar Maximally Filtered Graphs (PMFGs). URL: [www.mathworks.com/matlabcentral/fileexchange/38689-pmfg](http://www.mathworks.com/matlabcentral/fileexchange/38689-pmfg).
- Aste, T. (2014). An algorithm to compute Directed Bubble Hierarchical Tree (DBHT) clustering. URL: [www.mathworks.com/matlabcentral/fileexchange/46750-dbht](http://www.mathworks.com/matlabcentral/fileexchange/46750-dbht).
- Aste, T., Di Matteo, T., and Hyde, S. T. (2005). Complex networks on hyperbolic surfaces. *Physica A* **346**, 20–26.
- Aste, T., Shaw, W., and Di Matteo, T. (2010). Correlation structure and dynamics in volatile markets. *New Journal of Physics* **12**, 085009.
- Aste, T., Grammatica, R., and Di Matteo, T. (2012). Exploring complex networks via topological embedding on surfaces. *Physical Review E* **86**, 036109.
- Bartolozzi, M., Mellen, C., Di Matteo, T., and Aste, T. (2007). Multiscale correlations in different futures markets. *European Physical Journal B* **58**, 207–220.
- Borghesi, C., Marsili, M., and Miccichè, S. (2007). Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode. *Physical Review E* **76**, 026104.
- Buccheri, G., Marmi, S., and Mantegna, R. N. (2013). Evolution of correlation structure of industrial indices of US equity markets. *Physical Review E* **88**, 012806.
- Chetalova, D., Schäfer, R., and Guhr, T. (2014). Zooming into market states. Preprint, arXiv:1406.5386v1 [q-fin.ST].
- Coronnello, C., Tumminello, M., Lillo, F., Miccichè, S., and Mantegna, R. N. (2011). Sector identification in a set of stock return time series traded at the London Stock Exchange. *Acta Physica Polonica* **36**, 2653–2680.
- Di Matteo, T., and Aste, T. (2002). How does the eurodollars interest rate behave? *Journal of Theoretical and Applied Finance* **5**, 122–127.
- Di Matteo, T., Aste, T., and Mantegna, R. N. (2004). An interest rate cluster analysis. *Physica A* **339**, 181–188.

- Di Matteo, T., Aste, T., Hyde, S. T., and Ramsden, S. (2005). Interest rates hierarchical structure. *Physica A* **335**, 21–33.
- Di Matteo, T., Pozzi, F., and Aste, T. (2010). The use of dynamical networks to detect the hierarchical organization of financial market sectors. *European Physical Journal B* **73**, 3–11.
- Donetti, L., and Muñoz, M. A. (2004). Detecting network communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics: Theory and Experiment* **2004**(10), P10012.
- Fama, E. F., and French, K. R. (2004). The capital asset pricing model: theory and evidence. *Journal of Economic Perspectives* **18**(3), 25–46.
- Feller, W. (2008). *An Introduction to Probability Theory and Its Applications*, Volume 2. Wiley.
- Fenn, D. J., Porter, M. A., Mucha, P. J., McDonald, M., Williams, S., Johnson, N. F., and Jones, N. S. (2012). Dynamical clustering of exchange rates. *Quantitative Finance* **12**(10), 1493–1520.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports* **486**, 75–174.
- Guimerà, R., Sales-Pardo, M., and Amaral, L. A. N. (2004). Modularity from fluctuations in random graphs and complex networks. *Physical Review E* **70**(2), 025101.
- Hu, Y., Li, M., Zhang, P., Fan, Y., and Di, Z. (2008). Community detection by signaling on complex networks. *Physical Review E* **78**(1), 016115.
- Hubert, L., and Arabie, P. (1985). Comparing partitions. *Journal of Classification* **2**, 193–218.
- Jang, W., Lee, J., and Chang, W. (2010). Currency crises and the evolution of foreign exchange market: evidence from minimum spanning tree. *Physica A* **390**(4), 707–718.
- Kaufman, L., and Rousseeuw, P. J. (1987). Clustering by means of medoids. In *Statistical Data Analysis Based on the  $L_1$ -Norm and Related Methods*, Dodge, Y. (ed), pp. 405–416. North-Holland.
- Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., and Ben-Jacob, E. (2010). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PLoS ONE* **5**(12), e15032.
- Kohonen, T., Schroeder, M. R., and Huang, T. S. (2001). *Self-Organizing Maps*, 3rd edn. Springer.
- Lancichinetti, A., and Fortunato, S. (2009). Community detection algorithms: a comparative analysis. *Physical Review E* **80**(5), 056117.
- Livan, G., Inoue, J., and Scalas, E. (2012). On the non-stationarity of financial time series: impact on optimal portfolio selection. *Journal of Statistical Mechanics: Theory and Experiment* **2012**(7), P07025.
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *European Physical Journal B* **11**, 193–197.
- McDonald, M., Suleman, O., Williams, S., Howison, S., and Johnson, N. F. (2008). Impact of unexpected events, shocking news, and rumors on foreign exchange market dynamics. *Physical Review E* **77**, 046110.
- Mitrović, M., and Tadić, B. (2009). Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities. *Physical Review E* **80**(2), 026123.

- Munnix, M. C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T. H., Guhr, T., and Stanley, H. E. (2012). Identifying states of a financial market. *Scientific Reports* **2**, 644.
- Musmeci, N., Aste, T., and Di Matteo, T. (2014). Relation between financial market structure and the real economy: comparison between clustering methods. Working Paper, arXiv:1406.0496v1 [q-fin.ST].
- Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E* **69**(6), 066133.
- Newman, M. E. J., and Girvan, M. (2004). Community detection in graphs. *Physical Review E* **69**(2), 026113.
- Onnela, J. P., Chakraborti, A., Kaski, K., and Kertész, J. (2003a). Dynamic asset trees and Black Monday. *Physica A* **324**, 247–252.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J., and Kanto, A. (2003b). Dynamics of market correlations: taxonomy and portfolio analysis. *Physical Review E* **68**, 056110.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J., and Kanto, A. (2003c). Asset trees and asset graphs in financial markets. *Physica Scripta* **T106**, 48–54.
- Pozzi, F., Di Matteo, T., and Aste, T. (2012). Exponential smoothing weighted correlations. *European Physical Journal B* **85**(6), 175, 295.
- Pozzi, F., Di Matteo, T., and Aste, T. (2013). Spread of risk across financial markets: better to invest in the peripheries. *Scientific Reports* **3**, 1665.
- Reichardt, J., and White, D. R. (2007). Role models for complex networks. *European Physical Journal B* **60**, 217–224.
- Rosvall, M., and Bergstrom, C. T. (2007). An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7327–7331.
- Ruan, J., Dean, A., and Zhang, W. (2010). A general co-expression network-based approach to gene expression analysis: comparison and applications. *BMC Systems Biology* **4**(8), 1–21.
- Shi, J., and Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(8), 888–905.
- Song, W. M., Di Matteo, T., and Aste, T. (2011). Nested hierarchies in planar graphs. *Discrete Applied Mathematics* **159**, 2135–2146.
- Song, W. M., Di Matteo, T., and Aste, T. (2012). Hierarchical information clustering by means of topologically embedded graphs. *PLoS ONE* **7**(3), e31929.
- Tola, V., Lillo, F., Gallegati, M., and Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control* **32**(1), 235–258.
- Tumminello, M., Aste, T., Di Matteo, T., and Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10421–10426.
- Tumminello, M., Lillo, F., and Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior and Organisation* **75**, 40–58.
- Tumminello, M., Miccichè, S., Lillo, F., Piilo, J., and Mantegna, R. N. (2011a). Statistically validated networks in bipartite complex systems. *PLoS ONE* **6**(3), e17994.
- Tumminello, M., Miccichè, S., Lillo, F., Varho, J., Piilo, J., and Mantegna, R. N. (2011b). Community characterization of heterogeneous complex systems. *Journal of Statistical Mechanics: Theory and Experiment* **1**, P01019.

- Wagner, S., and Wagner, D. (2007). Comparing clusterings: an overview. Technical Report, ITI Wagner, Faculty of Informatics, Universität Karlsruhe (TH).
- West, D. B. (1996). *Introduction to Graph Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Zhou, H. (2003). Distance, dissimilarity index, and network community structure. *Physical Review E* **67**(6), 061901.