AUCO Czech Economic Review 4 (2010) 330–340
Acta Universitatis Carolinae Oeconomica

Dynamics of Stock Market Correlations

Dror Y. Kenett*, Yoash Shapira*, Asaf Madi, Sharron Bransburg-Zabary**, Gitit Gur-Gershgoren†, Eshel Ben-Jacob*‡**

Received 3 March 2010; Accepted 5 September 2010

Abstract We present a novel approach to the study the dynamics of stock market correlations. This is achieved through an innovative visualization tool that allows an investigation of the structure and dynamics of the market, through the study of correlations. This is based on the Stock Market Holography (SMH) method recently introduced. This qualitative measure is complemented by the use of the eigenvalue entropy measure, to quantify how the information in the market changes in time. Using this innovative approach, we analyzed data from the New York Stock Exchange (NYSE), and the Tel Aviv Stock Exchange (TASE), for daily trading data for the time period of 2000–2009. This paper covers these new concepts for the study of financial markets in terms of structure and information as reflected by the changes in correlations over time.

Keywords Correlation, Stock Market Holography, eigenvalue entropy, sliding window
JEL classification C60, C63, C65

1. Introduction

To date, the fact that financial systems exhibit distinct dynamical and chaotic behavior is well understood. Much work has been devoted to the analysis of financial data and financial systems, yet the dynamics of such systems remains a puzzling mystery. Commonly used and well documented methods include autocorrelations (Mantegna and Stanley 2000), non-linear time series analysis (Kodba et al. 2005), cross correlation (Coronnello et al. 2005; Coronnello et al. 2007; Garas and Argyrakis 2007; Gopikrishnan et al. 2000; Jung et al. 2006; Laloux et al. 1999; Mantegna 1999; Noh 2000; Pafka and Kondor 2004; Plerou et al. 2002; Utsugi et al. 2004), eigenvalue analysis, hierarchal clustering (Coronnello et al. 2005; Coronnello et al. 2007; Garas and Argyrakis 2007; Gopikrishnan et al. 2000; Jung et al. 2006; Laloux et al. 1999; Mantegna 1999; Noh 2000; Pafka and Kondor 2004; Plerou et al. 2002; Utsugi et al. 2004). The events of the recent past, seeing what is perhaps the biggest economic crisis since the great depression in the 1930's emphasize the importance of the attempts to study and understand these dynamical properties.

* Tel-Aviv University, School of Physics and Astronomy, The Raymond and Beverly Sackler Faculty of Exact Sciences, 69978 Tel-Aviv, Israel.
** Tel Aviv University, Faculty of Medicine, 69978 Tel Aviv, Israel.
† Israel Securities Authority, 22 Kanfei Nesharim St., 95464 Jerusalem, Israel; School of Business and Management, Ben Gurion University, 84105 Beer Sheva, Israel.
‡ Corresponding author. Phone: +972-3-6407845, E-mail: eshelbj@gmail.com.

Recently, we have investigated system level information embedded in the stock market (Shapira et al. 2009), such as the existence of modular organization into subgroups that share similar dynamical properties. Identification of such system level organization is essential for understanding the complexity of the market behavior. In our previous work (Shapira et al. 2009) we focused on the stationary correlations between stocks, mainly by using the Stock Market Holography (SMH) analysis—that is the correlations between stocks calculated for the entire time period investigated.

Here we re-analyze the data presented in our previous work (Shapira et al. 2009), belonging to the New York and Tel Aviv stock markets. For both markets we computed the matrices of stock correlations (correlations between the relative daily return of the different stocks) using the Pearson's pair-wise correlations. However, here we extended the length of the investigated time period, which is now a period of 9 years, from 1/2000–03/2009. The correlation matrices were investigated using the Stock Market Holography (SMH) methodology (Shapira et al. 2009).

The SMH method includes collective normalization of the correlations according to the correlations of each stock with all the others followed by dimension reduction algorithms (Principal Component Analysis algorithm – PCA, Chou 1975) which is applied on the matrices of normalized correlations. The results are presented by placing the stocks (and the index when it is included) in a reduced 3-dimensional PCA space (whose axes are the three leading principal components of the PCA). Using Principal Component Analysis (PCA) is similar in concept to the Random Matrix Theory (RMT) approach, used by many others (Garas and Argyrakis 2007; Jung et al. 2006; Laloux et al. 1999; Noh 2000; Plerou et al. 2002; Utsugi et al. 2004) to study stock movement cross-correlations. Both involve constructing the matrix of pair wise cross-correlations, and investigation of the principal eigenvalues of this matrix to identify the key driving forces of the market.

However, here we do not focus on the spectrum and statistics of the eigenvalues, rather on the structure and dynamics that govern the stocks in the reduced 3-dimensional space. Furthermore, the collective normalization we apply in the process of the SMH analysis uncovers hidden information about the system. Finally, unlike RMT, our SMH analysis tool enables a visual presentation of the stocks in the reduced correlation space. To retrieve information that can be lost in the dimension reduction process, the stocks in the reduced (Holographic) space are linked according to the correlations—color-coded lines (according to the correlations before normalization) are drawn between the stocks. Furthermore, it is possible to combine the SMH analysis with a simple sliding window approach, in order to uncover and study the dynamics of the market correlations.

Here we focus on the dynamics of the correlations. We do this using a sliding window approach. Keeping the Epps effect in mind (Epps 1979), we searched for the smallest possible time window which still contained in it a significant amount of data, and that was compliant with the Epps effect. For this purpose, we began with a time window of 500 time periods, and decreased it until we were able to account for the two constraints. Finally, it was found that a 22-day time window best meets these two criteria. For each time window, the correlation matrix is calculated for the given set of stocks. Then, the correlation matrix is used to calculate the average correlation between stock i and all the other stocks, which represents the relationship between the given stock and the market, for that time window. We then calculate the average correlation in the market in the given time window, and the STD of the correlations.

To gain a more comprehensive understanding of how the market evolves, we combine this sliding window approach with the SMH analysis. In each window we apply the SMH analysis, and are thus able to follow the time evolution of the correlations in the market in the special 3-D PCA space. Furthermore, using the idea of eigenvalue entropy (Kenett et al. 2009), we study how the information in the market evolves in time. This adds to the more qualitative sliding window SMH analysis a quantitative measure to study how the market changes in time.

2. Similarity matrices

We begin by calculating the stock raw correlations that are calculated using the Pearson correlation coefficient:

$$C(i,j) = \frac{(r(i) - \langle r(i)\rangle) \cdot (r(j) - \langle r(j)\rangle)}{\sigma(i) \cdot \sigma(j)}, \tag{1}$$

where $r(i)$ and $r(j)$ are the return of stock i and j, $\langle r(i)\rangle$ and $\langle r(j)\rangle$ denote the corresponding means, $\sigma(i)$ and $\sigma(j)$ are the corresponding standard deviations (STD). Note that $C(i,j)$ is a symmetric square matrix and $C(i,i) = 1$ for all i.

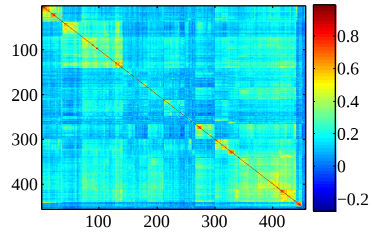
The correlation matrices are normalized using the affinity transformation, a special collective normalization procedure first proposed by Baruchi et al. (2006) and Baruchi et al. (2005). The idea is to normalize the correlations between each pair of stocks according to the correlations of each of the two stocks with all other stocks. This process is in fact calculation of the correlation of correlations or meta-correlation. The meta-correlations $MC(i,j)$ are the Pearson's correlation between rows i and j in the correlation matrix after reordering. In the reordering process, the elements $C(i,i)$ and $C(j,j)$ are taken out. The correlation vector for i is $\{C(i,j), C(i,1), C(i,2), \ldots\}$ and for j it is $\{C(i,j), C(j,1), C(j,2), \ldots\}$,

$$MC(i,j) = \frac{\sum_{k\neq i,j}^{N} (C(i,k) - \langle C(i)\rangle) \cdot (C(j,k)\langle C(j)\rangle)}{\left(\widehat{\langle C(i)^2\rangle} \cdot \widehat{\langle C(j)^2\rangle}\right)^{\frac{1}{2}}}. \tag{2}$$

In other words, the meta-correlation is a measure of the similarity between the correlations of stock i with all other stocks to the correlations of stock j with all other stocks. Using the meta-correlations, the normalized correlations $A(i,j)$ are

$$A(i,j) = \sqrt{C(i,j) \cdot MC(i,j)}. \tag{3}$$

The affinity transformation process emphasizes subgroups of variables (stocks) in the system, by removing the effect of the background noise of correlation. Groups (clusters) identified in the affinity matrix are significant in the system, and warrant



(a) The correlation matrix (b) The affinity matrix

Figure 1. Comparison of the correlation matrix to the affinity matrix for the 455 S&P500 stocks

further investigation. We demonstrate the strength of the affinity transformation in Figure 1, where we compare the S&P500 dataset correlation matrix to its affinity matrix. Both matrices are ordered similarly, and the groups weakly visible in the correlation matrix (Figure 1a) are emphasized and highlighted by the affinity transformation process (Figure 1b). The affinity transformation emphasizes the stock clusters, making them stand out in comparison to the background.

3. SMH analysis using a sliding window

While applying the Stock Market Holography (SMH), analysis on the entire time period, using the full stock time series, provides many insightful observations (Shapira et al. 2009), there is also room to consider applying the SMH analysis on shorter time periods. To this end, we combine the use of a sliding window algorithm together with the

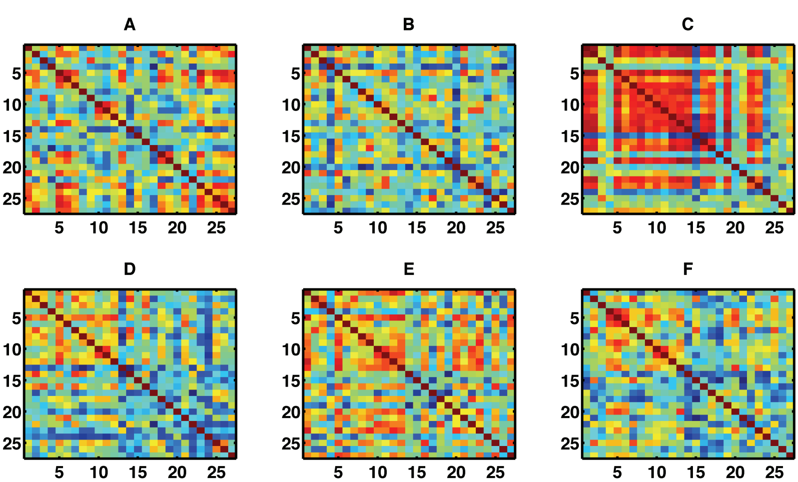


Figure 2. The normalized correlations matrices for the Tel-Aviv dataset

SMH analysis. At each time window, we compute the correlation matrix, normalize it to gain the affinity matrix, and then use the PCA algorithm to create the 3-dimensional space. Here we have used a 22-day time window, which corresponds to one working month of trading, however different time windows are also possible. While this is conceptually simple, it provides a lot of information and insights on the dynamics of the market. In Figure 2 we present some examples of different affinity matrices, calculated for different time windows, for the Tel-Aviv (TA) dataset. It is clear that the normalized correlations change quite significantly throughout time.

We are faced with two main problems when aiming to combine these two analysis methods. First, Due to the fact that at each time window we are in fact calculating new principal vectors, we first have to verify that the principal vectors at each time window truly do capture at least 75% of the variance of the system. Second, we in fact have n (n being the number of time windows) 3-dimensional spaces. In order to combine them, we choose one set of principal vectors, and project the results on the selected principal components. Here we have chosen to compute the 3 leading principal components for the entire time period of the given set of stocks, and then project on it the results of each time window SMH analysis. At each time window, we transpose the specific affinity matrix to coordinates on the three chosen principal vectors. Figure 3 presents the percentage of information in the first three principal vectors (15 different time windows shown). While the percentage varies for the different time windows, it remains above 75%.

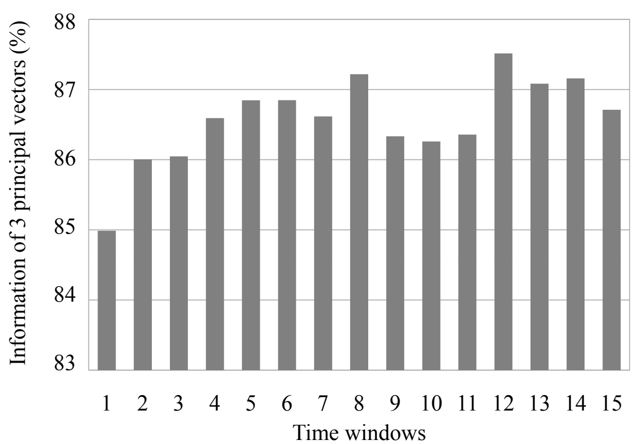
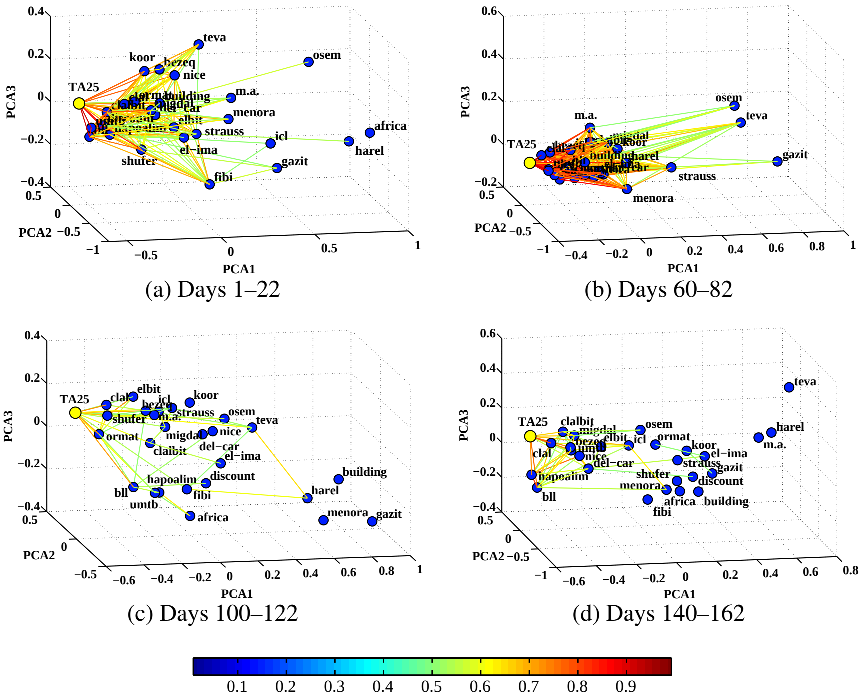


Figure 3. Percentage of information in the first three principal vectors

The outcome of this process is an animated movie of the stock correlations in the 3-dimensional affinity space. This tool allows an easy visual analysis of the dynamics of the system. An example of such movies is also presented in Figure 4, where we present four frames from such a movie for the TA dataset (see also http://tamar.tau.ac.il/∼dror).

Next, we can make use of the running window SMH analysis method to follow the



(a) Days 1–22

(b) Days 60–82

(c) Days 100–122

(d) Days 140–162

Note: The threshold for lines connecting the stocks is 0.5 for all panels.

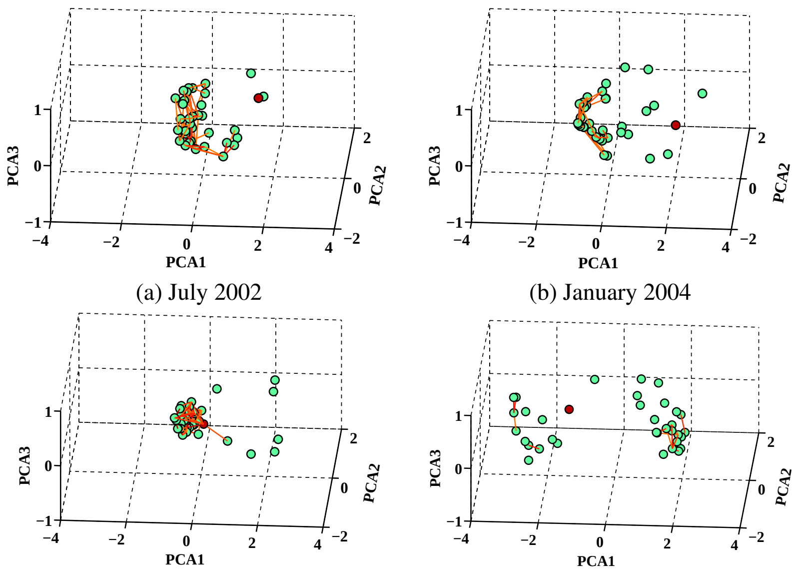
Figure 4. The SMH analysis using a running window for the TA stock dataset

stability of a given sector. For example, we focus on the energy sector (in this dataset, 34 stocks), and study how the intra-sector correlations evolve in time. Looking at Figure 5 we can see that the dispersion in the different panels shows how the correlations in the sector first become weaker, and then stronger, as time progresses.

4. Eigenvalue entropy

The concept of eigenvalue entropy has been used as a measure to quantify the deviation of the eigenvalue distribution from a uniform one (Alter et al. 2000). The idea was first used in the context of biological systems (Varshavsky et al. 2007; Varshavsky et al. 2006), and recently applied to the study of stock similarity matrices (Kenett et al. 2009). The spectral entropy, SE, is defined as

$$SE \equiv -\frac{1}{\log(N)} \sum_{i=1}^{N} \Omega(i) \log[\Omega(i)], \tag{4}$$



Note: S&P500 index marked in red. The threshold for lines connecting the stocks is 0.7 for all panels.

Figure 5. The SMH analysis using a running window for the energy sector stocks

where $\Omega(i)$ is given by

$$\Omega(i) = \frac{\lambda(i)^2}{\sum_{i=1}^{N} \lambda(i)^2}. \tag{5}$$

Note that the $1/\log(N)$ normalization was selected to ensure that $SE = 1$ for the maximum entropy limit of flat spectra (all λ are equal).

First, we make use of this measure to compare the information contained in the correlation matrix versus that contained in the affinity matrix (see Table 1, and also Kenett et al. 2009). Next, we study the entropy value of the similarity matrix computed for stocks only, versus that computed for the stocks and the index, as an additional "ghost" stock (see Shapira et al. 2009). These values are presented in Table 1. We find that more information (less entropy) is embedded in the affinity matrices in comparison to the correlation matrices, and that the inclusion of the index as a "ghost" stock adds significant information on the system.

Table 1. Entropy values for the S&P and TA datasets

	S&P		TA	
	Correlation	Affinity	Correlation	Affinity
Stocks+Index	0.1322	0.0972	0.1485	0.0408
Stocks	0.1344	0.1007	0.1720	0.0450

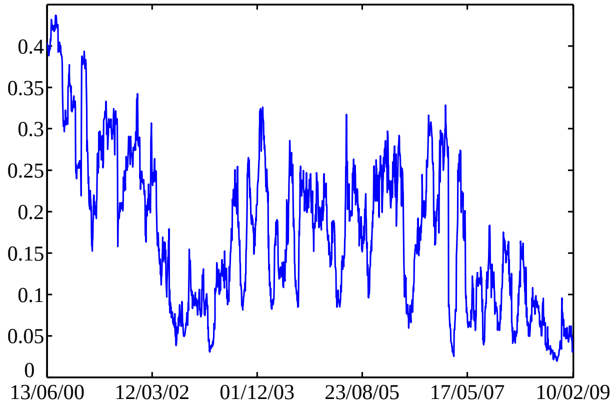


Figure 6. Sliding window entropy calculation for the S&P dataset

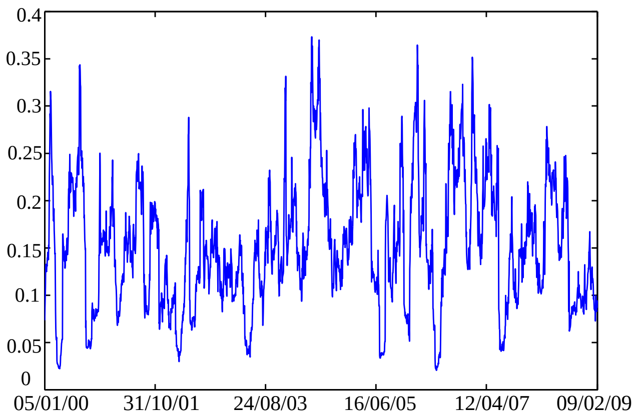


Figure 7. Sliding window entropy calculation for the TA dataset

Next, we make use of this quantitative measure to study how the information in the markets changes throughout time. This is done by repeating the sliding window analysis discussed above, using a 22 day window. In each time window we calculate the affinity matrix, and from it compute the eigenvalue entropy. We perform this sliding window entropy calculation for the two datasets separately. In Figure 6 we present the results for the S&P dataset. These results present evidence of significant changes in the entropy in the market, and moreover the existence of distinct time periods in respect to the market entropy.

In Figure 7 we present the results of the sliding window entropy calculation for the TA dataset. Once more, it is possible to observe the stochastic nature of the market entropy across time. Comparing Figure 7 to Figure 6, it s possible to note that the entropy for the TA dataset is not characterized by distinct time periods, as is in the case of the S&P dataset.

5. Discussion

We present here an investigation of the dynamics of stock market correlations using the SMH analysis (Shapira et al. 2009), and by means of entropy analysis (Kenett et al. 2009). This study was performed on an extension of the dataset investigated by Shapira et al. (2009). Our work is also motivated by the investigations of the stock market in terms of correlations pioneered by Mantegna and Stanley (2000).

The SMH analysis (Shapira et al. 2009) provides an innovative way of investigating the market structure and dynamics. Displaying the stocks in the special correlation based 3-dimensional space provides an easy and first-hand comprehension of the structure and relationships in the market. A key step in the SMH analysis is the affinity transformation. This collective normalization of the correlations reduces the noise embedded in the correlation matrix, and produces a better estimation of the real relationships between stocks. Furthermore, it allows for a comparison between different datasets, since the correlations are normalized in the same way. This allows, for example, to compare between different markets, or to compare between a set of stocks with and without the index.

Expanding the analysis using a sliding window algorithm, one can follow and study the dynamics of the correlations in the market, and how the market evolves throughout time. Unlike the case of the 2-D correlation matrices, the 3-D representation of the correlations through the SMH provides an easy and intuitive way to investigate how the correlations evolve in the market. This important visualization tool can be further used to study the stability of different sectors in the market, identify unusual formation of correlated groups of stocks, and uncover time periods in which the market structure changes. For example, this can tool can serve as an "early warning" mechanism for the regulators, trying to prevent unhealthy changes in the structure of the market.

In our previous work (Shapira et al. 2009), we discussed the presence of a special feedback mechanism between the index, and the stocks belonging to it. Here we made use of the eigenvalue entropy measure (Kenett et al. 2009) to further investigate this issue. We found that indeed, there is more information (less entropy) in the system, when we include the index in the analysis.

The eigenvalue entropy also provides us with a quantitative measure to study how the information in the market changes in time. Combining the eigenvalue entropy analysis with the sliding window approach achieves this goal. This analysis complements the sliding window SMH analysis—where the former is a quantitative tool, the latter is a qualitative tool to study the dynamics of the correlations in the market. Our findings show that the amount of information in the market changes quite significantly over time, and it is possible to observe time periods with significantly more information. Furthermore, we see that fluctuations of the entropy are quite different for the two investigated markets. The S&P dataset clearly exhibits periods characterized by different behavior of the entropy, while this is not the case for the TA dataset. This provides us with another way to quantify how a small emerging market (TASE) is different from a large mature on (NYSE).

In conclusion, we present here new tools for the empirical analysis of the dynamics of stock market correlations. The methods discussed here could be used to study the "healthiness" of the market, serve as a method to compare and evaluate the performance of different types of markets, and serve as a crises prediction mean.

Acknowledgment We have benefited from insightful conversation with H. Eugene Stanley and his thoughtful advice regarding the presentation of the results. This research has been supported in part by the Tauber Family Foundation and the Maguy-Glass Chair in Physics of Complex System at the Tel Aviv University.

References

Alter, O., Brown, P. and Botstein, D. (2000). Singular Value Decomposition for Genome-Wide Expression Data Processing and Modeling. *PNAS*, 97, 10101–10106.

Baruchi, I., Grossman, D., Volman, V., Shein, M., Hunter, J., Towle, V. L. and Ben-Jacob, E. (2006). Functional Holography Analysis: Simplifying the Complexity of Dynamical Networks. *Chaos*, 16(1), 015112.

Baruchi, I., Towle, V. L., and Ben-Jacob, E. (2005). Functional Holography of Complex Networks Activity: From Cultures to the Human Brain. *Complexity*, 10, 38–51.

Chou, Y. (1975). *Statistical Analysis*. New York, Holt, Rinehart & Winston.

Coronnello, C., Tumminello, M., Lillo, F., Micciche, S. and Mantegna, R. N. (2005). Sector Identification in a Set of Stock Return Time Series Traded at the London Stock Exchange. *Acta Physica Polonica B*, 36, 2653–2679.

Coronnello, C., Tumminello, M., Lillo, F., Micciche, S. and Mantegna, R. N. (2007). Economic Sector Identification in a Set of Stocks Traded at the New York Stock Exchange: A Comparative Analysis. *Noise and Stochastics in Complex Systems and Finance*, 6601, U198–U209.

Epps, T. W. (1979). Comovements in Stock-Prices in the Very Short Run. *Journal of the American Statistical Association*, 74, 291–298.

Garas, A. and Argyrakis, P. (2007). Correlation Study of the Athens Stock Exchange. *Physica A: Statistical Mechanics and its Applications*, 380, 399–410.

Gopikrishnan, P., Plerou, V., Liu, Y., Amaral, L. A. N., Gabaix, X. and Stanley, H. E. (2000). Scaling and Correlation in Financial Time Series. *Physica A: Statistical Mechanics and its Applications*, 287, 362–373.

Jung, W.-S., Chae, S., Yang, J.-S. and Moon, H.-T. (2006). Characteristics of the Korean Stock Market Correlations. *Physica A: Statistical Mechanics and its Applications*, 361, 263–271.

Kenett, D. Y., Shapira, Y. and Ben-Jacob, E. (2009). RMT Assessments of Market Latent Information Embedded in the Stocks' Raw, Normalized and Partial Correlations. *Hindawi Journal of Probability and Statistics*, 2009, DOI:10.1155/2009/249370.

Kodba, S., Perc, M. and Marhl, M. (2005). Detecting Chaos from a Time Series. *European Journal of Physics*, 26, 205–215.

Laloux, L., Cizeau, P., Bouchaud, J. P. and Potters, M. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83, 1467–1470.

Mantegna, R. N. (1999). Hierarchical Structure in Financial Markets. *European Physical Journal B*, 11, 193–197.

Mantegna, R. N. and Stanley, H. E. (2000). *An Introduction to Econophysics: Correlation and Complexity in Finance*. Cambridge, Cambridge University Press.

Noh, J. D. (2000). Model for Correlations in Stock Markets. *Physical Review E*, 61, 5981–5982.

Pafka, S. and Kondor, I. (2004). Estimated Correlation Matrices and Portfolio Optimization. *Physica A: Statistical Mechanics and Its Applications*, 343, 623–634.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. and Stanley, H. E. (2002). Random Matrix Approach to Cross Correlations in Financial Data. *Physical Review E*, 65, 066126.

Shapira, Y., Kenett, D. Y., and Ben-Jacob, E. (2009). The Index Cohesive Effect on Stock Market Correlations. *The European Physical Journal B*, 72, 657–669.

Utsugi, A., Ino, K. and Oshikawa, M. (2004). Random Matrix Theory Analysis of Cross Correlations in Financial Markets. *Physical Review E*, 70(2), 026110.

Varshavsky, R., Gottlieb, A., Horn, D. and Linial, M. (2007). Unsupervised Feature Selection under Perturbations: Meeting the Challenges of Biological Data. *Bioinformatics*, 23, 3343–3349.

Varshavsky, R., Gottlieb, A., Linial, M. and Horn, D. (2006). Novel Unsupervised Feature Filtering of Biological Data. *Bioinformatics*, 22, e507–e513.