LETTER

Multiscale horizontal-visibility-graph correlation analysis of stock time series

To cite this article: Weidong Li and Xiaojun Zhao 2018 *EPL* **122** 40007

A Letters Journal Exploring the Frontiers of Physics

May 2018

EPL, **122** (2018) 40007
doi: 10.1209/0295-5075/122/40007

www.epljournal.org

Multiscale horizontal-visibility-graph correlation analysis of stock time series

Weidong Li and Xiaojun Zhao[(a)]

School of Economics and Management, Beijing Jiaotong University - Beijing 100044, PRC

received 17 January 2018; accepted in final form 6 June 2018
published online 3 July 2018

PACS 05.45.Tp – Time series analysis
PACS 89.20.-a – Interdisciplinary applications of physics
PACS 89.75.-k – Complex systems

Abstract – This letter is devoted to measure the nonlinear interactions between non-stationary time series on multiple time scales. A graph-theoretic method by the node degrees relationship in the context of horizontal visibility graphs (HVGs) is proposed, which bridges the gap among time series analysis, multiscale analysis, and graph theory. We compare the new method with other measures, and study the degree properties and significance test. We then apply it to stock time series analysis, so as to quantify the information exchange between the daily closing price and daily trading volume in stock markets.

Introduction. – Quantifying the interactions among complex systems or the information exchange of individual components in a single system has been extensively studied in recent decades. It is a multidisciplinary topic and has attracted much attention, for instance in physiology, hydrology, geography, traffic science, economics and finance [1–6]. As a typical case in the context of stock markets, the trading volume of stocks has been found to be closely associated with the trading price [7]. Identifying the information linkage between these variables is beneficial to further understand the dynamical mechanism of stock markets, and can be used for risk management as well as the selection of assets portfolio.

The output time series of many complex systems, typically characterized by nonlinearity and non-stationarity, may break the *ad hoc* assumptions of many traditional approaches, which lead to the failure of these methods. In order to quantify the nonlinear interaction between time series, information-theoretic tools (*e.g.*, mutual information) and complex networks [8,9] have been proposed. Time series are mapped into graphs, and analyzed by the graph tools, *e.g.*, recurrence networks [10–12], visibility graphs (VGs) [13–16] and horizontal visibility graphs (HVGs) [17–19].

By construction, the recurrence networks do not make any use of time information, so they can commonly also contain connections between temporally distant observations which are mostly absent from (H)VGs. Long-distance connections are not necessarily required for stock time series analysis, since the stock markets change so fast and it is difficult to explain the meaning of these connections. The absence of long-distance connections in a HVG makes time series effectively decomposed into short segments, leading to some community structure of graphs [20,21]. The VG, however, is invariant under affine transformations of a time series, *i.e.*, allows for connections between observations even in case of linear trends that would not be present in the HVG [22]. Moveover, many statistical properties of the HVGs can be theoretically derived. By mapping time series into HVGs, the dynamics of the underlying systems can be traced out through the topology structure of the graphs.

In a complex system, the trajectory of the system in phase space is described by state vectors [23]; each state vector consists of more than one data point. While for HVGs, one datum of the original time series is represented by one node of the graph. This limitation of the HVGs can be solved by the multiscale analysis, by which several neighboring data points are averaged in order to obtain one node. Furthermore, in many real-world applications, the complex systems controlled by regulatory mechanisms that operate on different time scales, generate time series that exhibit highly variable fluctuations at multiple levels of resolution [24–26].

Recently, Gao *et al.* introduced a multiscale limited penetrable HVG for analyzing a single nonlinear time

[(a)]E-mail: xjzh@bjtu.edu.cn (corresponding author)

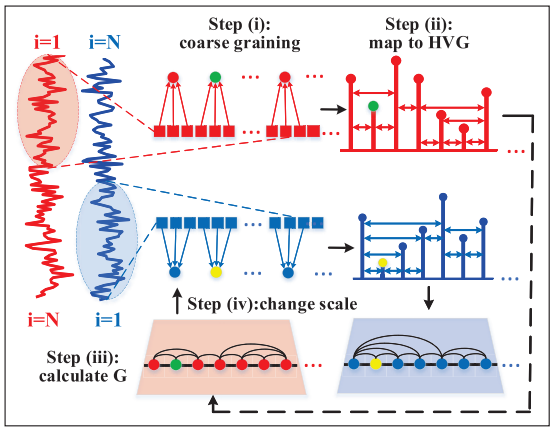

Fig. 1: (Colour online) Schematic sketch of the MHVGCA. Step (i): coarse grain two original time series by the scale s ($s = 3$ in this case). Step (ii): map the coarse-grained time series into HVGs, and compute the degree of each node. Step (iii): calculate the association measure G between two degree sequences. Step (iv): change the scale s, and repeat steps (i)–(iii) to obtain different $G(s)$ on multiple time scales.

series [27]. They also proposed a method for constructing a multiscale complex network from multivariate time series [28]. Zhang and Shang proposed a multiscale analysis for time irreversibility analysis based on phase-space reconstruction and HVGs [29]. Ahmadlou and Adeli proposed a VG similarity to measure generalized synchronization in coupled dynamic systems [30]. Zou *et al.* used joint and excess degrees of VGs to disentangle similarities and differences between the dynamical patterns exhibited by two simultaneously observed variables [31]. In this letter, we incorporate the multiscale analysis with HVGs, and propose a method of multiscale horizontal-visibility-graph correlation analysis (MHVGCA), in order to quantify the intrinsic interactions between two non-stationary time series. The MHVGCA is computationally efficient and does not require any *ad hoc* symbolization process. We compare the MHVGCA with other measures, and study the degree properties and significance test, then apply it to stock time series analysis.

Methodologies. –

Multiscale horizontal-visibility-graph correlation analysis (MHVGCA). Consider two non-stationary time series $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{N}$. To quantify their nonlinear association on multiple time scales, we propose a method of multiscale horizontal-visibility-graph correlation analysis (MHVGCA). The procedure of the MHVGCA consists of 4 steps (see the schematic sketch of the method in fig. 1).

(i) First coarse grain two time series by averaging them in small non-overlapping segments at the same time scale s, respectively,

$$X_i^{(s)} = \frac{1}{s}\sum_{j=1}^{s} x_{(i-1)s+j}, \quad Y_i^{(s)} = \frac{1}{s}\sum_{j=1}^{s} y_{(i-1)s+j}. \tag{1}$$

The coarse-grained time series are obtained by down sampling the original time series over s. The scale s can take integers $1, 2, \ldots, k$ ($k \ll N$). i ranges from 1 till to $\lfloor N/s \rfloor$, where $\lfloor N/s \rfloor$ is the largest integer not larger than N/s. By coarse-graining, many trivial details are smoothed, while the profile of the series is retained. When $s = 1$, the coarse-grained time series are equivalent to the original time series.

(ii) Next map the coarse-grained time series $\{X_i^{(s)}\}_{i=1}^{\lfloor N/s \rfloor}$ and $\{Y_i^{(s)}\}_{i=1}^{\lfloor N/s \rfloor}$ into HVGs and compute the node degrees, respectively. In the context of HVG, each datum of the coarse-grained time series is associated with a node. Two nodes are connected if they are horizontally visible to each other, *i.e.*,

$$H_{ij} = \begin{cases} 1, & \text{if } X_k^{(s)} < \mathrm{Inf}(X_i^{(s)}, X_j^{(s)}), \quad \forall k \in (i,j), \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

H_{ij} is set to 1 when two nodes i and j are horizontally visible, otherwise it is 0. No direction between nodes is defined in HVGs, and therefore, $H_{ij} \equiv H_{ji}$. When i and j are nearest neighbors, two nodes are naturally visible. The degree k_i of the i-th node gives the number of other nodes horizontally visible to this node:

$$k_i = \sum_{j=1, j\neq i}^{\lfloor N/s \rfloor} H_{i,j}. \tag{3}$$

The smallest value of k_i is 2, when the node falls in a valley and its nearest neighbors are both larger than it. Then this node is only horizontally visible to its two nearest neighbors, but not to itself.

(iii) In step (ii), we obtain two degree sequences from the coarse-grained time series, which are discrete and ordinal [32]. Next we use the Goodman and Kruskal's gamma [33] for ordinal variables analysis to quantify the similarity between the degree sequences:

$$G = \frac{N_s - N_d}{N_s + N_d}. \tag{4}$$

N_s represents the number of concordant pairs, *i.e.*, when $k_i^{\{X\}} < k_j^{\{X\}}$, $k_i^{\{Y\}} < k_j^{\{Y\}}$ (or when $k_i^{\{X\}} > k_j^{\{X\}}$, $k_i^{\{Y\}} > k_j^{\{Y\}}$). N_d represents the number of discordant pairs. G measures the strength of association of the cross-tabulated data when both variables are measured at the ordinal level. It makes no adjustment for either table size or ties. Values range from -1 (perfect inversion) to 1 (perfect agreement). $G \approx 0$ indicates the absence of association.

(iv) Change the scale s, and repeat the steps of (i)–(iii). Finally, we obtain a series of $G(s)$, which are used to quantify the strength of the interactions between time series on multiple time scales.

Comparison with other measures. We first compare the MHVGCA with classical methods for association analysis, including the Pearson correlation coefficient

Table 1: Comparison of 5 measures in the presence of non-stationarity. The values are the average of 10^2 simulations. Here we show the results of $\phi_1 = 10/N, \phi_2 = 0$, $\omega_1 = 10/N^2, \omega_2 = \omega_3 = 0$. MI is estimated by the method of k nearest neighbors. In the MHVGCA, the scale s is set to 1. For other scales, we obtain very similar results.

Method	No trends	$\phi_1 = 10/N$	$\omega_1 = 10/N^2$
Pearson's r	0.0007	0.8931	0.8991
Kendall's τ	0.0003	0.7135	0.6790
Spearman's ρ	0.0003	0.9019	0.8691
MI	0.0033	0.7779	0.7684
MHVGCA	0.0026	0.0035	0.0035

(Pearson's r), the Kendall rank correlation coefficient (Kendall's τ), the Spearman rank correlation coefficient (Spearman's ρ), and mutual information (MI). For these classical statistical association measures, Pearson's r and MI are directly estimated on the magnitude of the raw data; while Kendall's τ and Spearman's ρ are measured, indirectly, on the rank of the raw data. In comparison, the MHVGCA works indirectly with the degree of the raw data.

For two Gaussian white-noise series $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$ (mean values $u_x = u_y = 0$, variances $\sigma_x^2 = \sigma_y^2 = 1$) [34], we add up with linear trends $\{l_i\}_{i=1}^N$, and quadratic trends $\{t_i\}_{i=1}^N$, respectively:

$$\begin{cases} l_i = \phi_1 i + \phi_2, \\ t_i = \omega_1 i^2 + \omega_2 i + \omega_3, \end{cases} \tag{5}$$

and obtain the composite time series: $\{x_i + l_i\}$ and $\{y_i + l_i\}$, and $\{x_i + t_i\}$ and $\{y_i + t_i\}$. There exists a trade-off between the fluctuation (variance) of the original time series and the strength of the external trends. The original time series dominates when its fluctuation overwhelms the strength of the external trends, otherwise, the external trends would dominate.

For Gaussian white noise, all the measures give values close to zero. But for the composite time series, the values of these measures significantly increase when the strength of the linear or quadratic external trends gradually increases, except the MHVGCA (see table 1, where the results of $\phi_1 = 10/N, \phi_2 = 0$, and $\omega_1 = 10/N^2, \omega_2 = \omega_3 = 0$ are shown; we change the parameters and obtain similar results, which are not shown here). The reason behind the difference is simple. The traditional methods focus on distant associations between time series, the intrinsic associations of which are often disturbed by long-term external trends like l_i and t_i. By contrast, the MHVGCA emphasizes on small-scale associations, and therefore, the long-term trends have little impact on short periods.

Degree properties of HVGs. For white noise, the degrees of HVGs are exponentially distributed [17,35]:

$$P(k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2}, \quad k = 2, 3, \cdots. \tag{6}$$

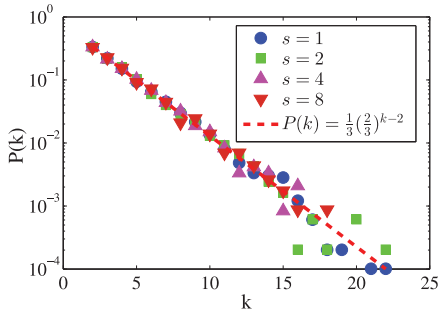

Fig. 2: (Colour online) Theoretical (dashed line) and empirical degree distribution of HVGs for Gaussian white noise. The scale for the coarse-graining is set to 1, 2, 4, and 8, respectively.

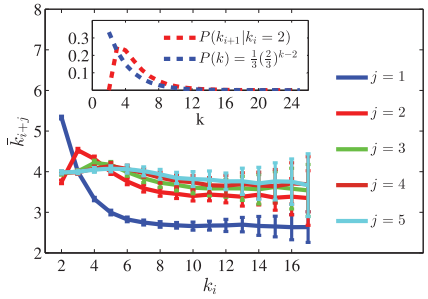

Fig. 3: (Colour online) The conditional mean values $\bar{k}_{i+j}|k_i$, $j = 1, 2, 3, 4, 5$. The error bars are estimated by 10^2 simulations (mean value $\pm$ standard deviation). Inset panel: the conditional probability $P(k_{i+1}|k_i = 2)$, and its comparison with the exponential degree distribution for white noise.

Particularly for Gaussian white noise ($\mu = 0$, $\sigma^2 = 1$), the coarse-grained time series are still Gaussian distributed ($\mu(s) = 0$, $\sigma^2(s) = 1/s$). More importantly, the coarse-grained time series are still white-noise series, as we use non-overlapping averaging for the multiscale analysis. In such a case, the node degrees of the coarse-grained time series are also exponentially distributed (see fig. 2).

We next state an anti-persistent property of the degree sequence of HVGs, *i.e.*, a small node degree is more likely to be followed by a large node degree, and vice versa. This is the case even for white noise. To certify this statement, we study the conditional mean values: $\bar{k}_{i+j}|k_i$, $j = 1, 2, \cdots$ of white-noise series. For $j = 1$, $\bar{k}_{i+1}$ significantly decreases with increasing k_i (see fig. 3). For $j = 2, 3, 4, 5$, the anti-persistence property gradually disappears. We also estimate the conditional probability $P(k_{i+1}|k_i = 2)$ as a special case (see the inset panel), which significantly deviates from the exponential distribution of white noise. If a node has the degree $k = 2$, it will be most likely to be followed by a node with degree 4 (the maximum point of the probability curve). Interestingly, the average degree of a white-noise series is $\bar{k} = \sum_{k=2}^{\infty} kP(k) = \sum_{k=2}^{\infty} k(1/3)(2/3)^k = 4$. The reason behind is that nodes are more likely to be only visible to their nearest neighbors (thus have degree 2) when they have small magnitude. In turn, nodes trend to be visible

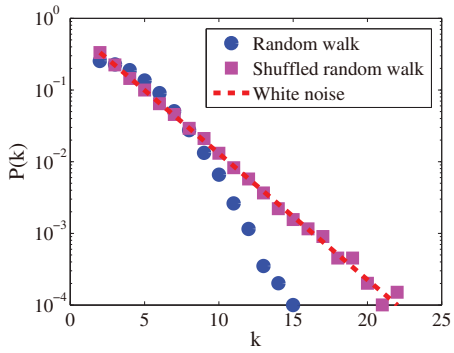

Fig. 4: (Colour online) The degree distribution of the HVG of a random walk process: $x_t = x_{t-1} + u_t$, where u_t is a Gaussian white-noise series. There exist correlations between neighboring values for the random walk process. We also give the degree distribution of the shuffled (*i.e.*, the structure has been destroyed), as well as the analytic degree distribution of white noise for comparison.

to many more nodes (thus have large degree) when their nearest neighbors have small magnitude.

Therefore, the degree of a node is generally attributed to two factors: the magnitude of the data itself, and the magnitude of its neighboring values [36]. To be more specific, a large magnitude of the data and small magnitude of its neighboring values would roughly render a large degree of the corresponding node. Unlike for the white-noise series, there exists temporal structure between neighboring values for short-term correlated or long-term correlated time series, hence disturbing the temporal structure would alter the node degrees as well as the degree distribution. In brief, randomizing the original time series changes the degree distribution, although it would not change the distribution of the original data (see fig. 4). The node degrees contain temporal information within a time series, and therefore, the MHVGCA defined on node degrees also contain such type of information. By contrast, some measures are likely to lose this information, *e.g.*, the mutual information.

Briefly, the MHVGCA includes 3 types of information: (i) the multiscale information; (ii) the information of temporal structure within time series; and (iii) the information on statistical associations between time series. Only two time series of complex systems that present overall associations on multiple scales, have similar temporal structures within each single time series, and present continuous and time-invariant interdependence structures between two time series, can indicate the complete synchronization between systems.

Significance test. In the MHVGCA, the association measure $G(s)$ ranges between -1 and 1. A value of -1 indicates perfect inversion, 1 indicates perfect agreement, and 0 indicates the absence of association. In real-world applications, it is often the case that $G(s)$ takes intermediate values rather than these three critical values. Thus,

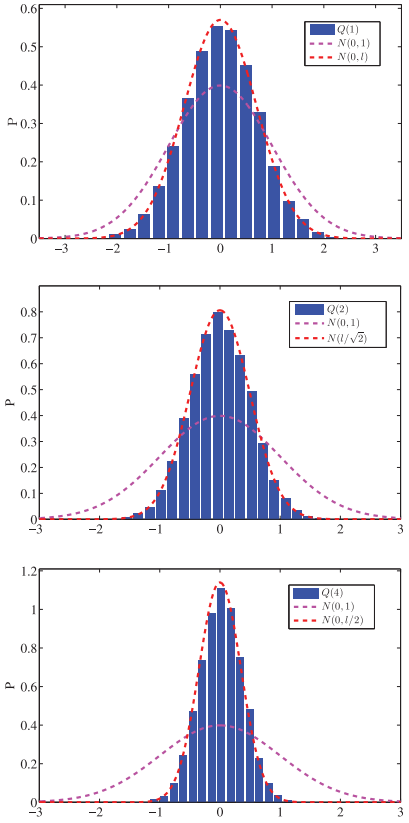

Fig. 5: (Colour online) We show estimates of the empirical distributions of $Q(s)$ for Gaussian white noise at scales $s = 1, 2, 4$ *vs.* Gaussian distributions with different variances. Here $l \approx 0.7$. For other scales, we obtain similar conclusions.

it is crucial to clarify the significance of the association between time series based on a hypothesis test.

On the one hand, the multiscale analysis does not alter the randomness within the time series, *i.e.*, the coarse-grained white-noise series are still white noise. On the other hand, the degree sequence for white noise, although it has anti-persistent temporal structure, does not affect much the association between time series [37]. In view of these, we construct a statistic for the significance test, to determine whether the association measure $G(s)$ significantly deviates from 0 or not, which simply consists of 4 steps:

(i) First, state the null and alternative hypotheses:

$$H_0 : G(s) = 0 \Leftrightarrow H_1 : G(s) \neq 0. \tag{7}$$

(ii) Next, construct a relevant test statistic Q. The critical values for the statistic Q can be estimated by using an approximation Student's t-distribution:

$$Q(s) \approx \frac{G(s)}{\sqrt{1 - G(s)^2}} \sqrt{\frac{N_s + N_d}{N}} \sim t(N_s + N_d - 2). \tag{8}$$

The statistic Q is generally used to test whether $G(s)$ is significantly different from 0 [33]. Student's t-distribution

can be approximated by the standard Gaussian distribution for large sample size. However, due to the anti-persistent temporal structure within the degree sequence, the variance of $Q(s)$ at scale 1 becomes a bit smaller, *i.e.*, $Q(1) \sim N(0, l)$, $0 < l < 1$. For large scales s, we approximately obtain $Q(s) \sim N(0, l/\sqrt{s})$, since the variance of the coarse-grained time series becomes smaller by the averaging process (see fig. 5). Here we derive from empirical estimates that $l \approx 0.7$. When there is temporal correlation within degree sequences, the variance of $Q(s)$ would be affected by this memory effects. A similar phenomenon has also been observed in ref. [37]. Conceptually, $G(s)$ is an association measure between the degree sequences, and if there were correlations within these sequences, they would affect the distribution under the no-dependence case by variance inflation. The observed anti-persistent behavior of the degree sequence can rule out the presence of relevant serial correlations in the underlying time series.

(iii) Then, select the significance level α (usually 0.05), a probability threshold below which the null hypothesis will be rejected, and obtain the critical values $-t_{\alpha/2}$ and $t_{\alpha/2}$, as well as the rejection region $(-\infty, -t_{\alpha/2}) \bigcup (t_{\alpha/2}, +\infty)$.

(iv) Finally, compute $Q(s)$ from the sample observations. Decide to either reject the null hypothesis in favor of the alternative or not reject it. The decision rule is to reject the null hypothesis H_0 if the observed value $Q(s)$ is in the rejection region, and to accept the hypothesis otherwise.

To be noted, the significance test is performed for the whole function $G(s)$. On multiple time scales s (usually only ranging from 1 to 20), the significance tests are not independently carried on, which can avoid the multiple testing problem in case of evaluating Q(s) independently and individually for a large number of s.

Synthetic time series analysis. We first consider two coupled VAR processes [38]:

$$\begin{cases} x_t = 0.4x_{t-1} + \lambda y_{t-1} + u_t, \\ y_t = 0.7y_{t-1} + v_t, \end{cases} \tag{9}$$

where $\lambda \in [0,\ 1]$ regulates the coupling strength. u_t and v_t are both Gaussian white-noise series with zero mean and unit variance. We employ MHVGCA under different coupling strengthes to observe the associations between two VAR processes (see fig. 6).

There exists a trade-off between the strength of real coupling of two processes and the variance of the white noise; the relationship between them can be considered as the signal-to-noise ratio, which compares the level of a desired signal to the level of background noise. When the real coupling between two processes dominates, the MHVGCA is likely to present accurate estimation of associations as expected, otherwise when the white noise dominates, spurious relationship estimated by the method may arise. The multiscale analysis decreases the variance of the white-noise processes to $1/s$ according to the properties of

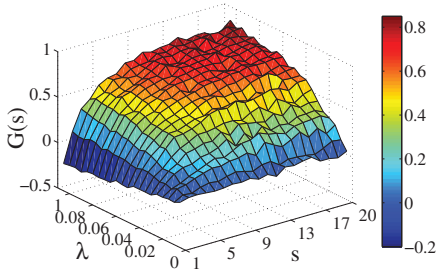

Fig. 6: (Colour online) The association measure $G(s)$ of MHVGCA for the coupled VAR processes. The length of each series is $N = 10^4$. We change the coupling strength λ to observe the values $G(s)$ of the new method.

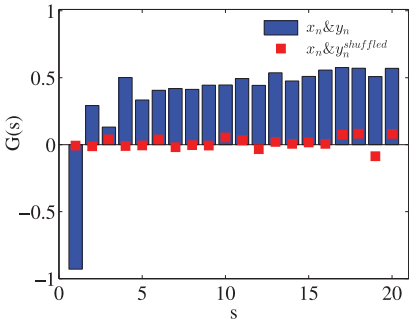

Fig. 7: (Colour online) The association measure $G(s)$ of MHVGCA for the classical Hénon map (blue bar), and the $G(s)$ for x_n& $y_n^{shuffled}$ for comparison(red square).

random walk processes. Further considering the signal-to-noise ratio, the real couplings are retained during the coarse-graining process, while the noise tends to be weakened. Therefore, the signal-to-noise ratio may become larger on larger scales for the VAR processes (for more detailed explanations, see our recent work [39]). Therefore, at small scales, the white noise dominates, thus the real coupling between VAR processes cannot be detected by the MHVGCA ($G(s) \approx 0$). At large scales, the real coupling dominates, and therefore, the $G(s)$ of the MHVGCA increases. Further, the $G(s)$ significantly increases with coupling strength increasing λ, as expected.

As our second example, we study the Hénon map,

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n, \\ y_{n+1} = bx_n. \end{cases} \tag{10}$$

The map depends on two parameters, a and b, which for the classical Hénon map have values of $a = 1.4$ and $b = 0.3$. There exists nonlinear information exchange between x_{n+1} and y_{n+1}, since x_{n+1} comprises $-x_n^2$ and y_{n+1} comprises x_n.

In fig. 7, we estimate the associations between time series on multiple time scales by the MHVGCA. On the scale 1, $G(1)$ is negative, while on larger scales, $G(s)$ becomes positive. Analysis on the unique scale would lead to information loss on multiple scales. Moreover, we shuffle one time series, *e.g.*, y_n, and estimate the interactions between

Table 2: The 9 global stock markets.

Index	# of Points	Country	Region
SSE	4843	China	Asia
SZSE	4843	China	Asia
N225	4914	Japan	Asia
DJI	5042	U.S.	America
NASDAQ	5034	U.S.	America
SP500	5034	U.S.	America
DAX	5075	Germany	Europe
CAC	5101	France	Europe
FTSE100	5053	UK	Europe

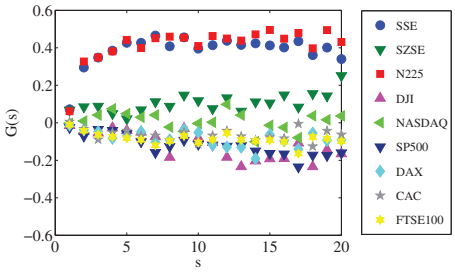

Fig. 8: (Colour online) The association measure $G(s)$ over the scale s between the daily closing price and daily trading volume in 9 stock markets.

x_n and $y_n^{shuffled}$. All the $G(s)$ values fluctuate around zero, which indicates the accuracy of the MHVGCA.

Stock time series analysis. – In this section, we study stock time series using MHVGCA. We mainly study the indices of 9 stock markets, *i.e.*, SSE, SZSE, N225, DJI, NASDAQ, SP500, DAX, CAC, and FTSE100, 3 located in Asia, 3 located in America, and 3 located in Europe (see table 2). The data points, including the daily closing price and daily trading volume, span completely 2 decades from 1/2/1997 to 12/30/2016.

Many researchers studied the logarithmic return rather than the daily closing price directly, due to: (i) The daily closing price series is non-stationary, since the data change significantly with time, and the fluctuation of the data tends to be proportional to the magnitude of the data. (ii) Many tools are merely applicable to analyze stationary time series. Here the MHVGCA is able to deal with non-stationary time series, and therefore, no transformation of the original data is required.

First, we obtain the degree distributions of the HVGs for the 9 stock daily closing price series, respectively. Although these 9 series evolve significantly different in the time domain, their degree distributions seem rather close to each other (not shown here). As indicated by Luque *et al.* [17], while the time series is defined in the time domain and the discrete Fourier transform is defined on the frequency domain, the (H)VG is defined on the "visibility domain". That is to say, the 9 stock closing price series are likely to have many common properties in the "visibility domain". At small node degrees k, all the closing price series show no significant difference from a white-noise series. But when the node degree $k \geq 9$, the closing price series begin to deviate from the white noise. Interestingly, this phenomenon is also observed in random walk processes, where the degree distribution deviating from white noise also begins at $k = 9$ (see fig. 4). This effect is partly due to the non-stationarity of both types of series. But not all non-stationary time series would have a similar degree distribution of random walk processes, *e.g.*, the white-noise, series added with a persistent increasing trend. Moreover, it is irrelevant for the data length.

We also study the degree distributions of the HVGs for the 9 stock daily trading volume series, respectively. Different from the daily closing price, all these degree distributions are rather close to that of a white-noise series. Thus, we derive that the degree distribution of the daily trading volume series can be approximated by the exponential distribution. Although the daily stock price and daily trading volume have a close long-term relationship [7], these two variables behave significantly different in their HVGs primarily addressing short time scales.

In fig. 8, we demonstrate the associations between the daily closing price and daily trading volume on multiple time scales by the MHVGCA in 9 stock markets. Three Asian stock markets perform significantly different from other markets, especially for the SSE and N225. For the SSE and N225, the associations between the two variables gradually increase with increasing scale, from 0.3 to roughly 0.5. This indicates the persistent information exchange between the daily closing price and daily trading volume. The information of the trading volume can thus be used for the prediction of the trading price in Asian markets. For the American stock markets, the DJI, NASDAQ, and SP500 all show negative relationship between the daily closing price and daily trading volume on large scales. The European stock markets are in an intermediate state between the Asian markets and American markets, where the associations between the daily closing price and daily trading volume are hardly significant. This is consistent with our results that the daily closing price can be approximated by random walk processes, while the daily trading volume is rather close to white noise.

Conclusion. – In this letter, we propose the MHVGCA to quantify interactions between non-stationary time series on multiple time scales by using HVGs. We compare the new method with other measures, and study the degree properties and significance test. We then apply the MHVGCA to 9 stock markets, and measure the interactions between the daily closing price and daily trading volume.

Our results indicate that in developed markets like the American and European stock markets, there exists little information exchange between the daily closing price and the daily trading volume at small temporal scales. An exception are the Asian markets, where the daily closing price and daily trading volume show interactions

on multiple time scales. Things change rapidly in stock markets, and the markets situation can change a lot at inter-annual scales. Thus, the MHVGCA applied for high-frequency stock time series analysis can be further explored. In addition, it is natural to quantify the similarity between degree sequences by some other measure that also accounts for ties in the values, since degree values are by definition discrete and have a possibly quite limited range, and ties are very frequent. Goodman and Kruskal's gamma does not take ties into account, while other extensions of Kendall's τ that do so may be considered in the future.

* * *

The authors would like to thank two anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this manuscript. The financial support by National Natural Science Foundation of China (61703029, 61671048), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (16YJC910007), Beijing Outstanding Talent Project (B16H200060), and Beijing Social Science (17YJC024) is gratefully acknowledged.

REFERENCES

[1] Thomas R. J., Mietus J. E., Peng C. K. and Goldberger A. L., *Sleep*, **28** (2005) 1151.
[2] Zhao X., Shang P., Lin A. and Chen G., *Physica A*, **390** (2011) 3670.
[3] Zhao X., Shang P. and Huang J., *EPL*, **102** (2013) 40005.
[4] Zhao X., Shang P. and Wang J., *Phys. Rev. E*, **87** (2013) 022805.
[5] Zhao X., Shang P. and Shi W., *Physica A*, **402** (2014) 84.
[6] Hao X., Zuo M. and Liu L., *Appl. Math. Lett.*, **82** (2018) 24.
[7] Podobnik B., Horvatic D., Petersen A. M. and Stanley H. E., *Proc. Natl. Acad. Sci. U.S.A.*, **106** (2009) 22079.
[8] Donges J. F., Schultz H. C. H., Marwan N., Zou Y. and Kurths J., *Eur. Phys. J. B*, **84** (2011) 635.
[9] Yan J., Zhang S. and Cai J., *Discrete Math.*, **341** (2018) 1160.
[10] Marwan N., Donges J. F., Zou Y., Donner R. V. and Kurths J., *Phys. Lett. A*, **373** (2009) 4246.
[11] Donner R. V., Zou Y., Donges J. F., Marwan N. and Kurths J., *New J. Phys.*, **12** (2010) 129.
[12] Donges J. F., Heitzig J., Donner R. V. and Kurths J., *Phys. Rev. E*, **85** (2012) 046105.
[13] Lacasa L., Luque B. and Ballesteros F., *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008) 4972.
[14] Lacasa L., Luque B., Luque J. and Nuno J. C., *EPL*, **86** (2009) 30001.
[15] Telesca L. and Lovallo M., *EPL*, **97** (2012) 50002.
[16] Donner R. V. and Donges J. F., *Acta Geophys.*, **60** (2012) 589.
[17] Luque B., Lacasa L., Ballesteros F. and Luque J., *Phys. Rev. E*, **80** (2010) 046103.
[18] Xie W.-J. and Zhou W.-X., *Physica A*, **390** (2011) 3592.
[19] Kloupis G., *Commun. Nonlinear Sci. Num. Simul.*, **51** (2017) 13.
[20] Santo F., *Phys. Rep.*, **486** (2009) 75.
[21] Goswami B., Boers N., Rheinwalt A., Marwan N., Heitzig J., Breitenbach S. M. F. and Kurths J., *Nat. Commun.*, **9** (2018) 48.
[22] Donges J. F., Donner R. V. and Kurths J., *EPL*, **102** (2013) 381.
[23] Takens F., *Dynamical Systems & Turbulence* (Springer) 1981.
[24] Mandelbrot B. B. and Wheeler J. A., *J. R. Stat. Soc.*, **147** (1983) 468.
[25] Peng C. K., Buldyrev S. V., Havlin S., Simons M., Stanley H. E. and Goldberger A. L., *Phys. Rev. E*, **49** (1994) 1685.
[26] Costa M., Goldberger A. L. and Peng C. K., *Phys. Rev. Lett.*, **89** (2002) 068102.
[27] Gao Z., Cai Q., Yang Y., Dang W. and Zhang S., *Sci. Rep.*, **6** (2016) 35622.
[28] Gao Z., Yang Y., Fang P., Zou Y., Xia C. and Du M., *EPL*, **109** (2015) 30005.
[29] Zhang Y., Shang P., Xiong H. and Xia J., *Fluct. Noise Lett.*, **17** (2018) 1850006.
[30] Ahmadlou M. and Adeli H., *Phys. D: Nonlinear Phenom.*, **241** (2012) 326.
[31] Zou Y., Donner R. V., Marwan N., Small M. and Kurths J., *Nonlinear Process. Geophys. Discuss.*, **1** (2014) 1113.
[32] Lacasa L., Nicosia V. and Latora V., *Sci. Rep.*, **5** (2014) 15508.
[33] Goodman L. A. and Kruskal W. H., *J. Am. Stat. Assoc.*, **49** (1954) 732.
[34] Shen T., Xin J. and Huang J., *Stoch. Anal. Appl.*, **36** (2018) 103.
[35] Xie W., Han R., Jiang Z., Wei L. and Zhou W., *EPL*, **119** (2017) 48008.
[36] Elsner J. B., Jagger T. H. and Fogarty E. A., *Geophys. Res. Lett.*, **36** (2009) 554.
[37] Zhao X., Shang P. and Lin A., *EPL*, **107** (2014) 40008.
[38] Zhao X., Shang P. and Wang J., *Nonlinear Dyn.*, **78** (2014) 1149.
[39] Zhao X., Sun Y., Li X. and Shang P., *Commun. Nonlinear Sci. Numer. Simul.*, **62** (2018) 202.