TOPICAL REVIEW — Statistical physics and complex systems

A mini-review on econophysics: Comparative study of Chinese and western financial markets*

Zheng Bo(郑 波)[a)†], Jiang Xiong-Fei(蒋雄飞)[a)b)], and Ni Peng-Yun(倪鹏云)[a)]

[a)] *Department of Physics, Zhejiang University, Hangzhou* 310027, *China*

[b)] *College of Information Engineering, Ningbo Dahongying University, Ningbo* 315175, *China*

(Received 14 March 2014; revised manuscript received 17 April 2014; published online 22 May 2014)

We present a review of our recent research in econophysics, and focus on the comparative study of Chinese and western financial markets. By virtue of concepts and methods in statistical physics, we investigate the time correlations and spatial structure of financial markets based on empirical high-frequency data. We discover that the Chinese stock market shares common basic properties with the western stock markets, such as the fat-tail probability distribution of price returns, the long-range auto-correlation of volatilities, and the persistence probability of volatilities, while it exhibits very different higher-order time correlations of price returns and volatilities, spatial correlations of individual stock prices, and large-fluctuation dynamic behaviors. Furthermore, multi-agent-based models are developed to simulate the microscopic interaction and dynamic evolution of the stock markets.

Keywords: complex systems, econophysics

PACS: 89.75.–k, 89.65.Gh **DOI:** 10.1088/1674-1056/23/7/078903

1. Introduction

In recent years, there has been a growing interest of physicists in complex economic and social systems. Financial markets are such important and representative examples with many-body interactions, somewhat similar to traditional physical systems. With a large amount of historical data piled up in the past years, it becomes possible to analyze the collective behavior of financial markets based on concepts and methods in statistical physics.

Some stylized facts such as the 'fat tail' in the probability distribution of price returns[1,2] are discovered in the empirical study. The long-range auto-correlation of volatilities, which is also known as volatility clustering, is quantified by computing the auto-correlation function.[3,4] The so-called leverage effect can be quantitatively studied, through the return-volatility correlation function.[5] Many researchers are devoted to the cross-correlations among individual stock prices, and out-stationary states such as financial crashes.[6–10] Meanwhile, various microscopic models and theoretical approaches have been developed, with certain degree of success, to describe the financial dynamics.[3,11–16] Very recently, experiments in laboratories have also been set up, to 'simulate' financial and economic systems.[17,18] As a modern trend in interdisciplines, econophysics not only generates a subfield in physics, but also leads to quantitative study of financial systems with advanced techniques.

To some extent, the collective behavior of the financial dynamics is rather robust, independent of particular stock markets, at least within the mature markets in western countries. On the other hand, it is also known that the emerging markets may behave differently. Especially, the Chinese stock market was newly set up in 1990 and shares a transiting economic and social system. Up to date, however, the dynamic behavior of the emerging markets is much less concerned, compared with that of the mature markets. In the past decade, one research direction in our group is in econophysics. Especially, we focus on the comparative study of Chinese and western stock markets, and explore the universal and non-universal dynamic behavior of the Chinese stock market, including the corresponding microscopic interactions and mechanisms. Many other authors have also contributed to this topic (see Ref. [19] and references therein), but in this mini-review article we mainly review the recent results in our group. Therefore, we apologize that the references are limited to very basic and relevant ones.

From the view of statistical physics, one may compute the time and spatial correlation functions for understanding the stationary state of a dynamic system. These correlation functions could be derived with phenomenological methods and/or from fundamental microscopic interactions. At this stage, it is natural that only the dynamic properties averaged over time or individual stocks or many events are practically concerned. In econophysics, economic and financial significance should be also taken into account. In this article, we will proceed in such a line. In Sections 2 and 3, we present phenomenologi-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11375149, 11075137, and J1210046) and the Natural Science Foundation of Zhejiang Province of China (Grant No. Z6090130).

†Corresponding author. E-mail: zheng@zimp.zju.edu.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd

cal results for the time and spatial correlations. In Section 4, we develop a representative multi-agent-based model. In Section 5, we study the non-stationary dynamic relaxation of large volatilities. In Section 6, the article closes with the conclusion.

2. Time correlations

In this section, we investigate the time correlations of price returns and volatilities, with the daily and minute-to-minute data of the Chinese indices and German DAX. We collect the daily data of the German DAX from 1959 to 2009, and the minute-to-minute data from 1993 to 1997. The daily data of the Shanghai Index are from 1990 to 2009, and the minute-to-minute data are from 1998 to 2006. The daily data of the Shenzhen Index are from 1991 to 2009, and the minute-to-minute data are from 1998 to 2003. The minute-to-minute data are recorded every minute in the German DAX, while every 5 minutes in the Chinese indices. A working day is about 450 minutes in Germany while exactly 240 minutes in China.[20]

We denote the price of a financial index at time t' as $P(t')$, then its logarithmic price return is $R(t') \equiv \ln P(t'+1) - \ln P(t')$. The absolute value of the price return, i.e., $|R(t')|$, is usually defined as the volatility. For comparison of different financial indices, we introduce the normalized price return as

$$r(t') \equiv [R(t') - \langle R(t') \rangle]/\sigma, \tag{1}$$

where $\langle \cdots \rangle$ represents the average over time t', and $\sigma = \sqrt{\langle R^2 \rangle - \langle R \rangle^2}$ is the standard deviation of $R(t')$. An important basic property of the financial dynamics is the probability distribution of price returns, denoted as $P(r)$. The $P(r)$ can also be separated into $P_+(r)$ and $P_-(r)$ for positive and negative price returns. It has been well known that $P(r)$ obeys a Lèvy distribution in the central regime, but usually with fat tails[2,4,20,21]

$$P_\pm(r) \sim |r|^{-\mu}. \tag{2}$$

Figure 1(a) shows the probability distribution of the daily data for both the Shanghai Index and German Dax. The $P_+(r)$ and $P_-(r)$ are symmetric, and the Shanghai Index and German Dax share the same exponent $\mu \approx 3.8$ for the tails.[20,21] Similar behaviors are obtained for the minute-to-minute data.[21] From the price returns of individual stocks, one may derive more accurate results. In some cases, the exponent μ may be different for $P_+(r)$ and $P_-(r)$.[20,22]

From the non-Gaussian fat-tail distribution of price returns, one may believe that the financial dynamics is not a random Gaussian process, rather with time and spatial correlations. Direct calculations show that the auto-correlation function of price returns decays exponentially with a correlating time of the order of minutes, but volatilities are long-range correlated in time, at least up to some months.[2–4] Let us define the auto-correlation function of volatilities as

$$A(t) = [\langle |r(t')||r(t+t')| \rangle - \langle |r(t')| \rangle^2]/A_0, \tag{3}$$

where $\langle \cdots \rangle$ is the average over time t', and $A_0 = \langle |r(t')|^2 \rangle - \langle |r(t')| \rangle^2$. Figure 1(b) displays $A(t)$ computed with the daily data of the Shanghai Index and German Dax.[20,21] In about two orders of magnitude, $A(t)$ exhibits a power-law behavior

$$A(t) \sim t^{-\beta}, \tag{4}$$

where $\beta = 0.32(3)$ and $0.32(2)$ for the Shanghai Index and German Dax, respectively. Similar results are achieved with the minute-to-minute data, after removing the intra-day pattern.[21] The long-range auto-correlation of volatilities is well known as the volatility clustering in economics and finance. In past years, many activities in different subjects including econophysics have been devoted to this phenomenon, including various microscopic modelings.[3,14,15,23–25]

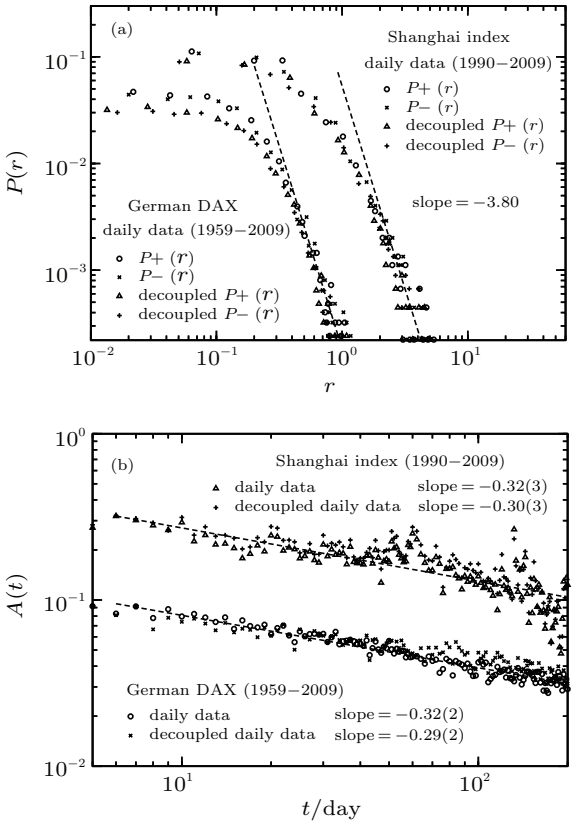


Fig. 1. (a) Probability distributions $P_\pm(r)$ of positive and negative returns are displayed for the daily data of the Shanghai Index and German DAX. $P_\pm(r)$ remain almost unchanged after the decoupling interactions in Eq. (6) are introduced. For clarity, $P_\pm(r)$ of the German DAX have been shifted to the left by 4.5 units. (b) The auto-correlation function of volatilities is displayed for the daily data of both the Shanghai Index and German DAX. The long-range auto-correlation remains after the decoupling interactions are introduced.[20]

For example, we could further explore the long-range auto-correlation of volatilities nonlocal in time, by introducing the persistence probability of volatilities.[20,26,27] The persistence probability of volatilities $P_+(t)$ (or $P_-(t)$) is defined as

the probability that $|r(t'+\widetilde{t})|$ has always been above (or below) $|r(t')|$ in time t, i.e., $|r(t'+\widetilde{t})| > |r(t')|$ (or $|r(t'+\widetilde{t})| < |r(t')|$) for all $\widetilde{t} < t$. The average is taken over t'. In general, $P_{-}(t)$ obeys a universal power-law behavior: $P_{-}(t) \sim t^{-\theta_{\rm p}}$.[26,27] In Fig. 3(a) of Ref. [20], the persistence probability $P_{-}(t)$ of volatilities is plotted for the daily data of the Shanghai Index and German DAX. The $P_{-}(t)$ exhibits a nice power-law behavior, with $\theta_{\rm p} = 0.84(2)$ and $0.87(3)$ for the Shanghai Index and German DAX, respectively. Within statistical errors, the exponent $\theta_{\rm p}$ is not different for the two stock markets. The fact $0.5 < \theta_{\rm p} < 1$ indicates that the volatility is long-range correlated in time.

Up to date, however, our knowledge on the dynamic behavior of the price return itself is still limited. Since the auto-correlation of price returns is extremely weak,[2,4] nonzero higher-order time correlations relevant for price returns become very important, especially the lowest-order one among them. In financial markets, this lowest-order nonzero correlation turns out to be the return-volatility correlation. In 1976, a negative return-volatility correlation is first discovered by Black,[28] although in somewhat different words. This is the so-called leverage effect, which implies that the past negative price returns increase future volatilities. To the best of our knowledge, the leverage effect exists in almost all stock markets in the world. In the Chinese stock market, however, a positive return-volatility correlation is detected, which is called the anti-leverage effect.[20,21,29]

The return-volatility correlation function is usually defined as

$$L(t) = (\langle r(t')|r(t'+t)|^2\rangle - \langle r(t')\rangle\langle |r(t')|^2\rangle)/Z,$$
$$Z = \langle |r(t')|^2\rangle^2, \tag{5}$$

where $\langle\cdots\rangle$ is the average over time t'. Replacing $|r(t'+t)|^2$ by $|r(t'+t)|$ leads to similar results. In Fig. 2, $-L(t)$ computed with the daily data of the Shanghai Index, Shenzhen Index, and German Dax is plotted.[20,21,29] For $t > 0$, we observe a negative $L(t)$, i.e., a leverage effect, for the German DAX, while a positive $L(t)$ for both the Shanghai Index and Shenzhen Index, i.e., 'an anti-leverage effect'. Fitting the data to an exponential form $L(t) = c\exp(-t/\tau)$, we obtain $\tau = 15$ and 7 days for the leverage and anti-leverage effects, respectively. Compared with the negligibly small correlating time of price returns (on the order of minute),[2,4] both the leverage and anti-leverage effects are rather prominent. Additionally, careful analysis shows that large volatilities seem to dominate the leverage and anti-leverage effects.[20] For $t < 0$, $L(t)$ fluctuates around zero for both the Chinese and German markets. This implies that $r(t')$ is very weakly correlated to volatilities in the past times, at least in the sense of local time correlations. Whether and how volatilities may affect the price movement remains open.[30]

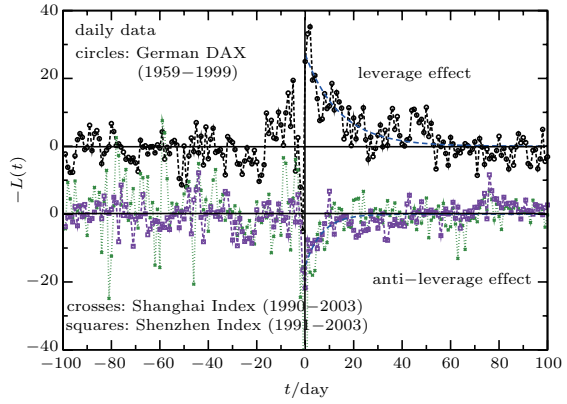


Fig. 2. $-L(t)$ computed with the daily data. Dashed lines show an exponential fit $L(t) = c\exp(-t/\tau)$.[29]

Usually, the leverage effect is considered to be a daily phenomenon that is only computed with the daily data.[5] To further confirm and support our findings, however, we also analyze the minute-to-minute data. As shown in Fig. 1(b) of Ref. [29], $L(t)$ computed with the minute-to-minute data contains high-frequency fluctuations. To reveal the dynamic behavior of the slow mode, we average the data points in time windows of four days. Then the leverage and anti-leverage effects emerge, in good agreement with those of the daily data.

Why do the Chinese and German markets exhibit different return-volatility correlations? Germany is a developed country, and the stock market is mature. Investors show risk aversion, and therefore may be nervous of trading as the stock price is falling. This induces a higher volatility. When the stock price rises, investors feel pleased and are conservative in trading. Thus, the stock market tends to be calm. This should be the social origin of the leverage effect. However, China has just experienced the first stage of capitalism, the stock market is emerging, and investors are somewhat excessively speculative in investing. Therefore, investors rush for trading as the stock price increases. When the price drops, investors remain inactive in trading and wait for raise of the stock price. This explains the anti-leverage effect.

More academically, what kinds of mechanisms or interactions lead to the leverage and anti-leverage effects? It has been argued that both the long-range auto-correlation of volatilities and asymmetric probability distribution of price returns are necessary for a leverage effect.[22,31] By the definition of the return-volatility correlation function, the long-range auto-correlation of volatilities and asymmetric probability distribution of returns together may indeed induce or alter the leverage effect or anti-leverage effect. In Ref. [20], however, we show that such an argument is not dominating in stock markets. We propose that a retarded volatility model may provide a phenomenological origination of the leverage and anti-leverage effects.

For this purpose, we reformulate the standard retarded

volatility model as[20]

$$r_0(t') = \left[1 + \sum_{t=1}^{\infty} K(t) r(t'-t)\right] r(t'), \tag{6}$$

where $r(t')$ is the original price return of a real financial market, and $r_0(t')$ is the decoupled one. The decoupling interaction $K(t)$ should be properly chosen, for example, $K(t) = -CL(t)$, as suggested in Ref. [20]. Our finding is that the leverage or anti-leverage effect can be eliminated by adjusting the constant C appropriately. In Fig. 1 of Ref. [20], $L(t)$ calculated from the decoupled return $r_0(t')$ is shown. Obviously, it fluctuates around zero for the daily and minute-to-minute data of both the German DAX and Chinese indices. Reversely, assuming that the original price return $r(t')$ is without a nonzero return-volatility correlation, the leverage or anti-leverage effect may be generated by the retarded volatility model.

More importantly, the decoupling interaction in Eq. (6) does not change other characteristics of the time series $r(t')$, since $K(t)$ is only a perturbation, in the sense that $K(t)|r(t'-t)| \ll 1$ and $\sum_{t=1}^{\infty} K(t) r(t'-t) \ll 1$. In Fig. 1(a), the probability distributions $P_{\pm}(r)$ of positive and negative price returns are displayed for both the original daily returns $r(t')$ and decoupled daily returns $r_0(t')$ of the Shanghai Index and German DAX. Clearly, the decoupling interaction $K(t)$ does not modify either the power-law tails or the shapes of $P_{\pm}(r)$. Similar results are also observed for the minute-to-minute data. From Fig. 1(b) in this article and Fig. 3(a) in Ref. [20], the same conclusion can be drawn for the long-range auto-correlation of volatilities, and the persistence probability of volatilities.

Finally, an important observation is that the Chinese stock market is undergoing a transition from an emerging market to a mature one.[32] Careful analysis shows that before the year 2000, the Chinese market exhibited a strong anti-leverage effect; in some years after 2000, the return-volatility correlation was weak; in recent years, it has gradually changed to the leverage effect.

3. Spatial structure

Besides the time correlations, it is an important topic to explore the spatial structure in financial markets. In this section, the so-called spatial structure does only refer to the interaction structure among individual stocks, rather than the geographical structure. Especially, unlike most traditional physical systems, where one derives spatial correlations between subunits from their interactions, the underlying "interactions" for the stock markets are not yet known. At the first stage, therefore, one needs phenomenological methods. For example, the hierarchical structure of stock markets has been investigated through the minimal spanning tree method and its variants.[10,33,34] With the random matrix theory (RMT), business sectors may be identified,[6,7,35] for example, for mature markets such as the New York Stock Exchange (NYSE) and the Korean Stock Exchanges, and also for emerging markets such as the National Stock Exchange (NSE) in India. In particular, we have investigated the spatial structure of the Chinese stock market based on the RMT method and combination of the network techniques.[32,36–38] As an important emerging market, the Chinese stock market exhibits much stronger cross-correlations than the mature ones, and the structure of the standard business sectors is weak. Instead, unusual sectors such as special treatment (ST) and blue-chip sectors are detected. Moreover, we observe that a business sector may split into two subsectors which are anti-correlated with each other within this eigenmode, and such a splitting phenomenon is rather prominent in the Chinese market.

We have collected the daily data of 259 stocks in the Shanghai Stock Exchange (SSE) from 1997 to 2007, and in the NYSE from 1990 to 2006. Meanwhile, we have collected the daily data of 66 financial indices, including 57 indices in stock markets and 9 treasury bond rates in the US from Sep., 1997 to Oct., 2008. We name the 66 indices the global market indices (GMI).[38] To ensure different stocks with an equal weight, we introduce the normalized price return of the i-th stock as

$$r_i(t) = [R_i(t) - \langle R_i(t)\rangle]/\sigma_i, \tag{7}$$

where $R_i(t)$ is the price return, σ_i denotes the standard deviation of $R_i(t)$, and $\langle\cdots\rangle$ is the average over time t. The elements of the cross-correlation matrix $\boldsymbol{C}$ are defined by the equal-time correlations as

$$C_{ij} \equiv \langle r_i(t) r_j(t)\rangle. \tag{8}$$

By the definition, $\boldsymbol{C}$ is a real symmetric matrix with $C_{ii} = 1$, and $C_{ij} \in [-1, 1]$. The mean value $\overline{C}_{ij}$ of the non-diagonal elements for the SSE is 0.37, much larger than 0.16 and 0.26 for the NYSE and GMI, respectively. It confirms that stock prices in emerging markets are more correlated than mature ones. In fact, from Fig. 1(a) of Ref. [36], it is shown that the curve of the probability distribution $P(C_{ij})$ of the SSE is shifted greatly in the positive direction, compared with that of the NYSE and even NSE.

We now compute the eigenvalues of the cross-correlation matrix $\boldsymbol{C}$, in comparison to those of the so-called Wishart matrix, which is derived from non-correlated time series. Assuming N time series with length T, and in the large-N and large-T limit with $Q \equiv T/N \geq 1$, the probability distribution $P_{\rm rm}(\lambda)$ of the eigenvalue λ for the Wishart matrix is given by

$$P_{\rm rm}(\lambda) = \frac{Q}{2\pi\lambda}\sqrt{(\lambda_{\rm max}^{\rm ran} - \lambda)(\lambda - \lambda_{\rm min}^{\rm ran})}, \tag{9}$$

with the upper and lower bounds of

$$\lambda^{\rm ran}_{\rm min(max)} = \left[1 \pm \left(1/\sqrt{Q}\right)\right]^2. \quad (10)$$

For a dynamic system, large eigenvalues of the cross-correlation matrix, which deviate from $P_{\rm rm}(\lambda)$, imply that there exist non-random interactions. In fact, in both mature and emerging stock markets, the bulk of the eigenvalue spectrum $P(\lambda)$ of the cross-correlation matrix is similar to $P_{\rm rm}(\lambda)$ of the Wishart matrix, but some large eigenvalues deviate significantly from the upper bound $\lambda^{\rm ran}_{\rm max}$. Let us arrange the eigenvalues in the order of $\lambda_\alpha > \lambda_{\alpha+1}$. As shown in Table 1 of Ref. [38], the largest eigenvalue λ_0 of the SSE (China) is 97.3, while that of the NYSE (US) and GMI is 45.6 and 21.5, respectively.

It has been well known that the large eigenvalues deviating from the bulk correspond to different modes of motion in stock markets. The components in the eigenvector of the largest eigenvalue λ_0 are uniformly distributed. Therefore, the largest eigenvalue represents the market mode, which is driven by interactions common for stocks in the entire market. The components in the eigenvectors of other large eigenvalues are localized. A particular eigenvector is dominated by a sector of stocks, usually associated with a business sector. By $u_i(\lambda_\alpha)$, we denote the component of the i-th stock in the eigenvector of λ_α. To identify the sector, one may introduce a threshold $u_{\rm c}$, to select the dominating components in the eigenvector by $|u_i(\lambda_\alpha)| \geq u_{\rm c}$.[36,38]

According to Ref. [36], standard business sectors can hardly be detected in the SSE (China). Instead, one finds that there exist three unusual sectors, i.e., the ST, blue-chip, and SHRE sectors, corresponding to the second, third, and fourth largest eigenvalues, respectively. Since the Chinese stock market is an emerging market, the companies are not operated strictly with the registered business. On the other hand, investors seriously look at the performance of the companies and the dominating business and areas. The fourth largest eigenvalue λ_3 of the SSE may reflect the fact that Shanghai and especially its real estate business have played an important role in China in the past years. However, it is puzzling what are the dominating stocks for the eigenvectors of other large eigenvalues.

Very recently, we observe that the components in an eigenvector may carry positive and negative signs, and thus a sector may split into two subsectors.[38] Practically, we introduce two thresholds $u^{\pm}_{\rm c} = \pm u_{\rm c}$: $u_i(\lambda_\alpha) \geq u^{+}_{\rm c}$ and $u_i(\lambda_\alpha) \leq u^{-}_{\rm c}$, to identify the positive and negative subsectors respectively. The results of the SSE are shown in Table 1. With this technique, we may explore the subsector structure in the SSE up to λ_6. A number of standard business subsectors such as the high technology and finance are also observed. However, the SSE is indeed dominated by unusual sectors and subsectors such as the ST, blue-chip, traditional industry, SHRE, weak and strong cyclical industry. For comparison, the subsectors of the NYSE (US) are listed in Table 3 of Ref. [38]. The subsector structure in the NYSE is somewhat different from that in the SSE. Firstly, most subsectors are the standard business subsectors. For more than half of λ_i, only one dominating subsector remains for sufficiently large thresholds $u^{\pm}_{\rm c}$. In other words, the splitting phenomenon is less prominent in the NYSE.

What is the physical meaning of the positive and negative subsectors? The cross-correlation between two stocks can be decomposed into different eigenmodes

$$C_{ij} = \sum_{\alpha=1}^{N} \lambda_\alpha C^\alpha_{ij}, \;\; C^\alpha_{ij} = u^\alpha_i u^\alpha_j, \quad (11)$$

where $u^\alpha_i \equiv u_i(\lambda_\alpha)$, and C^α_{ij} represents the cross-correlation in the α-th eigenmode. Since the eigenvalue λ_α is positive, it gives the weight to the α-th eigenmode, and the sign of C^α_{ij} is essential in the sum. The C^α_{ij} is positive if the components u^α_i and u^α_j have the same sign in a particular eigenmode. Otherwise, it is negative. When C^α_{ij} is negative, two stocks are referred to as anti-correlated in this eigenmode. Therefore, all stocks in the same subsector are positively correlated in this eigenmode, while the stocks in different subsectors are anti-correlated. This is the physical meaning of the positive and negative subsectors. For example, the sector of λ_2 in the SSE is composed of the traditional industry and high technology subsectors. The former represents those traditional industry companies with a long-term and stable interest, but a lower asset risk and expected revenue, while the latter includes the high technology companies with novel business and conceptions, but a higher asset risk and potential profit. In the sense of investing, these two subsectors are anti-correlated in the eigenmode of λ_2.

Table 1. Subsectors in the SSE. The fraction is the number of well identified stocks over the total number of stocks in the subsector. Null: no obvious category; ST: specially treated; Trad: traditional industry; Tech: high technology; SHRE: Shanghai real estate; Weak: weakly cyclical industry; Stro: strongly cyclical industry; Fin: finance; and IG: industrial goods.[38]

	λ_1		λ_2		λ_3		λ_4		λ_5		λ_6	
	+	−	+	−	+	−	+	−	+	−	+	−
Sector	Null	ST	Trad	Tech	ST	SHRE	Weak	Stro	Fin	Null	IG	Null
$u^{\pm}_{\rm c} = \pm 0.08$	26	31/35	22/23	23/25	24/27	27/27	23/26	24/26	14/18	25	15/17	25
$u^{\pm}_{\rm c} = \pm 0.10$	7	20/23	16/17	12/13	11/12	20/20	13/15	15/16	10/14	17	8/9	18

To verify the anti-correlation in an eigenmode more intuitively, we may compute the average $\overline{C}_{ij}$ within the positive or negative subsector, and between the positive and negative subsectors. The results for the NYSE are shown in Fig. 3, and those for the SSE look similar. The $\overline{C}_{ij}$ within the positive or negative subsector is obviously larger than that between the positive and negative subsectors, especially for small α, i.e., large eigenvalues. This strongly suggests that there indeed exists an anti-correlation between the positive and negative subsectors. Very recently, the analysis of the subsector structure is extended to the four stock markets in Greater China, and different characteristics and their economic indications are revealed.[32]

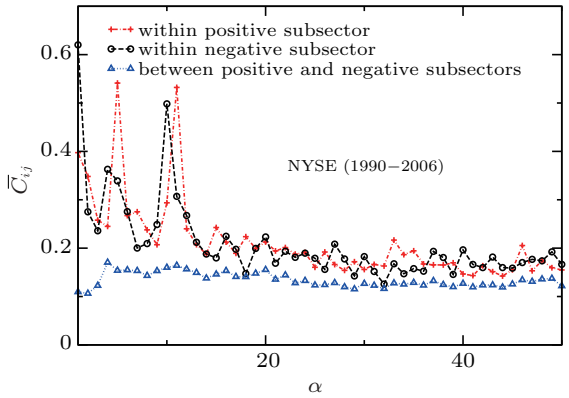


Fig. 3. The average cross-correlation $\overline{C}_{ij}$ for the NYSE.[38]

A further investigation step in this direction would be to develop a cross-correlation decomposition method. Since the eigenvectors of the cross-correlation matrix $\boldsymbol{C}$ form a complete set of bases, the price motion in a financial market can be generally decomposed into three kinds of eigenmodes: the market mode is described by the eigenmode of the largest eigenvalue λ_0; the sector modes include the eigenmodes of other large eigenvalues; and the quasi-random modes are the eigenmodes of all the eigenvalues within the upper bound of the Wishart matrix in Eq. (10). Such a cross-correlation decomposition method is powerful in the study of the time correlations of individual stock prices and stock-market indices, and the interaction structure of business sectors.[39]

4. Multi-agent-based models

The leverage and anti-leverage effects are crucial for the understanding of the price dynamics, and important for risk management and optimal portfolio choice.[5,20,28,29,40] However, the origination of the return-volatility correlation is still disputed, especially at the microscopic level. So far, various macroscopic models have been proposed to understand the return-volatility correlation, and the retarded volatility model is an enlightening one.[5,20,29] However, it is a model with only one degree of freedom, and both the initial time series of returns and the feedback return-volatility interaction are actually input. In recent years, many researches have been devoted to the return-volatility correlation, but how to produce the return-volatility correlation with a microscopic model remains open.

More recently, a multi-agent-based model is proposed to reproduce the probability distribution of price returns and trades in stock markets.[41] It is an outstanding model with key parameters determined from empirical findings rather than from being set artificially. In this line, we construct a multi-agent-based model with asymmetric trading and herding to explore the microscopic origination of the leverage and anti-leverage effects.[42] In the past decades, although the asymmetric trading and herding behaviors may have been touched on macroscopically, they have not been taken into account in the microscopic modeling yet. Especially, we propose effective methods to determine the key parameters from historical market data.

Our model is basically built on agents' daily trading, i.e., buying, selling, and holding stocks. Besides the investment horizon, most crucially, two important behaviors of investors are taken into account for understanding the return-volatility correlation. (i) Investors' asymmetric trading in bull and bear markets; and (ii) investors' asymmetric herding in bull and bear markets.

The logarithmic price return on day t is denoted as $R(t)$. In stock markets, the information for investors is highly incomplete, therefore an agent's decision of buy, sell, or hold is assumed to be random. Since intraday trading is not persistent in empirical trading data, we consider that only one trading decision is made by each agent in a single day.[42,43] In our model, there are N agents, and each operates one share every day. On day t, each agent i makes a trading decision $S_i(t) = 1, -1, 0$, corresponding to buy, sell, or hold, respectively. The price return $R(t)$ in our model is defined by

$$R(t) = \sum_{i=1}^{N} S_i(t). \tag{12}$$

The investment horizon is introduced since agents' decision makings are based on the previous stock performance of different time horizons. It has been found that the relative portion γ_i of agents with i days investment horizon follows a power-law decay, $\gamma_i \propto i^{-\eta}$ with $\eta = 1.12$.[41] The maximum investment horizon is denoted as M. We introduce a weighted average return $R'(t)$ to describe the integrated investment basis of all agents,

$$R'(t) = k \sum_{i=1}^{M} \left[\gamma_i \sum_{j=0}^{i-1} R(t-j) \right], \tag{13}$$

where k is a proportional coefficient. According to Ref. [44], the investment horizons of investors range from a few days to several months. We estimate the maximum investment horizon M to be 150 in our model. For M between 50 and 500, the results remain robust.

(i) Asymmetric trading Let us denote the probabilities of buy, sell, or hold as $P_{\rm buy}(t)$, $P_{\rm sell}(t)$, and $P_{\rm hold}(t)$. In Ref. [41], investors' probabilities of buy and sell are assumed to be equal, i.e., $P_{\rm buy} = P_{\rm sell} = p$, with p being a constant. We adopt the value of p determined in Ref. [41], $p = 0.0154$. We assume $P_{\rm buy}(t) = P_{\rm sell}(t)$ as well, but now $P_{\rm buy}(t)$ and $P_{\rm sell}(t)$ evolve with time since the agents' trading is asymmetric in bull and bear markets. As the trading probability $P_{\rm trade}(t) = P_{\rm buy}(t) + P_{\rm sell}(t)$, we set its average over time $\langle P_{\rm trade}(t)\rangle = 2p$.

We define the market to be bullish if $R'(t) > 0$, and bearish if $R'(t) < 0$. Thus, $P_{\rm trade}(t+1)$ should take the form

$$\begin{cases} P_{\rm trade}(t+1) = 2p\alpha, & R'(t) > 0, \\ P_{\rm trade}(t+1) = 2p, & R'(t) = 0, \\ P_{\rm trade}(t+1) = 2p\beta, & R'(t) < 0. \end{cases} \tag{14}$$

Here α and β are constants, and $\langle P_{\rm trade}(t)\rangle = 2p$ requires $\alpha + \beta = 2$. If $\alpha = \beta = 1$, the trading is symmetric.

(ii) Asymmetric herding The herding behavior implies that investors are divided into groups. A herding degree $D(t) = n_{\rm A}(t)/N$ may be introduced to quantify the herding behavior, with $n_{\rm A}(t)$ being the average number of agents in each group on day t. Since herding should be related to previous volatilities,[12,45] we set $n_{\rm A}(t+1) = |R'(t)|$. Hence the herding degree on day $t+1$ is $D(t+1) = |R'(t)|/N$. This herding degree is symmetric for $R'(t) > 0$ and $R'(t) < 0$. If herding behaviors in bull and bear markets are asymmetric, it should be redefined as

$$D(t+1) = |R'(t) - \Delta R|/N, \tag{15}$$

where ΔR is the degree of asymmetry.

Now the key step is the determination of the parameters α and ΔR. We emphasize that α and ΔR will be determined from the historical market data, such as the price returns and trading volumes, rather than from statistical fittings of the simulated results. Six representative stock-market indices are studied with our model. We collect the daily data from 1950 to 2012 for the S&P 500 Index, from 1991 to 2006 for the Shanghai Index, from 2003 to 2012 for the Nikkei 225 Index, from 2004 to 2012 for the FTSE 100 Index, from 2001 to 2012 for the Hangseng Index, and from 2008 to 2012 for the DAX Index.

Since the methods for the determination of α and ΔR are somewhat lengthy, we omit the details in Ref. [42]. Here we just present the results in Table 2. We note that only the value of α for the Shanghai Index is significantly different from 1, and also only the value of ΔR for the Shanghai Index is negative. The number of agents in our simulations is typically set as $N = 10000$. Our model produces the time series of price returns $R(t)$ in the following procedure. Initially, the price returns of the first 150 time steps are set to be 0. On day $t+1$, we calculate $R'(t)$ according to Eq. (13), then $P_{\rm trade}(t+1)$ and $D(t+1)$ according to Eqs. (14) and (15), respectively. Next, we randomly divide all agents into $1/D(t+1)$ groups. The agents in a group make the same trading decision (buy, sell, or hold) with the same probability ($P_{\rm buy}$, $P_{\rm sell}$, or $P_{\rm hold}$). After all agents made their decisions, we calculate $R(t+1)$ with Eq. (12). Repeating this procedure, we obtain the time series $R(t)$. A total of 20000 data points of $R(t)$ are produced in each simulation, but the first 10000 are abandoned for equilibration.

Table 2. The values of α and ΔR determined for the six indices.[42]

Index	α	ΔR
S&P 500 (1950–2012)	1.01 ± 0.01	3
Shanghai (1991–2006)	1.09 ± 0.01	−2
Nikkei 225 (2003–2012)	1.01 ± 0.01	2
FTSE 100 (2004–2012)	0.99 ± 0.01	2
Hangseng (2001–2012)	1.02 ± 0.02	2
DAX (2008–2012)	0.98 ± 0.02	1

In Fig. 4, the simulation results for the return-volatility correlation function are displayed for the Shanghai Index and S&P 500, in good agreement with the empirical ones. As shown in Fig. 3 of Ref. [42], results for the other four indices are similar. It is most important that the parameters α and ΔR determined from empirical prices returns and trading volumes, automatically generate the leverage or anti-leverage effect in real stock markets. Careful simulations show that both the investors' asymmetric trading and herding are essential generation mechanisms for the leverage and anti-leverage effects. However, the investors' trading is approximately symmetric for the five stock markets exhibiting the leverage effect, thus contributing very little. Further analysis in Ref. [42] convinces that our model also correctly produces the fat-tail distribution of price returns and long-range auto-correlation of volatilities.

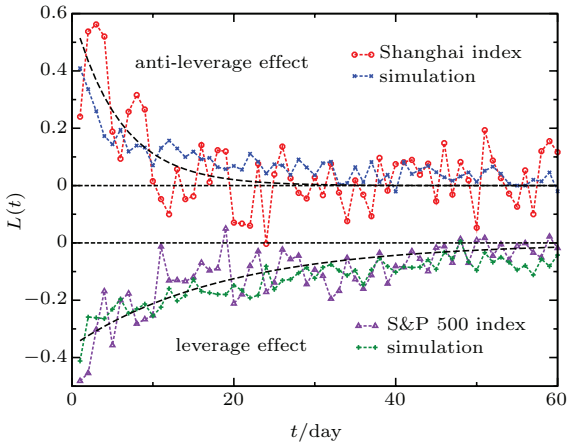


Fig. 4. Return-volatility correlation functions for the Shanghai and S&P 500 indices, and for the corresponding simulations. The Shanghai and S&P 500 indices are simulated with $(\alpha, \Delta R) = (1.0, 3)$ and $(\alpha, \Delta R) = (1.1, -2)$, respectively. Dashed lines show an exponential fit $L(t) = c\exp(-t/\tau)$.[42]

5. Large-fluctuation dynamics

Assuming that a financial market is in a stationary state, one may analyze its static statistical properties as in the preceding sections. For a comprehensive understanding of the financial market, however, it is also important to investigate the non-stationary dynamic properties. A typical example is the so-called financial crash.[8] In the past years, many activities have been devoted to such a topic.[9,24,46,47] A weakness of these studies is that there are usually not so many events for a statistical average. Stimulated by this, we systematically analyze the large-fluctuation dynamics in financial markets, based on the minute-to-minute and daily data of the Chinese Indices and German DAX.[48] In our study, a large fluctuation is identified when its volatility is large compared with the average one, but not yet as extremely large as a crash. Therefore, we gain a sufficient number of events for a statistical average. On the time scale of minutes, those large volatilities may have nothing to do with real financial crashes or rallies. Even on the daily time scale, a large volatility may not correspond to a real financial crash or rally either. However, as the magnitude of the large volatility increases, it approaches a real financial crash or rally. To our knowledge, the dynamic behavior of rallies has not been analyzed in detail.

We will investigate the dynamic relaxation both before and after the large fluctuations, and focus on the time-reversal symmetry or asymmetry at different time scales. To achieve more reliable results, we introduce the remanent and anti-remanent volatilities to describe the large-fluctuation dynamics, different from those in Refs. [9] and [47]. More importantly, we examine different categories of the large fluctuations, and explore the origin of the time-reversal asymmetry on the daily time scale.

The data that we will use for our analysis in this section are the same as those in Section 2. The so-called "Chinese Indices" are the averages of the Shanghai Index and Shenzhen Index. Denoting a financial index at time t as $P(t)$, the return and volatility are defined as $R(t) = \ln P(t+1) - \ln P(t)$ and $|R(t)|$, respectively. Naturally, the dynamic properties of volatilities may depend on the time scale. We introduce the remanent and anti-remanent volatilities as

$$v_{\pm}(t) = [\langle |R(t' \pm t)| \rangle_{\rm c} - \sigma]/Z, \tag{16}$$

where $Z = \langle |R(t')| \rangle_{\rm c} - \sigma$, σ is the average volatility, and $\langle \cdots \rangle_{\rm c}$ represents the average over t' with specified large volatilities. The large volatilities are selected by the condition $|R(t')| > \zeta$, and the threshold ζ is well above σ, e.g., $\zeta = 2\sigma, 4\sigma, 6\sigma$, and 8σ. The remanent volatility $v_+(t)$ describes how the system relaxes from a large fluctuation to the stationary state, while the anti-remanent volatility $v_-(t)$ depicts how it approaches a large fluctuation.

Large shocks in volatilities are usually followed by a series of aftershocks. Thus, we assume that both $v_+(t)$ and $v_-(t)$ obey a power law,

$$v_{\pm}(t) \sim (t + \tau_{\pm})^{-p_{\pm}}, \tag{17}$$

where $p_{\pm}$ are the exponents, and $\tau_{\pm}$ are positive constants. In most cases, the constants $\tau_{\pm}$ are rather small. To reduce the fluctuations, we integrate Eq. (17) from 0 to t. Thus, the cumulative function of $v_{\pm}(t)$ is written as

$$V_{\pm}(t) \sim [(t + \tau_{\pm})^{1-p_{\pm}} - \tau_{\pm}^{1-p_{\pm}}], \quad p_{\pm} \neq 1. \tag{18}$$

Due to the so-called intra-day pattern,[4,29] direct calculations of $V_{\pm}(t)$ with the minute-to-minute data may suffer from periodical fluctuations at a working day. Such a kind of intra-day patterns should be removed.[4,21,48] The $V_{\pm}(t)$ computed with the minute-to-minute and daily data for the Chinese Indices and German DAX are shown in Figs. 1 and 2 of Ref. [48]. The curves for the minute-to-minute data perfectly fit with Eq. (18), while those for the daily data have certain fluctuations. The exponents $p_{\pm}$ and constants $\tau_{\pm}$ are listed in Table 3. For the minute-to-minute data of both the Chinese Indices and German DAX, both p_+ and p_- increase with the threshold ζ. Especially, p_+ and p_- of every ζ are equal within statistical errors. Therefore, the dynamic behavior at the microscopic time scale, typically in minutes, is symmetric before and after large volatilities. The exponents $p_{\pm}$ of the German DAX are somewhat larger than those of the Chinese Indices. The exponents $p_{\pm}$ of the daily data also vary with ζ, similar to those of the minute-to-minute data. However, the ζ dependence of p_+ becomes obviously weaker than p_- does, i.e., $p_- \neq p_+$. In other words, the time-reversal symmetry before and after the large fluctuations is violated at the daily time scale, for both the Chinese Indices and German DAX. Again, $p_{\pm}$ of the German DAX are larger than those of the Chinese Indices.

Table 3. Measured $p_{\pm}$ with the minute-to-minute and daily data of the Chinese Indices (CHN) and German DAX. For CHN(min), DAX(min) and CHN(day), $\tau_{\pm} = 0$.[48]

		$\zeta = 2\sigma$	$\zeta = 4\sigma$	$\zeta = 6\sigma$	$\zeta = 8\sigma$
CHN(min)	p_-	0.11(1)	0.15(1)	0.17(1)	0.20(1)
	p_+	0.11(1)	0.15(1)	0.18(1)	0.22(1)
DAX(min)	p_-	0.16(1)	0.23(1)	0.27(1)	0.29(1)
	p_+	0.16(1)	0.22(1)	0.26(1)	0.29(1)
CHN(day)	p_-	0.27(3)	0.31(4)	0.36(4)	0.51(6)
	p_+	0.26(2)	0.32(3)	0.33(4)	0.36(5)
DAX(day)	τ_-	13.11	9.06	4.07	3.78
	p_-	0.41(3)	0.47(4)	0.60(5)	0.77(7)
	τ_+	10.66	9.23	7.28	3.98
	p_+	0.40(2)	0.42(3)	0.45(5)	0.46(5)

It is puzzling how the time-reversal asymmetry arises at the daily time scale. Our first thought is to classify the large

fluctuations $|R(t')|$ by $R(t') < 0$ and $R(t') > 0$, i.e., the so-called crashes and rallies. Such a classification cannot be distinguished from the exponents $p_{\pm}$. Importantly, the large fluctuations at the daily time scale could be also classified into endogenous events and exogenous events.[8,48] An exogenous event is associated with the market's response to external forces, and an endogenous event is generated by the dynamic system itself. Naturally, different stock markets may respond differently to the external forces. Looking carefully at the history of the Shanghai stock market, for example, we find that there are nine exogenous events among the sixteen large volatilities selected by the threshold $\zeta = 8\sigma$, as shown in Table 2 of Ref. [48]. For the large volatilities corresponding to the smaller thresholds such as $\zeta = 2\sigma$ and 4σ, it is not meaningful to naively identify the external forces. In Fig. 5, $V_{\pm}(t)$ of $\zeta = 6\sigma$ and 8σ are displayed for the endogenous and exogenous events of the Shanghai Index. Obviously, the dynamic relaxation of the exogenous events is faster. For the German DAX, we obtain similar results. The estimated exponents $p_{\pm}$ for both the endogenous and exogenous events are given in Table 4 of Ref. [48].

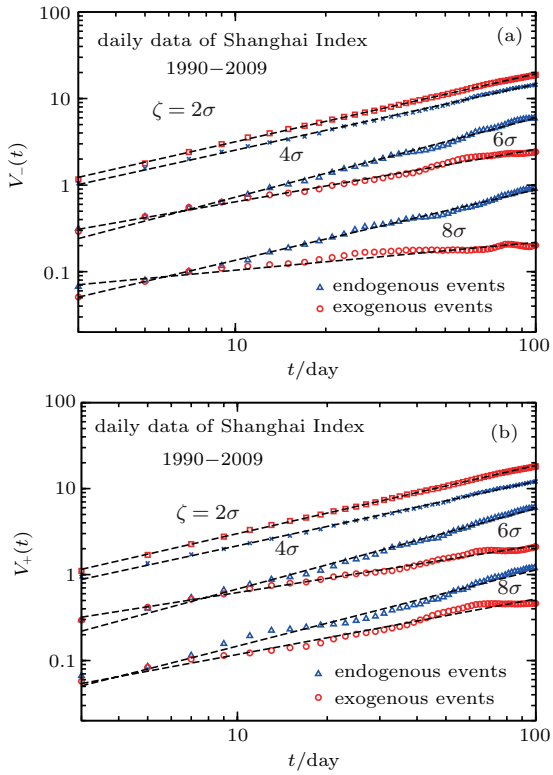


Fig. 5. (a) $V_-(t)$ and (b) $V_+(t)$ for the daily data of the Shanghai Index. For $\zeta = 6\sigma$ and 8σ, $V_{\pm}(t)$ are displayed for endogenous and exogenous events separately.[48]

Our first observation is that the time-reversal asymmetry at the daily time scale is mainly induced by the exogenous events, i.e., $p_- \approx p_+$ for the the endogenous events and $p_- > p_+$ for the exogenous events. In particular, $p_{\pm}$ of the endogenous events are almost independent of the threshold ζ, and it probably indicates that the dynamic system remains in the stationary state. However, the exogenous events may drive the dynamic system to a non-stationary state, and lead to the ζ-dependent $p_{\pm}$ and time-reversal asymmetry. The second observation is that all the exogenous events in the Shanghai Index correspond to the market-policy changes, while those in the German DAX are induced by the political and economic accidents.[48] This leads to somewhat different dynamic behaviors of the exogenous events for the Chinese and German markets. For example, it seems that once exogenous or endogenous events occur, the German market does not distinguish between exogenous and endogenous events in the future evolution.

6. Conclusion

By virtue of concepts and methods in statistical physics, we have phenomenologically investigated the time correlations and spatial structure of Chinese and western financial markets based on empirical high-frequency data. We discover that the Chinese stock market shares common basic properties with the western stock markets. For example, the exponent μ for the power-law tails of the probability distribution of price returns and β for the power-law decay of the autocorrelation function in Eqs. (2) and (4) are estimated to be $\mu = 3.8$ and $\beta = 0.32$ for both the Shanghai Index and German DAX. The persistence exponent is $\theta_{\rm p} = 0.84(2)$ and $0.87(3)$ for the Shanghai Index and German DAX, respectively.

For higher-order time correlations of price returns and volatilities, spatial correlations of individual stock prices, and large-fluctuation dynamic behaviors, however, the Chinese stock market exhibits very different characteristics from the western ones. For example, the return-volatility correlation function shows a leverage effect for the western markets, while an anti-leverage effect for the Chinese market. Interestingly, the Chinese market gradually changed to the leverage effect in recent years. From the cross-correlation matrix of individual stock prices, we derive that the Chinese market forms rather unusual sectors such as the ST sector and blue-chip sector, in contrast to the standard business sectors in the American market. Furthermore, a business sector may split into two subsectors which are anti-correlated with each other, and such a splitting phenomenon is rather prominent in the Chinese market. The large-fluctuation dynamic behavior is time-reversal symmetric at the time scale of minutes, while asymmetric at the daily time scale, and the latter is mainly induced by exogenous events. All exogenous events in the Chinese market correspond to the market-policy changes, while those in the German market are induced by political and economic accidents.

Based on the phenomenological results, microscopic many-body models are developed to simulate the microscopic

interaction and dynamic evolution of the stock markets. In particular, a multi-agent-based model with asymmetric trading and herding mechanisms is constructed to explore the microscopic origination of the leverage and anti-leverage effects, and the two key parameters in the model are uniquely determined from the historical market data.

In the next decades, physics will greatly explore its territory with its advanced theories and methodologies. The complex financial dynamics should be such a field that is already ready for quantitative and fundamental study from a microscopic level, owing to the large amount of empirical data that has piled up in the past years. On the one hand, one should further develop the theoretical approaches, such as the cross-correlation decomposition and empirical mode decomposition methods, to phenomenologically tackle the financial dynamics, and construct more general multi-agent-based models to explore the microscopic interactions and the dynamic evolution of price returns in financial systems. On the other hand, it is a great challenge to push forward the experimental study in econophysics. In a certain sense, a financial market is a large-scale experimental system which is not repeatable. How to simplify such a system, and 'simulate' it in laboratories without losing its characteristics indeed needs our wisdom.

References

[1] Mantegna R N and Stanley H E 1995 *Nature* **376** 46
[2] Gopikrishnan P, Plerou V, Amaral L A N, Meyer M and Stanley H E 1999 *Phys. Rev. E* **60** 5305
[3] Giardina I, Bouchaud J P and Mézard M 2001 *Physica A* **299** 28
[4] Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng C K and Stanley H E 1999 *Phys. Rev. E* **60** 1390
[5] Bouchaud J P, Matacz A and Potters M 2001 *Phys. Rev. Lett.* **87** 228701
[6] Plerou V, Gopikrishnan P, Rosenow B, Amaral L A N and Stanley H E 1999 *Phys. Rev. Lett.* **83** 1471
[7] Plerou V, Gopikrishnan P, Rosenow B, Amaral L A N, Guhr T and Stanley H E 2002 *Phys. Rev. E* **65** 066126
[8] Sornette D 2003 *Phys. Rep.* **378** 1
[9] Lillo F and Mantegna R N 2003 *Phys. Rev. E* **68** 016119
[10] Tumminello M, Aste T, Di Matteo T and Mantegna R N 2005 *Proc. Natl. Acad. Sci.* **102** 10421
[11] Challet D and Zhang Y C 1997 *Physica A* **246** 407
[12] Cont R and Bouchaud J P 2000 *Macroeconomic Dynamics* **4** 170
[13] Eguiluz V M and Zimmermann M G 2000 *Phys. Rev. Lett.* **85** 5659
[14] Zheng B, Qiu T and Ren F 2004 *Phys. Rev. E* **69** 046115
[15] Ren F, Zheng B, Qiu T and Trimper S 2006 *Phys. Rev. E* **74** 041111
[16] Li G H, Zhang H and Luo M K 2012 *Chin. Phys. B* **21** 128901
[17] Zhao L, Yang G, Wang W, Chen Y, Huang J P, Ohashi H and Stanley H E 2011 *Proc. Natl. Acad. Sci. USA* **108** 15058
[18] Sonnemann U, Camerer C F, Fox C R and Langer T 2013 *Proc. Natl. Acad. Sci. USA* **110** 11779
[19] Liang Y, Yang G and Huang J P 2013 *Front. Phys.* **8** 438
[20] Shen J and Zheng B 2009 *Europhys. Lett.* **88** 28003
[21] Qiu T, Zheng B, Ren F and Trimper S 2007 *Physica A* **378** 387
[22] Roman H E, Porto M and Dose C 2008 *Europhys. Lett.* **84** 28001
[23] Zheng B, Ren F, Trimper S and Zheng D F 2004 *Physica A* **343** 653
[24] Ren F and Zhou W X 2010 *New J. Phys.* **12** 075030
[25] Ren F, Zheng B and Chen P 2010 *Physica A* **389** 2744
[26] Ren F and Zheng B 2003 *Phys. Lett. A* **313** 312
[27] Ren F, Zheng B, Lin H, Wen L Y and Trimper S 2005 *Physica A* **350** 439
[28] Black F 1976 *Studies of Stock Price Volatility Changes, Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, Alexandria, p. 177
[29] Qiu T, Zheng B, Ren F and Trimper S 2006 *Phys. Rev. E* **73** 065103
[30] Tan L, Zheng B, Chen J J and Jiang X F 2014 *How Volatilties Nonlocal in Time Control the Price Movement in Complex Financial Systems*, preprint
[31] Giraitis L, Leipus R, Robinson P M and Surgailis D 2004 *J. Financ. Econom.* **2** 177
[32] Ouyang F Y, Zheng B and Jiang X F 2014 *Physica A* **402** 236
[33] Bonanno G, Caldarelli G, Lillo F and Mantegna R N 2003 *Phys. Rev. E* **68** 046130
[34] Aste T, Shaw W and Matteo T D 2010 *New J. Phys.* **12** 085009
[35] Pan R K and Sinha S 2007 *Phys. Rev. E* **76** 046116
[36] Shen J and Zheng B 2009 *Europhys. Lett.* **86** 48005
[37] Qiu T, Zheng B and Chen G 2010 *New J. Phys.* **12** 043057
[38] Jiang X F and Zheng B 2012 *Europhys. Lett.* **97** 48006
[39] Jiang X F, Chen T T and Zheng B 2014 *Scientific Reports* **4** 5321
[40] Park B J 2011 *Journal of Banking & Finance* **35** 2657
[41] Feng L, Li B, Podobnik B, Preis T and Stanley H E 2012 *Proc. Natl. Acad. Sci. USA* **109** 8388
[42] Chen J J, Zheng B and Tan L 2013 *PloS One* **8** e79531
[43] Eisler E and Kertesz J 2007 *Europhys. Lett.* **77** 28001
[44] Menkhoff L 2010 *Journal of Banking & Finance* **34** 2573
[45] Blasco N, Corredor P and Ferreruela S 2012 *Quant. Financ.* **12** 311
[46] Weber P, Wang F Z, Vodenska-Chitkushev I, Havlin S and Stanley H E 2007 *Phys. Rev. E* **76** 016109
[47] Mu G H and Zhou W X 2008 *Physica A* **387** 5211
[48] Jiang X F, Chen T T and Zheng B 2013 *Physica A* **392** 5369