Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

On the Application of the Cross-Correlations in the Chinese Fund Market: Descriptive Properties and Scaling Behaviors

by

Weibing Deng, Wei Li, Xu Cai, and Qiuping Wang

Preprint no.: 46 2011

Advances in Complex Systems, Vol. 14, No. 1 (2011) 97–109
© World Scientific Publishing Company
DOI: 10.1142/S0219525911002871

ON THE APPLICATION OF THE CROSS-CORRELATIONS IN THE CHINESE FUND MARKET: DESCRIPTIVE PROPERTIES AND SCALING BEHAVIORS

WEIBING DENG*,†,‡,¶, WEI LI†,§, XU CAI†
and QIUPING A. WANG*,‡

*ISMANS, 44, Ave. Bartholdi,
72000 Le Mans, France

†Complexity Science Center & Institute of Particle
Physics, Hua-Zhong (Central China)
Normal University, Wuhan 430079, China

‡LPEC, UMR CNRS 6087, Universite du Maine,
72085 Le Mans, France

§Max-Planck Institute for Mathematics in the Sciences,
Inselstr. 22–26, D-04103 Leipzig, Germany
¶dengwb@phy.ccnu.edu.cn

Received 21 April 2010
Revised 31 October 2010

On the basis of the relative daily logarithmic returns of 88 different funds in the Chinese fund market (CFM) from June 2005 to October 2009, we construct the cross-correlation matrix of the CFM. It is shown that the logarithmic returns follow an exponential distribution, which is commonly shared by some emerging markets. We hereby analyze the statistical properties of the cross-correlation coefficients in different time periods, such as the distribution, the mean value, the standard deviation, the skewness and the kurtosis. By using the method of the scaled factorial moment, we observe the intermittence phenomenon in the distribution of the cross-correlation coefficients. Also by employing the random matrix theory (RMT), we find a few isolated large eigenvalues of the cross-correlation matrix, and the distribution of eigenvalues exhibits the power-law tails. Furthermore, we study the features of the correlation strength with a simple definition.

Keywords: Cross-correlation; random matrix theory; scaled factorial moment; correlation strength; power-law.

1. Introduction

The financial markets have been perceived as a collection of nonlinear interaction units [1]. Physicists are making great efforts to probe the empirical laws, construct the financial models and explore the economical theories, which will help us to understand such complex interactions. Especially, the financial markets have

¶Corresponding author.

been found to exhibit some universal features, which are similar to those observed in physical systems with large numbers of interacting units [2]. For example, the logarithmic return distributions of the real world markets display the peak-center and fat-tail properties [3, 4], and there exists the intermittent behavior in the price changes [5], etc.

The Chinese fund market (CFM), with a history of about 10 years, contributes substantially to the Chinese financial market, and promotes significantly to establish a sound and complete financial system. Particularly in the recent years, on the basis of its inherent features of moderate risks and considerable benefits, with the trading volume and the industry size growing steadily, the CFM plays a more and more important role in the Chinese financial market [6].

Through issuance of the funds, the fund management companies raise the money from the investors, and then, they use these collected financing to invest in stocks, bonds, financial derivatives, real estate industry or some other prosperous companies, etc. Afterwards, the fund management companies and the investors share the investment risk and revenue.

There are many common economic factors in the fund market [7], which influence the fund transactions, including the investment portfolio, the risk prediction, and the deposit interest rates, etc. The funds with the same common economic factors are strongly correlated with each other and tend to be grouped into a community. Therefore, to investigate the nature of correlations among the funds is very important to comprehend the pricing mechanism in the fund market [8]. Generally, the cross-correlation matrix [9] has been used to analyze such interactions among the funds, and if we can obtain significant information embedded in the cross-correlation matrix, we will better understand the CFM.

Previous studies about the financial networks have focused on the topological properties and the formation principles [8], which revealed that the degree distribution of the stock network follows the power-law [10, 11]. Whether the stock logarithmic returns from pricing models can explain interactions among stocks was also investigated, which may show the deterministic factors that significantly affect the formation process of the stock networks [12, 13].

At the same time by using the scaled factorial moment [14] and the random matrix theory [15], we investigate the characteristics of the cross-correlation matrix about the relative logarithmic return. The whole text is organized as follows. We show the construction of the cross-correlation matrix in Sec. 2. Section 3 presents the distribution, the mean value, the standard deviation, the skewness and the kurtosis of the elements in the cross-correlation matrix. In Sec. 4, applying the scaled factorial moment method, we analyze the features about the distribution of the correlation coefficients in different time scales. Section 5 depicts the eigenvalue statistics of the cross-correlation matrix. In Sec. 6, we investigate the correlation strength with a simple definition. The conclusion is given in the last part, Sec. 7.

2. Construction of the Cross-Correlation Matrix

Here, we chose a 4-year research period ranging from June 2005 to October 2009, and collected the daily closed prices of 88 funds from the web-site of the Chinese fund market [16]. Let $p_i(t)$ be the price of fund i at moment t, then the logarithmic return of the fund i after a time interval Δt is defined as

$$r_i(t) = \ln p_i(t + \Delta t) - \ln p_i(t), \tag{1}$$

where the logarithmic return $r_i(t)$ denotes the logarithm fluctuation of the price during the time interval Δt, which can be taken as 1 day, 1 week, or 1 month.

To reduce the influence of different investment portfolio, we define the relative logarithmic return

$$R_i(t) = r_i(t) - \frac{1}{N}\sum_j r_j(t), \tag{2}$$

where $R_i(t)$ denotes the relative logarithmic return of the fund i at time t, which refers to the logarithmic return $r_i(t)$ as compared to average logarithmic returns of the whole 88 funds at time t.

Firstly, we calculated the relative daily logarithmic returns of the 88 funds from June 2005 to October 2009, where only the absolute values are considered. The distribution of the absolute relative logarithmic return $P(|R|)$ is shown in Fig. 1. The Kolmogorov–Smirnov test has been employed to test the distribution of the relative logarithmic returns, which indicates the P value of 0.36, being greater than 0.05. Thus the distribution of the relative logarithmic returns is indeed exponential.

Such similar distributions have been reported in the Indian [17], Japanese [18], German [19], Brazilian markets [20, 21], and US market [22, 23], thus we should be

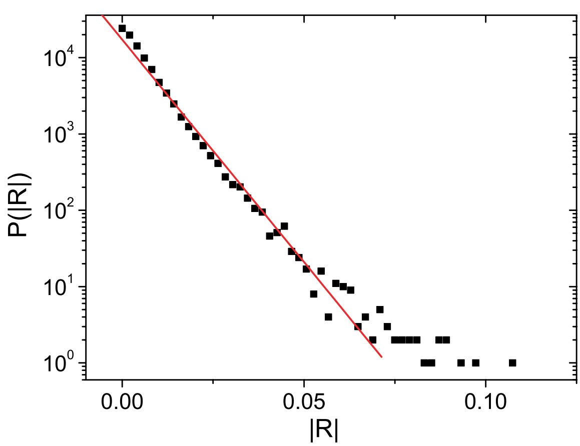

Fig. 1. The distribution of the absolute relative daily logarithmic return in a single-logarithmic scale.

aware of the presence of the exponential distribution and recognize it as another "stylized fact" in the set of analytical tools for financial data analysis [24].

Then the elements of the cross-correlation matrix C are

$$C_{ij} = \frac{\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle}{\sqrt{(\langle R_i^2 \rangle - \langle R_i \rangle^2)(\langle R_j^2 \rangle - \langle R_j \rangle^2)}}, \tag{3}$$

where C is a real symmetric matrix with positive eigenvalues, and the values of C_{ij} range from -1 to 1: $C_{ij} = 1$ means the perfect positive correlations, while $C_{ij} = -1$ shows the perfect negative correlations and $C_{ij} = 0$ predicts no correlations.

3. Basic Properties of the Cross-Correlation Coefficients

The cross-correlation coefficients C_{ij} reflect the correlation degree of the relative logarithmic return between funds i and j, while the distribution of the cross-correlation coefficients shows the correlation situation of the whole fund market. To compare the distributions of the correlation coefficients in different years, the correlation has been computed over one year long data blocks from June 2005 to May 2009, and the distributions are shown in Fig. 2.

In the fund market, the cross-correlations among different funds are not independent. For example, in the same subset, the price of the funds may rise or

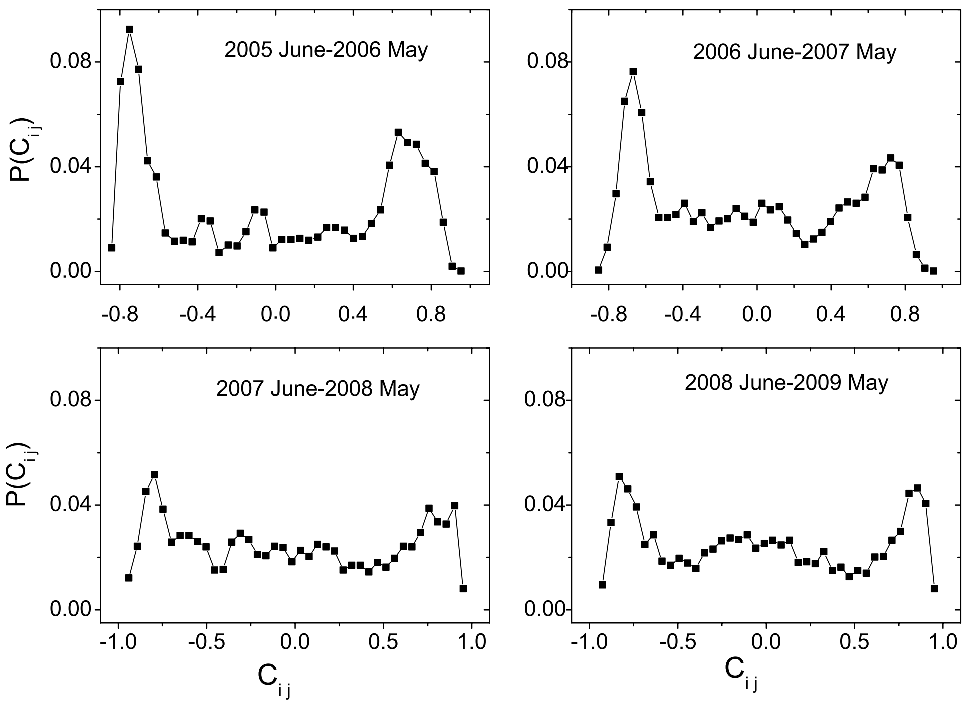

Fig. 2. The time-dependent distributions of the cross-correlation coefficients C_{ij} in every year from June 2005 to May 2009.

fall simultaneously, which may result in the correlation matrix for each subset of funds might have a positive determinant. However, the situations do not always happen like that, and here the time span we considered is one year, which dramatically minimizes the effect of this dependence. Therefore, this dependence of the cross-correlation has little influence on their distributions of the one-year-long database.

As one can observe that the cross-correlation coefficients mainly distribute in the regimes of either large or small values, which means there may exist some communities in the fund market who share the same or the inverse investment portfolio. In other words, there exists some leading fund management companies in the fund market, and the others just imitate their actions, like the herding behaviors in the stock market [25, 26].

In order to gain some information embedded in the fund market, we shift to a continuous analysis of the 88 funds from June 2005 to October 2009, and characterize the cross-correlation distribution by its first four moments, the mean value, standard deviation, skewness and kurtosis of the cross-correlation coefficients.

Firstly, the whole continuous time span has been divided into several bins, the length of every bin is $T = 5$, since only five transaction days in a week. Then, we

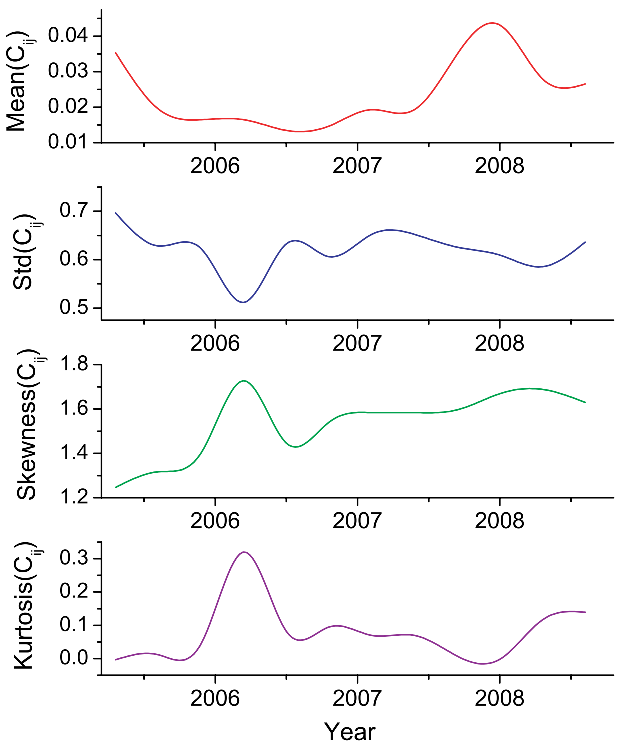

Fig. 3. The time-dependent descriptive statistics of the cross-correlation coefficients, including the mean value, the standard deviation, the skewness and the kurtosis.

calculate the first four moments of the cross-correlation coefficients in each bin, therefore a specific point just reflects the trend of that week, it belongs to. So many discrete points have been connected by a smooth line. From Fig. 3, an interesting finding is that the effects of two major bubbles in the Chinese financial market are clearly visible, where the mean values of the cross-correlation coefficients are larger than others during the initial time of bubbles, while the values of the skewness and the kurtosis also become larger during the intermediate periods of the bubbles. This interesting phenomenon may imply that people might take similar actions during the bubble periods.

Furthermore, we also considered whether these four different measures are correlated by calculating the Pearson's linear and Spearman's rank-order correlation coefficients, the values of which between the mean value and the standard deviation are 0.95 and 0.91, while the ones between the skewness and kurtosis are 0.93 and 0.97, respectively. Thus, the first two and the last two measures are very strongly correlated with each other.

4. Scaled Factorial Moment Method

The study of non-statistical fluctuations in the multi-particle production has entered into a new era since Bialas and Peschanski introduced an attractive methodology, named as the scaled factorial moment method [27], which has the feature that it can measure the non-statistical fluctuations avoiding the statistical noise [28].

Now, it has been widely used in the analysis of event by event correlations and fluctuations, such as the DNA sequence [29], the hadronic decay [30], and the spectra of complex networks [31], as it allows for a direct access to the dynamical fluctuations of the multiplicity with the function of dismissing the statistical fluctuations due to finite number of events [14]. While applied to the financial systems, it provides a way to analyze the intermittency or the self-similarity behaviors in the dynamics of distribution of the correlation coefficients, despite the limited financial data.

The intermittency of a time series can be defined as its normalized difference in scaling parameters. Many observed time series are intermittent in the sense that observations differ dramatically from previous observations from time to time, for example, fast increases in the heart rate resulting from physiological activity exhibit intermittency. It is an essential property of the system, which also reflects the self-similarity behavior, and has been estimated to characterize data. The concept of intermittency we considered here has connections with the concept of multi-scaling or multi-fractal in the stochastic processes to a certain degree.

Imitating the methods in the high energy nuclear collisions, the scaled factorial moment has been used to search for the intermittency in the cross-correlation coefficients. We firstly divide the value range of the cross-correlation coefficients Δ into M intervals δ, with $M = \Delta/\delta$, n is the number of cross-correlation coefficients that fall in δ of one event. Here "one event" refers to the ensemble of cross-correlation coefficients among a single fund and the other funds, then the q-order scaled factorial moment is defined as

$$F_q = \frac{\langle n(n-1)\cdots(n-q-1)\rangle}{\langle n\rangle^q}, \tag{4}$$

where the brackets mean an average over all events.

If power-law scaling follows, just as

$$F_q \sim \delta^{-\varphi_q}, \qquad \varphi_q > 0, \tag{5}$$

then we can say that the intermittency phenomena is observed in the behavior of multiplicity fluctuations, where φ_q is the intermittent exponent. The intermittency phenomena also indicates the self-similarity characteristics of the system [14].

Calculating the second-order scaled factorial moment F_2, we showed the relationship between F_2 and the intervals δ in Fig. 4, in which the power-law scaling

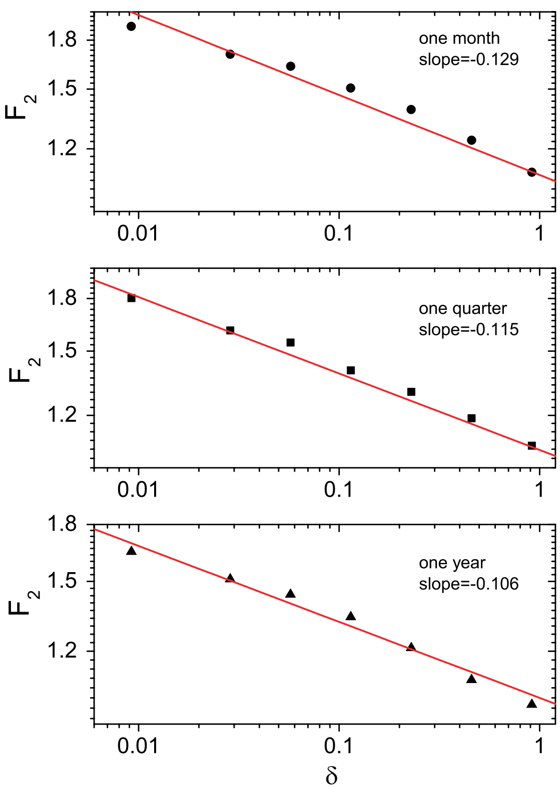

Fig. 4. The intermittency behavior of the cross-correlation coefficients by using the scaled factorial moment method in different time scales, one month, one quarter and one year.

has been shown in the double-logarithm scale, with

$$\ln F_2 = \alpha - \beta \ln \delta, \tag{6}$$

$$F_2 \sim \delta^{-\beta}, \quad \beta > 0. \tag{7}$$

This scaling has been found in different time scales: one month, one quarter and one year, which suggests it is insensitive to time span. Hence, we see that the intermittence phenomena exists in the fluctuation of cross-correlation among the relative fund logarithmic returns, and so does the self-similarity characteristics in the dynamics of the cross-correlation coefficients. This may further show the evidence that the fluctuations of the fund prices obey the long-range correlations [32].

5. Eigenvalue Statistics of the Cross-Correlation Matrix

In statistics, a random matrix is a matrix whose elements are random variables, and many important features of the physical systems can be represented mathematically as the random matrix problems [33]. In the 1950s, Wigner applied the random matrices to model the energy levels at nuclear reactions in quantum physics [34–36]. Since then, the random matrices have been of great importance in statistics.

The cross-correlation matrices are of great importance to extract the underlying information in the empirical data. There have been many reports on the applications of random matrix theory (RMT) to the investigation of the correlation and covariances of stock price changes [37], which shows the distribution of the eigenvalues and the components of most of the corresponding eigenvectors largely follow that of the relevant ensemble of random matrices [38], and there exists many eigenvalues which are larger than the maximum expected for a completely random matrix, etc. [39]. However, the assumptions on the random matrix can strongly affect its spectrum. Recently, many researches have been carried out to investigate which spectrum can be yielded by the covariance matrices, built from time series with non-Gaussian statistics, for example, the Lévy α-stable statistics. Politi *et al.* [40] have reviewed the analytic derivation of the spectral density of free stable Wishart–Lévy random matrices, which was already solved by Burda *et al.* [41, 42], and furthermore, they have validated numerically the analytic results by the Monte Carlo simulations.

Here, we analyze the cross-correlation matrix of the relative daily logarithmic returns of the 88 funds from June 2005 to October 2009, and the main goal is to find more evidence on the applicability of the RMT to the analysis of the CFM. Therefore, we calculate the eigenvalues of the cross-correlation matrix and show the distribution in Fig. 5.

We have tested the spectrum by using a KS-type goodness-of-fit test for power law distribution hypothesis [43], firstly, the power-law exponent has been calculated by applying the maximum likelihood estimation (MLE), then the KS test has been employed to calculate the maximum distance between the hypothesized cumulative distribution and the empirical distribution. It is found that $K = 0.1526$ is below

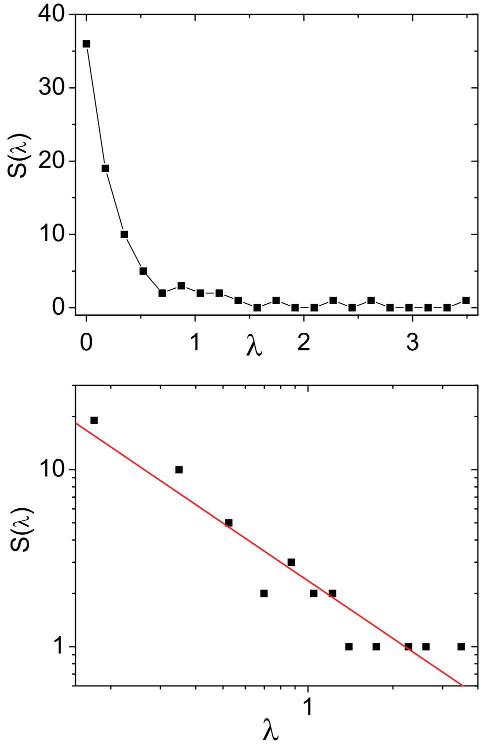

Fig. 5. The eigenvalue distribution of the cross-correlation matrix, where in the top panel the natural scale is adopted and in the bottom one the double-log scale is done, with slope −1.12.

the 0.9 quantile and 10 samples, therefore, with the observed significance level greater than 10%, there is insufficient evidence to reject the hypothesis that the distribution is a power-law. Furthermore, we can find that most eigenvalues of the cross-correlation matrix are very small, while a few of them are much bigger, which may indicate that a few funds' price variation can have a dramatic effect on other ones' prices.

6. Correlation Strength

From the above analysis, we know that the funds are correlated. In this part, we investigated the correlation from the weight point of view [28]. This study may provide information for the importance of a fund i in the whole Chinese fund market, and can be quantified by the correlation strength S_i,

$$S_i = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} |C_{ij}|, \tag{8}$$

where $N = 88$, the total number of the funds.

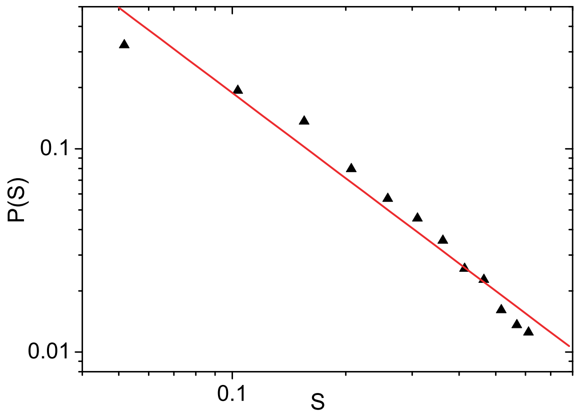

Fig. 6. The distribution of the correlation strength of the Chinese fund market in the double logarithm scales, which follows the power-law with slope −1.39.

We plotted the distribution of the correlation strength in Fig. 6, and employed the same method in Sec. 5 to test the distribution, which shows the observed significance level is also greater than 10%. There is insufficient evidence to reject the hypothesis that the distribution is a power-law, with $P(S) \propto S^{-\gamma}$, $\gamma = 1.39$.

On the one hand, the power-law decaying distribution may again illustrate that there exists a couple of funds having strong influence in the fund market, which powerfully affect the fluctuation of the other funds' prices. We listed the top 15 funds with large correlation strengths in Table 1, and the results are close to that of the fund rating in five years by the Morningstar corporation and the Lipper company [44].

On the other hand, the presence of a scaling invariance in the correlation of the fund market would be related to the ultimate hierarchical structure, which is similar to the results in [45, 46].

Table 1. List of the top 15 funds that have large correlation strength.

Compositor	Name of the funds	Correlation strength
1	HUA XIA DA PAN	0.62
2	PENG HUA CHINA 50	0.60
3	FU GUO TIAN LI	0.58
4	JIA SHI GROWING	0.57
5	HUA AN BAO LI	0.56
6	HUA BAO XING YE	0.53
7	HUA XIA HUI BAO	0.51
8	TAI DA HE YIN	0.50
9	ZHONG XIN JING DIAN	0.49
10	GUO TOU RUI YIN	0.46
11	JING SHUN GREAT WALL	0.45
12	YI FANG DA STRATEGY	0.45
13	GUANG FA WEN JIAN	0.43
14	YIN HE WEN JIAN	0.42
15	GUO TAI JIN YING	0.42

7. Conclusion

We have studied the characteristics of the cross-correlation matrix for the relative daily logarithmic return of the Chinese fund market by using the concepts and methods of data analysis.

We have discovered the relative daily logarithmic return of the 88 funds exhibits the exponential distribution, and studied the basic properties of the cross-correlation coefficients, which include the distribution, its first four moments and their correlations with one another. The descriptive statistical measures such as mean value, standard deviation, kurtosis and skewness of the cross-correlation coefficients show distinct behavior in the main two bubbles of the Chinese financial market.

By using the scaled factorial moment method, we observed the intermittency phenomena. This may suggest that the distribution of the cross-correlation coefficients is time-dependent, and has the self-similarity property that brings the scale symmetries embedded in the topological structures of economical systems.

We have examined the structure of the cross-correlation matrix by employing the random matrix theory (RMT), which suggests that the distribution of eigenvalues of the cross-correlation matrix exhibits the power-law tails, and there exist a few isolated large eigenvalues.

The correlation strength of each fund is defined as the average of weights, namely the absolute values of the cross-correlation coefficients, which images the importance of the fund in the whole fund market. We have found the strong correlations present in the CFM. Some funds have the strong influence in the fund market, which greatly affect the fluctuation of the other funds' prices.

Our work may shed light on interpreting the pricing mechanism in the Chinese fund market. Future studies may pay attention to the community structure analysis of the CFM.

Acknowledgments

Weibing Deng would like to show his gratitude to Prof. Sornette for the valuable suggestions of the previous work. The authors also thank the other members in the Complexity Science Center of Central China Normal University for helpful discussions.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 10647125, 10635020, 10975057 and 10975062), the Programme of Introducing Talents of Discipline to Universities under Grant No. B08033, and the PHC CAI YUAN PEI Programme (LIU JIN OU [2010] No. 6050).

References

[1] Lux, T. and Marchesi, M., *Nature* **397** (1999) 498.
[2] Mantegna, R. N., *European Physical Journal B* **11** (1999) 193.
[3] Fama, E. F., *Journal of Business* **38** (1965) 34.

[4] Huisman, R., Koedijk, K. G., Kool, C. J. M. and Palm, F., *Journal of Business and Economic Statistics* **19** (2001) 208.
[5] Kozuki, N. and Fuchikami, N., *Physica A* **329** (2003) 222.
[6] China Fund Industry Report 2007–2008. Chap. 1.4 Development status in fund industry, 2007; China Fund Industry Report 2008–2009. Chap. 1.3 Review of Fund Market, 2008.
[7] China Fund Management Industry, 2007. Chap. 1 Fund Management in China: Past and Present.
[8] Eom, C., Oh, G. and Kim, S., *Physica A* **383** (2007) 139.
[9] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N. and Stanley, H. E., *Physical Review Letters* **83** (1999) 1471.
[10] Vandewalle, N., Brisbois, F. and Tordoir, X., *Quantitative Finance* **1** (2001) 372.
[11] Kim, H. J., Kim, I. M., Lee, Y. and Kahng, B., *Journal of the Korean Physical Society* **40** (2002) 1105.
[12] Sharpe, W. F., *Journal of Finance* **19** (1964) 425.
[13] Lintner, J., *Review of Economics and Statistics* **47** (1965) 13.
[14] Kirejczyk, M., *Acta Physical Polonica B* **35** (2004) 10.
[15] Utsugi, A., Ino, K. and Oshikawa, M., *Physical Review E* **60** (2004) 026110.
[16] http://data.cnfund.cn, The Chinese Fund Market Database.
[17] Matia, K., Pal, M., Salunkay, H. and Stanley, H. E., *Europhysics Letters* **66** (2004) 909.
[18] Kaizoji, T. and Kaizoji, M., *Advances in Complex Systems* **6** (2003) 303.
[19] Remer, R. and Mahnke, R., Application of Heston model and its solution to German DAX data, presented at the 4th Application of Physics in Financal Analysis (APFA) Conference (2003).
[20] Vicente, R., de Toledo, C. M., Leite, V. B. P. and Caticha, N., arXiv:cond-mat/0402185, 6 February 2004.
[21] Miranda, L. C. and Riera, R., *Physica A* **297** (2001) 509.
[22] Dragulescu, A. and Yakovenko, V. M., *Quantitative Finance* **2** (2002) 443.
[23] Silva, A. C. and Yakovenko, V. M., *Physica A* **324** (2003) 303.
[24] Silva, A. C., Prange, R. E. and Yakovenko, V. M., *Physica A* **344** (2004) 227.
[25] Eguiluz, V. M. and Zimmermann, M. G., *Physical Review Letters* **85** (2000) 5659.
[26] Lillo, F., Moro, E., Vaglica, G. and Mantegna, R. N., *New Journal of Physics* **10** (2008) 043019.
[27] Bialas, A. and Peshanski, R., *Nuclear Physics B* **273** (1986) 703; *ibid.* **308** (1988) 857.
[28] Jiang, J., Ma, K. and Cai, X., *Physica A* **375** (2007) 274.
[29] Mohanty, A. K. and Narayana Rao, A. V. S. S., *Physical Review Letters* **84** (2000) 1832.
[30] Chen, G., Hu, Y., Liu, L., Kittel, W. and Metzger, W. J., arXiv:hep-ph/0112092 v3, 19 December 2001.
[31] Zhao, F. C., Yang, H. J. and Wang, B. H., *Physical Review E* **72** (2005) 046119.
[32] Deng, W. B., Li, W. and Cai, X., *International Journal of Modern Physics C* **21** (2010) 79.
[33] Wilcox, D. and Gebbie, T., *Physica A* **375** (2007) 584.
[34] Mehta, M. L., *Random Matrices* (Academic Press, New York, 1991).
[35] Brody, T. A., Flores, J., French, J. B., Mello, P. A., Pandey, A. and Wong, S. S. M., *Reviews of Modern Physics* **53** (1981) 385.
[36] Guhr, T. Müller-Groeling, A. and Weidenmüller, H. A., *Physics Reports* **299** (1998) 190.

[37] Laloux, L., Cizeau, P., Bouchad, J. P. and Potters, M., *Physical Review Letters* **83** (1999) 1467.
[38] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. and Stanley, H. E., *Physical Review E* **65** (2002) 066126.
[39] Noh, J. D., *Physical Review E* **61** (2000) 5981.
[40] Politi, M., Scalas, E., Fulger, D. and Germano, G., *European Physical Journal B* **73** (2010) 13.
[41] Burda, Z., Jurkiewicz, J. and Waclaw, B., *Physical Review E* **71** (2005) 026111.
[42] Burda, Z., Jurkiewicz, J., Nowak, M. A., Papp, G. and Zahed, I., *Physical Review E* **75** (2007) 051126.
[43] Goldstein, M. L., Morris, S. A. and Yen, G. G., *European Physical Journal B* **41** (2004) 255.
[44] www.morningstarcorp.com, www.lipperweb.com.
[45] Faloutsos, M., Faloutsos, P. and Faloutsos, C., *Computer Communications Review* **29** (1999) 251.
[46] Albert, R., Jeong, H. and Barabasi, A. L., *Nature* **401** (1999) 130.