Eur. Phys. J. B **20**, 573–579 (2001)

THE EUROPEAN PHYSICAL JOURNAL B

EDP Sciences

The distribution and scaling of fluctuations for Hang Seng index in Hong Kong stock market

B.H. Wang[1,2,a] and P.M. Hui[3]

[1] Department of Modern Physics and Nonlinear Science Center, University of Science and Technology of China, Hefei 230026, PR China
[2] CCAST (World Laboratory), PO Box 8730, Beijing 100080, PR China
[3] Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, PR China

Received 9 August 2000 and Received in final form 28 August 2000

Abstract. The statistical properties of the Hang Seng index in the Hong Kong stock market are analyzed. The data include minute by minute records of the Hang Seng index from January 3, 1994 to May 28, 1997. The probability distribution functions of index returns for the time scales from 1 minute to 128 minutes are given. The results show that the nature of the stochastic process underlying the time series of the returns of Hang Seng index cannot be described by the normal distribution. It is more reasonable to model it by a truncated Lévy distribution with an exponential fall-off in its tails. The scaling of the maximium value of the probability distribution is studied. Results show that the data are consistent with scaling of a Lévy distribution. It is observed that in the tail of the distribution, the fall-off deviates from that of a Lévy stable process and is approximately exponential, especially after removing daily trading pattern from the data. The daily pattern thus affects strongly the analysis of the asymptotic behavior and scaling of fluctuation distributions.

PACS. 89.65.Gh Economics, business, and financial markets – 05.40.Fb Random walks and Lévy flights

1 Introduction

Financial markets are typical complex systems in which the large-scale dynamical properties are dependent on the evolution of a large number of nonlinear-coupled subsystems. Understanding the dynamics of these strongly fluctuating complex systems is an important scientific challenge of current interest. Techniques developed in studying complex physical systems are useful in analyzing financial data. They can yield new and insightful results. The difficulties in understanding the dynamics of a financial market come not only from the complexity of the different elements comprising the systems, but also from the different ways in which external factors come in. Even two markets in the same country or within one region, may behave quite differently. Remarkably, the statistical properties of certain observables such as the transaction price, the volume of shares traded, the trading frequency, and the values of market indices, appear to behave in a similar way for quite different markets [1,2]. This implies that there may be some "universal" behavior in financial markets and some unifying regularities governing complex economic systems.

A problem of both practical and theoretical interests concerns the distribution of the variations in share prices and the dynamic evolution of the distribution. The index returns $Z(t)$ are defined as the difference between two successive logarithms of the index $X(t)$,

$$Z(t) \equiv \text{fluc}(t) = \beta\{\ln[X(t+\Delta t)] - \ln[X(t)]\}, \quad (1)$$

where Δt is the time interval separating the non-overlapping index records and β is a magnification factor included for convenience in displaying figures. For small changes of $X(t)$, the return is approximately the forward relative change,

$$Z(t) \equiv \text{fluc}(t) \approx \beta[X(t+\Delta t) - X(t)]/X(t).$$

The most widely accepted models state that the variation of share prices is a random process. Investigating the time series of returns on varying time scales Δt is useful in probing the underlying nature of the stochastic process [3–7].

Bachelier proposed the first model for the stochastic process of returns [8]. The model describes the variation of share prices as an uncorrelated random walk with independent, identically Gaussian distributed random variables. It is natural if one considers the return over a time scale to be the result of many independent "shocks". However, empirical studies showed that the distribution of the returns has pronounced tails in striking contrast to that

[a] e-mail: bhwang@ustc.edu.cn

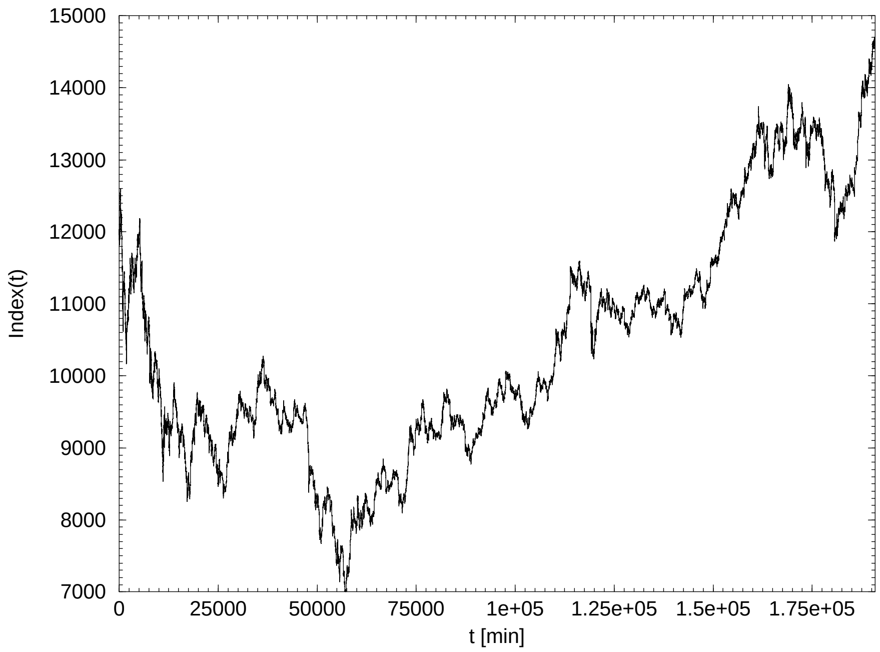

Fig. 1. Data analyzed: The Hang Seng index in Hong Kong for the period January 3, 1994 to May 28, 1997 as the function of time t at intervals of 1 minute. The total trading time includes 190821 minutes.

of the Gaussian distribution. Large events are frequent in financial data. In his pioneering analysis of cotton prices, Mandelbrot observed that in addition to being non-Gaussian, the process of returns shows a time scaling behavior [9]. He then proposed that the distribution of returns is consistent with a Lévy stable distribution [10,11]. In 1995, Mantegna and Stanley analyzed a large set of data of the S&P500 index. It has been reported that [3,12–14] the central part of the distribution of S&P 500 returns appears to be well fitted by a Lévy distribution, but the asymptotic behavior of the distribution shows faster decay than that predicted by a Lévy distribution. Hence, it was proposed [3] that a truncated Lévy distribution – a Lévy distribution in the central part follows by an approximately exponential truncation – is the model for the return distribution.

In order to probe the extent of universality in the dynamics of complex behavior in financial markets and to provide a basic and appropriate framework for developing economic models of financial markets, we investigate the distribution of the fluctuations in the Hang Seng index – the most important financial index in the Hong Kong stock market. It is observed that the distribution of returns in the Hang Seng index shows apparent scaling behavior, which cannot be modelled by a normal distribution. The non-Gaussian dynamics of the stochastic process underlying the time series of returns of the Hang Seng index, is better modelled by a truncated Lévy distribution. A power-law behavior is observed for the probability of zero return for time intervals Δt spanning at least two orders of magnitude. However, the power-law fall-off behavior in the tails deviate from that of Lévy stable process. The two tails of the distribution drop more slowly than a Gaussian, but faster than a Lévy process with an exponent outside the Lévy stable region. The exponential truncation ensures the existence of a finite second moment. The observations are useful for establishing dynamical models of the Hong Kong stock market.

2 Lévy distribution characteristics and scaling of index returns

Figure 1 shows the minute-by-minute records of a total of 838 trading days for the Hang Seng index of the Hong Kong stock market from January 3, 1994 to May 28, 1997. The trading time for a trading day in the data was not the same in the whole period. Although for all trading days, the Hong Kong stock market opened from 10:00AM to 12:30PM for the morning session, occupying a time interval of 150 minutes and opened from 2:30PM in the afternoon, the closing times were not the same. For the period from January 3, 1994 to June 30, 1994, the market closed at 3:30PM with the total trading time per day being 210 minutes. For the period from July 1, 1994 to August 30, 1995, the market closed at 3:45PM with the total trading time per day being 225 minutes. For the period from September 1, 1995 to December 30, 1996, the market closed at 3.55PM with the total trading time per day being 235 minutes. Finally, for the period from January 1, 1997 to May 28, 1997, the market closed at 4:00PM with the total trading time per day being 240 minutes. We only count the time during trading hours, and remove the lunch breaks, evenings, weekends and holidays from the data set, *i.e.*, the closing time and the time of the opening of the next session are considered to be neighboring time

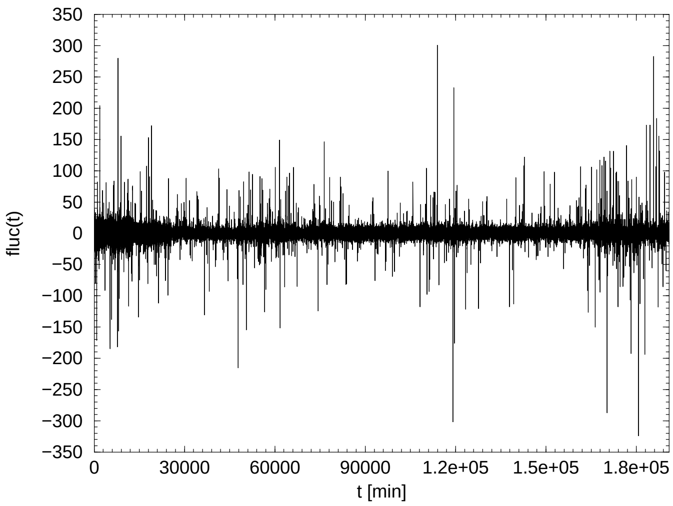

Fig. 2. Time series of 1 minute returns of the Hang Seng index (1994.1–1997.5). The ordinate is the index returns $Z(t) \equiv \mathrm{fluc}(t) = \beta\{\ln[X(t+\Delta t) - \ln[X(t)\}$, where $X(t)$ is the index at time t, $\Delta t = 1$ min, $\beta = 10^4$ is an amplification factor.

steps. A total trading time of 190821 minutes is considered and 190821 data points $X(t)$ of the Hang Seng index are analyzed.

The returns Z(t) of the Hang Seng index (1994.1–1997.5) defined in equation (1) over a sampling time interval $\Delta t = 1$ minute are shown in Figure 2. One can see that large events are frequent in the fluctuations of financial data when compared with a normal process.

In order to characterize quantitatively the observed process, we plot the probability distribution $P(Z)$ of the returns over different time scales Δt. Figure 3a shows eight distributions of returns with Δt taking on Δt=$1, 2, \cdots, 64, 128$ minutes, respectively in semi-log scale. The number of data points in each set decreases from the value of 190820 (for $\Delta t = 1$ minute) to the value of 1491 (for $\Delta t = 128$ minutes). As expected for a random process, the distributions are roughly symmetrical with a wider spread for increasing Δt. It may also be noted that the two tails of the distributions are larger than that of a normal process.

Because larger Δt implies less data points, it is difficult to determine the parameters characterizing the distributions only by investigating the spreads. Hence, we studied [3] the peak values at the center of the distributions, *i.e.*, the probability of zero return $P(Z = 0)$ as the function of Δt. With this choice, we can investigate the point of each probability distribution which is least affected by noise. Figure 3b shows $P(0)$ *versus* Δt in a log-log plot. It can be seen that all the data can be well fitted by a straight line with a slope -0.618 ± 0.025. Thus scaling behavior of non-Gaussian process is observed for Δt from 1 min. to 128 min. This observation agrees with theoretical model leading to a Lévy distribution.

If we assume that the central part of the distribution of returns can be described by a Lévy stable symmetrical distribution with an index α and a parameter γ,

$$P_\alpha(Z, \Delta t) \equiv (1/\pi) \int_0^\infty \mathrm{e}^{-\gamma \Delta t |q|^\alpha} \cos(qZ) \mathrm{d}q, \tag{2}$$

where $\mathrm{e}^{-\gamma \Delta t |q|^\alpha}$ is the characteristic function of a Lévy symmetrical stable process, the probability of zero return is given by

$$P(0) = P_\alpha(0, \Delta t) = \Gamma(1/\alpha)/[\pi\alpha(\gamma\Delta t)^{1/\alpha}], \tag{3}$$

where Γ is the Gamma function. Using the value -0.618 ± 0.025 for the slope of the fitted line to the data (Fig. 3b), we obtain the index $\alpha = 1.619 \pm 0.05$.

To check whether the Lévy scaling can be extended to the entire probability distribution of returns in the Hang Seng index, we notice that under the transformation:

$$Z_\mathrm{s} \equiv Z/[(\Delta t)^{1/\alpha}]$$

and

$$P_\mathrm{s}(Z_\mathrm{s}) \equiv (\Delta t)^{1/\alpha} P_\alpha(Z, \Delta t) = (\Delta t)^{1/\alpha} P_\alpha[(\Delta t)^{1/\alpha} Z_\mathrm{s}, \Delta t], \tag{4}$$

the distributions for different time scales Δt will collapse onto one curve. Figure 3c shows the re-scaled distributions for the same data in Figure 3a in the scaled variables, *i.e.*, $P_s(Z_s)$ *versus* Z_s. Data collapse is evident, except for some data points in the tails for large Δt. The closer to the central point $Z_\mathrm{s} = 0$, the stronger is the extent of data collapse. These observations imply that the Lévy distribution is a better description of the dynamics of the random process underlying the variation of returns in the central part of the probability distribution $P(Z)$ over Δt spanning at least two orders of magnitude.

3 Removing daily pattern

There exists daily pattern of market activities in large financial markets [13–18]. A possible explanation for the daily pattern is the reaction to the information gathered during the hours when the market is closed, together with the fact that many liquidity traders are active near the closing hours [13,14,17]. We observed similar daily pattern in the absolute changes in the Hang Seng index $|Z(t)|$.

In order to quantify the correlation in absolute index changes, it is important to remove this trend. Otherwise one may find peaks in the power spectrum at the frequencies of 1/day and larger [14].

The intra-day pattern $A(t_\mathrm{day})$, where t_day denotes the time in a day, is defined as the average of the absolute index changes at time t_day of a day for all days:

$$A(t_\mathrm{day}) \equiv \sum_{j=1}^{N} |Z^j(t_\mathrm{day})|/N, \tag{5}$$

where the index j runs over all the trading days considered ($N = 838$). Daily pattern $A(t_\mathrm{day})$ for Hang Seng index is

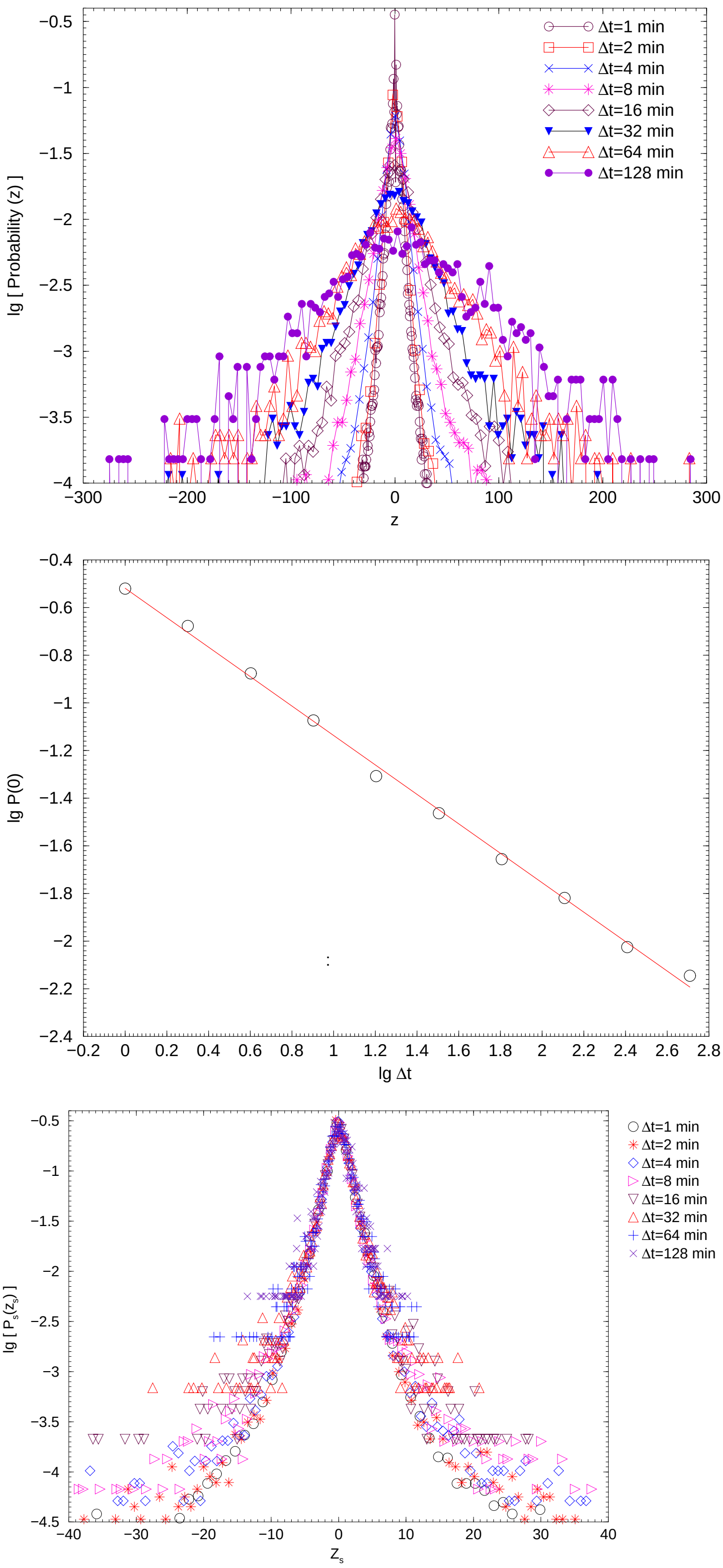

Fig. 3. Probability distributions of the returns and their scaling behavior. (a): The probability distributions of index returns for $\Delta t = 1$, 2, 4, 8, 16, 32, 64, 128 min. (b): The central peak value $P(0)$ as a function of Δt. A power-law behavior is observed. The slope of the best-fit straight line is -0.618 ± 0.025 from which we obtain the scaling exponent $\alpha = 1.619 \pm 0.05$ characterizing the Lévy distribution. (c): Re-scaled plot of the probability distributions shown in (a). Data collapse is evident after using rescaled variables with α =1.619. The abscissa is the re-scaled returns, the ordinate is the logarithm of re-scaled probability.

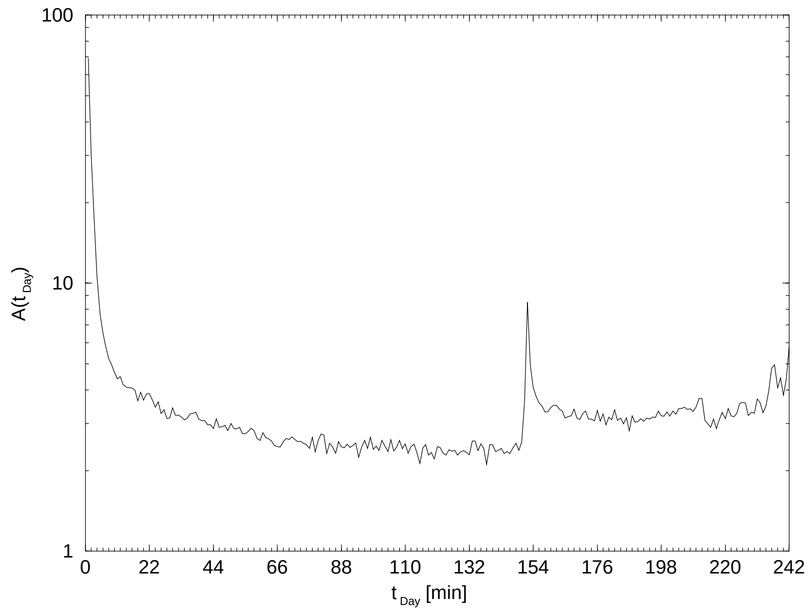

Fig. 4. 1-minute interval daily pattern for the absolute price changes of the Hang Seng index (1994.1–1997.5). The abscissa $t_{\rm day}$ is the time within a trading day. The ordinate $A(t_{\rm day})$ is the average of the absolute index change $|Z(t_{\rm day})|$ at time $t_{\rm day}$ of the day for all days.

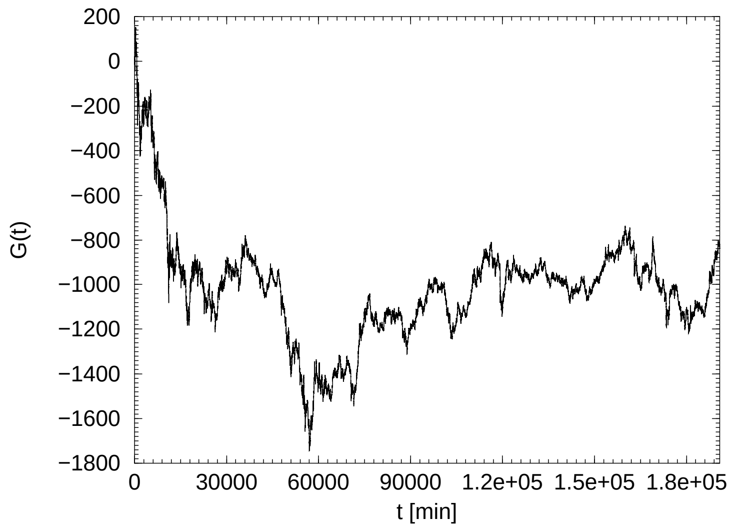

Fig. 5. The integrated price changes of Hang Seng index returns (1994.1–1997.5) after removing daily pattern.

shown in Figure 4, which shows an apparent intra-day oscillating pattern.

Similar pattern has also been observed independently in reference [19].

In order to remove the effects caused by the daily pattern, we investigate the new returns $g(t)$ defined by

$$g(t) \equiv Z(t)/A(t_{\rm day}) \tag{6}$$

instead of the old returns $Z(t)$. The time integration of the new returns

$$G(t) = \int_0^t g(t')\mathrm{d}t' \tag{7}$$

is shown in Figure 5. It can be regarded as a random walk with steps $g(t)$. Figure 6 shows the central part of the distributions after removing the daily pattern for the new returns $g(t)$ for $\Delta t = 1, 2, 4, \cdots, 64, 128$ min. Similar to Figure 3b and c, we may obtain the scaling of the peak values of the probability distributions with different Δt as Figure 3b and achieve data collapse after removal of daily pattern by properly rescaling the variables.

4 Accumulative distribution and truncated Lévy scaling

In order to determine if an exponential truncated Lévy distribution can be used to describe the stochastic process and to investigate the kind of asymptotic behavior

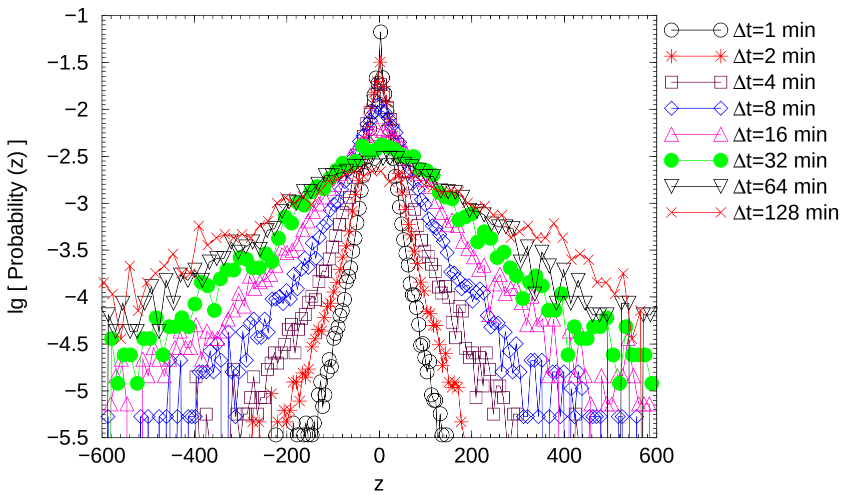

Fig. 6. The probability distributions of the returns $g(t)$ over time intervals $\Delta t = 1, 2, \cdots, 128$ after removing daily pattern.

outside the Lévy stable region, we study the accumulate distribution $P(g > Z)$ of the fluctuations of financial data.

For a stable symmetric Lévy distribution $(0 < \alpha < 2)$, the two tails show a power-law asymptotic behavior

$$P(Z) \sim Z^{-(1+\alpha)}, \quad (8)$$

and hence the second moment diverges. This leads to an asymptotic power-law for the accumulate distribution for both the positive and negative tails [13] in the form

$$P(g > Z) \sim Z^{-\alpha} \quad (9)$$

Figure 7 shows the accumulate probability distribution of returns $P(g > Z)$ for $\Delta t = 1$ min. for the Hang Seng index before and after the removal of daily pattern. For data in the region $10 \le Z \le 200$ before removing daily pattern, regression fits yield $\alpha = 2.32$ (positive tail) and $\alpha = 2.22$ (negative tail). These results appear to be outside the Lévy stable range of $0 < \alpha < 2$. For the data in the region $3 \le Z \le 15$ after removing daily pattern, regression fits yield $\alpha = 5.0$ (positive tail) and $\alpha = 4.0$ (negative tail). These results are further outside the Lévy stable range when compared with those before the daily pattern is removed.

From the above analysis, we conclude that after removing the daily pattern, the two tails in the distribution of the returns in the Hang Seng index fall-off with an exponent well outside the Lévy stable range. The asymptotic behavior of the stochastic process underlying the fluctuations in the Hang Seng index are quite different before and after removing daily pattern. Removing daily pattern is equivalent to decreasing the fluctuations, which leads to a stronger confinement and so makes the second moment of the distribution finite. Thus, the scaling exponent of the asymptotic distribution deviates further from the Lévy stable range. It is therefore important to remove the daily pattern in analyzing the asymptotic behavior of fluctuation distribution.

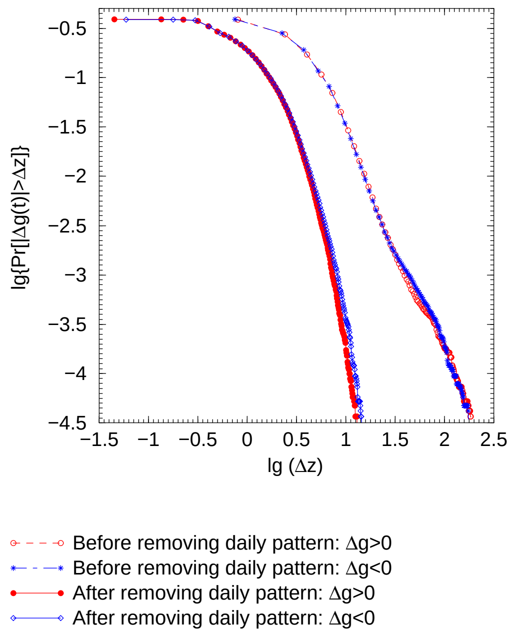

Fig. 7. The accumulate probability distributions $P(g > Z)$ of 1-minute returns for Hang Seng index before and after removing daily pattern. For data before removing daily pattern, regression fits in the region $10 \le Z \le 200$ yield $\alpha = 2.32$ (positive tail) and $\alpha = 2.22$ (negative tail). For the data after removing daily pattern, regression fits in the region $3 \le Z \le 15$ yield $\alpha = 5.0$ (positive tail) and $\alpha = 4.0$ (negative tail).

We are grateful to Dr. Lei-Han Tang for providing us with the data. BHW acknowledges the support from a Special Funds for Major State Basic Research Projects in China (973 Project), the National Basic Research Climbing-up Project "Nonlinear Science", and the National Natural Science Foundation in China (NNSFC) under Key Project Grant No. 19932020 and General Project Grant No. 19974039 and No. 59876039. PMH also acknowledges the support from NNSFC under Grant No. 19932020 and No. 19974039. BHW would like to thank the

Department of Physics at CUHK for its hospitality during a visit when this work was initiated. The visit was supported in part by a Grant (CUHK 4191/97P) from the Research Grants Council of the Hong Kong SAR Government.

References

1. A. Pagen, J. Empirical Finance **3**, 15 (1996).
2. R.N. Mantegna, H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000); R.N. Mantegna, Physica A **179**, 232 (1991).
3. R.N. Mantegna, H.E. Stanley, Nature **376**, 46 (1995).
4. J.-P. Bouchaud, D. Sornette, J. Phys. I France **4**, 863 (1994).
5. A. Arneodo, J.-F. Muzy, D. Sornette, Eur. Phys. J. B **2**, 277 (1998).
6. T. Lux, J. Econ. Behavior Organization **33**, 143 (1998).
7. T. Lux, J. Econ. Dyn. Control **22**, 1 (1997).
8. L. Bachelier, Ann. Sci. École Norm. Sup. **3**, 21 (1900).
9. B.B. Mandelbrot, J. Business **36**, 394 (1963).
10. P. Lévy, *Théorie de l'Addition des Variables Aléatoires* (Gauthier-Villars, Paris, 1937).
11. D. Sornette, A. Johansen, J.-P. Bouchaud, J. Phys. I France **6**, 167 (1996).
12. R.N. Mantegna, H.E. Stanley, Physica A **239**, 255 (1997).
13. P. Gopikrishnan, M. Meyer, L.A.N. Amaral, H.E. Stanley, Eur. Phys. J. B **3**, 139 (1998); P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Meyer, H.E. Stanley, Phys. Rev. E **60**, 5305 (1999).
14. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Phys. Rev. E **60**, 1390 (1999).
15. R.A. Wood, T.H. McInish, J.K. Ord, J. Finance **40**, 723 (1985).
16. L. Harris, J. Financial Economics **16**, 99 (1986).
17. A. Admati, P. Pfleiderer, Rev. Financial Studies **1**, 3 (1988).
18. P.D. Ekman, J. Futures Markets **12**, 365 (1992).
19. Z.F. Huang, Physica A **287**, 405 (2000).