Indication of multiscaling in the volatility return intervals of stock markets

Fengzhong Wang,[1] Kazuko Yamasaki,[1,2] Shlomo Havlin,[1,3] and H. Eugene Stanley[1]
[1]*Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA*
[2]*Department of Environmental Sciences, Tokyo University of Information Sciences, Chiba 265-8501, Japan*
[3]*Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel*
(Received 30 July 2007; published 29 January 2008)

The distribution of the return intervals τ between price volatilities above a threshold height q for financial records has been approximated by a scaling behavior. To explore how accurate is the scaling and therefore understand the underlined nonlinear mechanism, we investigate intraday data sets of 500 stocks which consist of Standard & Poor's 500 index. We show that the cumulative distribution of return intervals has systematic deviations from scaling. We support this finding by studying the m-th moment $\mu_m \equiv \langle (\tau/\langle\tau\rangle)^m \rangle^{1/m}$, which show a certain trend with the mean interval $\langle\tau\rangle$. We generate surrogate records using the Schreiber method, and find that their cumulative distributions almost collapse to a single curve and moments are almost constant for most ranges of $\langle\tau\rangle$. Those substantial differences suggest that nonlinear correlations in the original volatility sequence account for the deviations from a single scaling law. We also find that the original and surrogate records exhibit slight tendencies for short and long $\langle\tau\rangle$, due to the discreteness and finite size effects of the records, respectively. To avoid as possible those effects for testing the multiscaling behavior, we investigate the moments in the range $10 < \langle\tau\rangle \leq 100$, and find that the exponent α from the power law fitting $\mu_m \sim \langle\tau\rangle^{\alpha}$ has a narrow distribution around $\alpha \neq 0$ which depends on m for the 500 stocks. The distribution of α for the surrogate records are very narrow and centered around $\alpha = 0$. This suggests that the return interval distribution exhibits multiscaling behavior due to the nonlinear correlations in the original volatility.

DOI: 10.1103/PhysRevE.77.016109 PACS number(s): 89.65.Gh, 05.45.Tp, 89.75.Da

I. INTRODUCTION

The price dynamics of financial markets has long been a focus of economics and econophysics research [1–9]. Studying the volatility time series is not only crucial for revealing the underlined mechanism of financial markets dynamics, but also useful for traders. For example, it helps traders to estimate the risk and optimize the portfolio [8,10]. The volatility series is known to be long-term power-law correlated [11–22]. To better understand these correlations and characterize temporal scaling features in volatilities, recently Yamasaki *et al.* [23] and Wang *et al.* [24,25] studied the statistics of return intervals τ between volatilities that are above a given threshold q, which is an alternative way to analyze long-term correlated time series (see Ref. [26] and references therein). They find that scaling and memory in the return intervals of daily and intraday financial records are similar to that found in the climate and earthquake data [27–29]. Very recently, some related studies on financial markets, such as escape time [30], exit time [31,32], first passage time [32,33], and level crossing [34], have been performed.

Studies of financial records show that the scaling in the return intervals distribution can be well approximated by a scaling function [23–25]. However, financial time series are known to show complex behavior and are not of uniscaling nature [35] and nonlinear features [36]. Recent studies [37–39] of stock markets show that the distribution of activity measure such as the intertrade time has multiscaling behavior. Thus a detailed analysis of the scaling properties of the volatility return intervals is of interest. It might improve our understanding of the return intervals statistics and shed light on the underlined complex mechanism of the volatility. Our analysis suggests that for all Standard & Poor's index constituents, the cumulative distributions of the return intervals depart slightly but systematically from a single scaling law. We also find that the moments $\mu_m \equiv \langle (\tau/\langle\tau\rangle)^m \rangle^{1/m}$ are consistent with the deviations from scaling. However, using the corresponding surrogate records [40–42] which remove the nonlinearities, μ_m almost does not depend on $\langle\tau\rangle$ and no deviation from scaling occurs. Therefore our results suggest that nonlinear correlations in the volatility account for the deviations from a scaling law.

The paper is organized as follows: In Sec. II we introduce the database and define the volatility. In Sec. III we discuss the scaling and investigate the deviations from scaling in the cumulative distributions of the return intervals. We also describe the stretched exponential form suggested for the distribution and the generation of the surrogate records. Section IV deals with the moments of the return intervals. We quantify the deviation from the scaling that exhibits multiscaling behavior. We simulate the return intervals with different sizes and show the finite size effect for long $\langle\tau\rangle$. We also study the discreteness effect for short $\langle\tau\rangle$ and explore the relation between the moment and its order. In Sec. V we present a discussion.

II. DATABASE

We analyze the Trades And Quotes (TAQ) database from New York Stock Exchange (NYSE), which records every trade for all the securities in United States stock markets. The duration is from Jan. 1, 2001 to Dec. 31, 2002, which has a total of 500 trading days. We study all 500 companies which consist of Standard & Poor's 500 index (S&P 500) [43], the benchmark for American stock markets. The volatility is defined the same as in [24]. First we take the absolute value for the logarithmic price change, then remove the intraday U-shape pattern, and finally normalize it with the standard deviation. Here the price is the closest tick to a minute mark. Thus the sampling time is 1 min and a trading day usually has 391 points after removing the market closing hours. For each stock, the size of data set is about 200 000 records.

III. SCALING IN RETURN INTERVALS

The probability density function (PDF) for the return intervals τ of the financial volatilities is well-approximated by the following form:

$$P_q(\tau) = \frac{1}{\langle\tau\rangle} f(\tau/\langle\tau\rangle), \tag{1}$$

as analyzed by Yamasaki *et al.* [23] and Wang *et al.* [24,25]. Here $\langle\cdot\rangle$ stands for the average over the data set and mean interval $\langle\tau\rangle$ depends on the threshold q. It was suggested that the scaling function can be approximated by a stretched exponential [23–25],

$$f(x) = ce^{-(ax)^\gamma}, \tag{2}$$

for financial records, which is consistent with other long-term correlated records [26–29]. Here a and c are fitting parameters and γ is the exponent characterizing the long-term correlation [26–29]. From the normalization of PDF follows [44]

$$1 = \int_0^\infty P_q(\tau)d\tau = \int_0^\infty \frac{1}{\langle\tau\rangle} ce^{-(a\tau/\langle\tau\rangle)^\gamma} d\tau. \tag{3}$$

From the definition of $\langle\tau\rangle$ follows

$$\langle\tau\rangle = \int_0^\infty \tau P_q(\tau)d\tau = \int_0^\infty \tau \frac{1}{\langle\tau\rangle} ce^{-(a\tau/\langle\tau\rangle)^\gamma} d\tau. \tag{4}$$

Using Eqs. (3) and (4), the parameters a and c can be expressed by γ,

$$a = \Gamma(2/\gamma)/\Gamma(1/\gamma),$$

$$c = \gamma a/\Gamma(1/\gamma) = \gamma\Gamma(2/\gamma)/\Gamma(1/\gamma)^2. \tag{5}$$

Here $\Gamma(a) \equiv \int_0^\infty t^{a-1}e^{-t}dt$ is the Gamma function. Thus if the stretched exponential distribution is valid for the scaled interval $\tau/\langle\tau\rangle$, it is completely determined by γ. For $\gamma=1$, the record has no long-term correlations and the return interval distribution indeed follows an exponential distribution, represented by a Poissonian statistics, as expected.

Though the scaling in the return intervals distribution is a good approximation, we find slight deviations that as shown below are attributed to nonlinear features. To explicitly explore the quality of the scaling in return interval distributions, we study all S&P 500 constituents and show the results of four representative stocks: Citigroup (C), General Electric (GE), Coca Cola (KO), and Exxon Mobil (XOM). All other stocks studied here usually show similar features. First, we examine the cumulative distribution [45] of the scaled intervals,

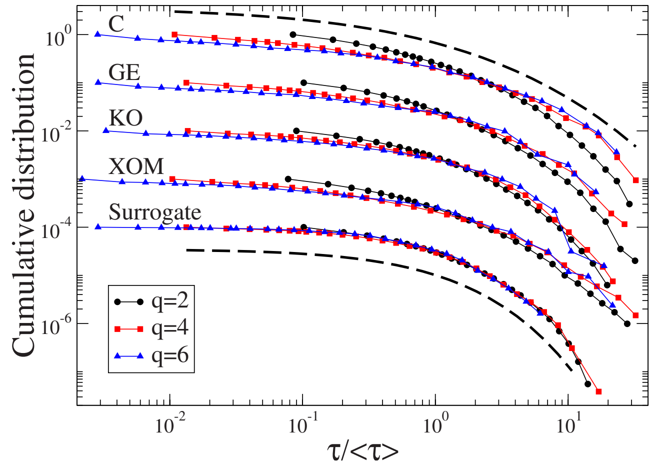

FIG. 1. (Color online) Cumulative distribution [45] of the scaled intervals $\tau/\langle\tau\rangle$ for stock C, GE, KO, XOM, and GE's surrogate. Symbols are for three thresholds, $q=2$ (circles), 4 (squares), and 6 (triangles), respectively. As two examples, we fit the cumulative distribution of stock C ($q=2$) and the surrogate ($q=2$) to a stretched exponential distribution [Eq. (2)] with exponent $\gamma=0.25$ and $\gamma=0.50$ correspondingly. Except for stock C, all symbols and curves are vertically shifted for better visibility. For the surrogate records, symbols collapse almost perfectly to one curve. However, all original data exhibit similar systematic deviations from the scaling. This suggests that the nonlinear correlations in the volatility series affect the scaling in its return intervals.

$$D(\tau/\langle\tau\rangle) \equiv \int_\tau^\infty P_q(\tau)d\tau = \int_{\tau/\langle\tau\rangle}^\infty f(x)dx. \tag{6}$$

If the scaling function $f(\tau/\langle\tau\rangle)$ is valid, the cumulative distributions should also collapse to a single curve. Otherwise, the cumulative distributions, which integrate deviations, may show clearer deviations from scaling. Indeed, in Fig. 1 we show cumulative distributions for three thresholds $q=2$, 4, and 6. Note that the volatility is normalized by its standard deviation, the threshold q is in units of standard deviations, and therefore $q=6$ is a quite large volatility. It is clearly seen that those distributions are close to each other but do not collapse to a single curve. More important, they show apparent deviations from the scaling, which are systematic with the threshold. For small scaled intervals ($\tau/\langle\tau\rangle < 1$), the cumulative distribution decreases with q, while for large scaled intervals ($\tau/\langle\tau\rangle > 1$), it increases with q [46]. In other words, the scaled interval prefers to be larger for higher threshold. This systematic trend suggests multiscaling in the return intervals, which might be related to the nonlinear correlations in the volatility.

To better understand the systematic trends and test if it is not due to finite size effect or discreteness of minutes, we also measure the cumulative distribution of return intervals for surrogate records of volatilities using the Schreiber method [40–42] where nonlinearities are removed. For a given time series, we store the power spectrum and randomly shuffle the sequence, then we apply the following iterations. Each iteration consists of two consecutive steps:

(i) We perform the Fourier transform of the shuffled series, replace its power spectrum with the original one, then take the inverse Fourier transform to achieve a series. This step enforces the desired power spectrum to the series, while the distribution of volatilities usually is modified.

(ii) By ranking, we exchange the values of the resulting series from step (i) with that of the original record. The largest value in the resulting series is replaced by the largest one in the original series, the second largest value is replaced by the second largest one, and so on. This step restores the original distribution but now the power spectrum is changed.

To achieve the convergence to the desired power spectrum and distribution, we repeat these two steps 30 times. In this way, a "surrogate" series is generated. Because of the Wiener-Kinchine theorem [47], the surrogate record has the same linear correlations as the original, as well as the distribution. The only difference is that the original record has the nonlinear correlations (if they exist) but the surrogate does not have any nonlinear features.

In Fig. 1 we also plot the cumulative distribution for the surrogate with the same three thresholds as the original. Since the surrogate records lost the nonlinear correlations, they are similar to each other, we only show results for GE's surrogate. It is seen that the collapse of the surrogate for different q values is significantly better than that of the original and the deviation tendency with the threshold in the original records disappears. This indicates that the scaling deviations in the original are due to the nonlinear correlations in the volatility. To further test this hypothesis, we analyze the moments $\mu_m \equiv \langle (\tau/\langle\tau\rangle)^m\rangle^{1/m}$ in Sec. IV and show similar and consistent deviations from scaling. We also compare our results to the stretched exponential distribution (dashed lines). This curve is very close to the empirical results, in particular for the surrogate records which contain only the linear correlations. This suggests that PDF of return intervals is well approximated by a stretched exponential.

IV. THE MOMENTS OF SCALED INTERVALS

The cumulative distribution shows a clear systematic trend with q, which is difficult to see from the PDF directly [23–28]. To further analyze the systematic tendency in the distribution, we calculate the moments μ_m averaged over a stock data set as a function of $\langle\tau\rangle$, where a mean interval $\langle\tau\rangle$ corresponds to a threshold q and therefore characterizes a return interval series. We study moments for a wide range of $\langle\tau\rangle$, from 3 min (to avoid the artificial effects due to discreteness close to $\tau=1$) to thousands of minutes (a few trading days or even a week). Assuming a single scaling function for the PDF $P_q(\tau)$, Eq. (1), it follows that

$$\mu_m \equiv \langle (\tau/\langle\tau\rangle)^m\rangle^{1/m} = \left\{\int_0^\infty (\tau/\langle\tau\rangle)^m \frac{1}{\langle\tau\rangle} f(\tau/\langle\tau\rangle) d\tau\right\}^{1/m} = \left\{\int_0^\infty x^m f(x) dx\right\}^{1/m}, \qquad (7)$$

which only depends on m and on the form of the scaling function $f(x)$ but independent of $\langle\tau\rangle$. Thus if μ_m depends on $\langle\tau\rangle$, it suggests deviation from the assumption of scaling.

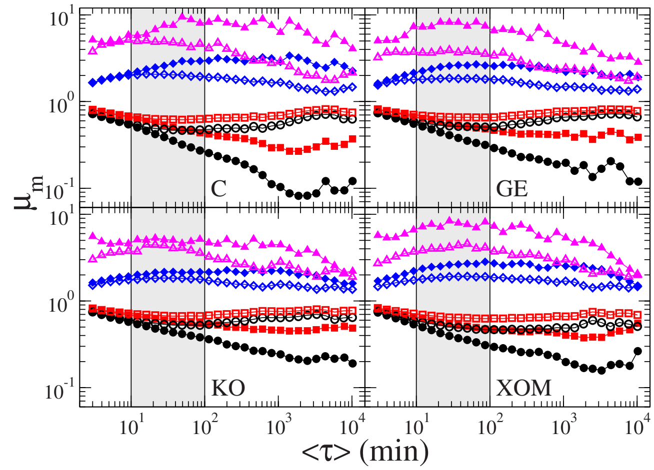

FIG. 2. (Color online) Moment μ_m for the scaled intervals of stock C, GE, KO, and XOM. We show results from the original volatility series (filled symbols) and their surrogate (empty symbols). For each case, four moments, $m=0.25$ (circles), 0.5 (squares), 2 (diamonds), and 4 (triangles), are demonstrated. Symbols for the original are clearly away from the horizontal line and their deviations are much larger than that for surrogate, therefore the nonlinear correlations in the original are related to those deviations. To avoid effects from the resolution and size limit, we choose the shadow area, $10<\langle\tau\rangle\leq 100$ (corresponding to $1.6<q\leq 4.2$ for the 500 stocks), to study the multiscaling behavior.

A. Moments vs mean interval $\langle\tau\rangle$

First we examine the relation between the moments μ_m and the mean interval $\langle\tau\rangle$. Figure 2 shows four representative moments $m=0.25$, 0.5, 2, and 4 for stocks C, GE, KO, and XOM. Ignoring small fluctuations, which is usually due to limited size data, all moments μ_m for the original records deviate significantly from a horizontal line, which is expected for a perfect scaling of the PDF. They depend on $\langle\tau\rangle$ and show some systematic tendency. For $m>1$, moments have similar convex structure, first μ_m increases with $\langle\tau\rangle$ and then decreases, where the crossover starts earlier for larger m. For $m<1$, moments also show a similar tendency but in the opposite direction compared to $m>1$. These deviations from scaling in μ_m are consistent with the deviations seen in the cumulative distributions shown in Fig. 1. Moments of large m $(m>1)$ represent large $\tau/\langle\tau\rangle$ in the PDF and they initially (for $\langle\tau\rangle\leq 100$) increase with $\langle\tau\rangle$ ($\langle\tau\rangle$ increases with the threshold q), while moments of small m $(m<1)$ represent small $\tau/\langle\tau\rangle$ and they initially (for $\langle\tau\rangle\leq 100$) decrease with $\langle\tau\rangle$.

To further test if the systematic deviations are not due to finite size effects and discreteness, we also examine moments for the surrogate records which are more flat for most ranges, as shown in Fig. 2. For the same order m, the moment of the surrogate obviously differs from that of the original, especially in the medium range of $\langle\tau\rangle$ $(10<\langle\tau\rangle\leq 100)$. This discrepancy suggests that the nonlinear correlations exist in the original volatility and accounts for the scaling deviations. Nevertheless, all moments of surrogate show small curvature from a perfect straight line at both short and long $\langle\tau\rangle$, which are much weaker compared to the original records. The weak curvature suggests that some additional effects, not related to the nonlinear correlations, affect the moments. For small $\langle\tau\rangle$, the resolution discrete limit seems to have some influence on the moments. We will discuss this effect in Sec. IV C. For large $\langle\tau\rangle$, the moments are gradually approaching the horizontal line and are more fluctuating, the effect seems to be related to limited size of the record. On the other hand, for moments of high order such as $m=4$, the trend in the medium range is not as obvious as for the lower order moments. This is probably due to a lack of statistics for high order moments. This effect will be discussed further in Sec. IV B.

B. Multiscaling

For the original volatility records, the systematic tendency in the distribution of τ and the moments implies that the return intervals may have multiscaling features. To avoid as much as possible the effect of discreteness and finite size, we calculate the moments only for some medium range of $\langle\tau\rangle$ where the effects are small. Since there is no nonlinear correlations in the surrogate records, the curvature in their moments is only due to the additional effects, we use the surrogate curve as our reference. For small $\langle\tau\rangle$, the increasing (decreasing) range for $m>1$ $(m<1)$ almost ends at $\langle\tau\rangle=10$ min. For large $\langle\tau\rangle$, the curves start to decrease (increase) from different positions, but at $\langle\tau\rangle=100$, all curves do not or just start to decrease (increase). Thus we choose to study μ_m in the region, $10<\langle\tau\rangle\leq 100$, represented by the shadow areas in Fig. 2. Note that a certain $\langle\tau\rangle$ represents a certain threshold q. For the 500 constituent stocks of S&P 500, $\langle\tau\rangle=10$ corresponds to $q=1.6\pm 0.2$ and $\langle\tau\rangle=100$ corresponds to $q=4.2\pm 0.3$. In this range, we find a clear trend for the original records while the surrogate is almost horizontal. To quantify the tendency, we fit the moments with a power-law,

$$\mu_m \sim \langle\tau\rangle^{\alpha}. \tag{8}$$

If the distribution of $\tau/\langle\tau\rangle$ follows a scaling law, the exponent α should be some value very close to 0. If α is significantly different from 0, it suggests multiscaling.

To examine the multiscaling behavior for the whole market, we calculate α for all 500 stocks of S&P 500 constituents and plot the histogram for $m=0.25$ to 2. Figure 3 shows that each histogram has a narrow distribution, which suggests that α are similar for the 500 stocks. For the original records, almost all α significantly differ from 0, thus the moments clearly depend on the mean interval. Moreover, the mean value of α shifts with order m from $\langle\alpha\rangle=-0.22\pm 0.08$ for $m=0.25$ to $\langle\alpha\rangle=0.08\pm 0.05$ for $m=2$ which means the dependence varies with the order m. This behavior suggests multiscaling in the return intervals distribution. Indeed, histograms for the surrogate records are more centered around values close to $\alpha=0$ (from $\langle\alpha\rangle=-0.07\pm 0.04$ for $m=0.25$ to $\langle\alpha\rangle=0.00\pm 0.02$ for $m=2$). The uniscaling behavior for the surrogate suggests that the nonlinear correlations in the volatility are responsible for the multiscaling behavior in the original.

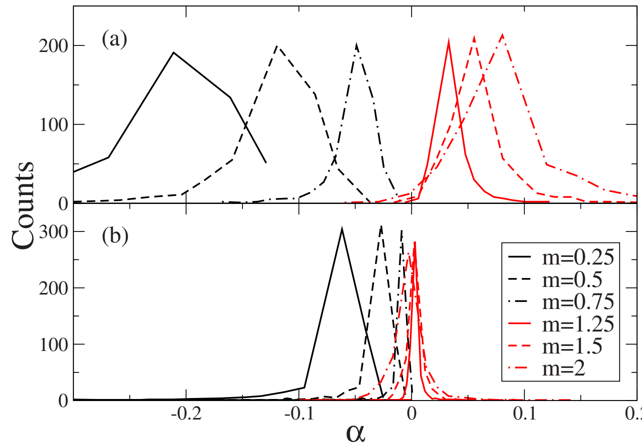

FIG. 3. (Color online) Distribution of multiscaling exponent α for S&P 500 constituents. The exponent α is obtained from the power-law fitting for moments in the medium range $10<\langle\tau\rangle\leq 100$. (a) Histogram of α for the original volatility and (b) for surrogate. The distributions have a systematic shift with m in (a) while all of them almost collapse in (b). This suggests the multiscaling behavior in the return intervals in the original records. The significant discrepancy between the original and the surrogate records manifests that nonlinear correlations from the original volatility account for the multiscaling behavior.

To remove fluctuations and show the tendency more clearly, we plot the dependence of $\langle\alpha\rangle$ on m, where $\langle\alpha\rangle$ is the average α over all 500 stocks. In Fig. 4 we show this relation for a wide range of m, $0.1\leq m\leq 10$, and the plot shows two different behaviors. For small m (roughly $m\leq 2$), $\langle\alpha\rangle$ for the original records clearly deviates from 0 and demonstrates the multiscaling behavior, while $\langle\alpha\rangle$ for the surrogate is closer to 0. For large m $(m>2)$, the two curves have a similar decreasing trend. Since large τ dominates high order moments, this similarity may be due to finite size effects. To test the finite size effects we simulate surrogate return intervals by assuming an independent and identically distributed (i.i.d.) process of stretched exponential distribution, with three sizes (number of all τ in the series), 2×10^6, 2×10^5 (the size of the empirical data set), and 2×10^4. Without loss of generality, we choose $\gamma=0.3$, which is the correlation exponent for GE of $q=2$. To be consistent with the 500 stocks, we perform 500 realizations and plot their average exponent $\langle\alpha\rangle$. As shown in the inset of Fig. 4, all $\langle\alpha\rangle$ show a similar decreasing trend as that of the empirical curves. However, it is seen that the trend starts earlier for smaller size, and thus the size limit has a strong influence on high order moments. Figure 4 also shows the error bars for the two records, which is the standard deviation of 500 α values. Note that the error bars for the volatility records do not overlap those of their corresponding surrogate, indicating the significance of our results.

C. Discreteness effect

For small $\langle\tau\rangle$ ($\langle\tau\rangle\leq 10$), the behavior of μ_m as a function of $\langle\tau\rangle$ was attributed to the discreteness. Here we examine

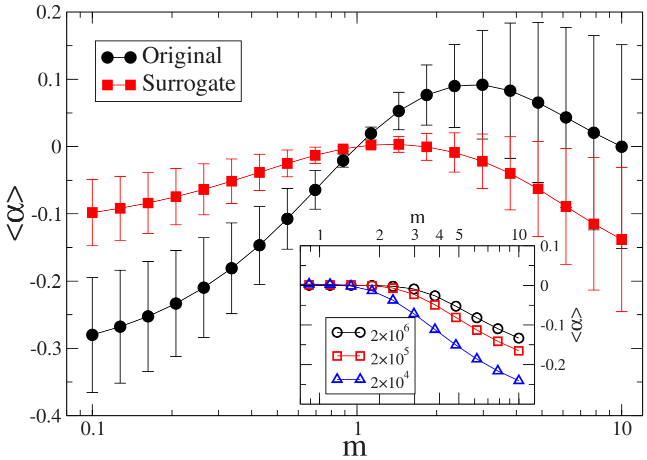

FIG. 4. (Color online) Dependence of average multiscaling exponent $\langle\alpha\rangle$ on order m. The average $\langle\alpha\rangle$ was taken over the 500 α of the S&P 500 constituents. The error bars are standard deviations over the 500 α. Results for the original (circles) and surrogate records (squares) are displayed. For large m, the two curves have a similar tendency which is attributed to the finite size effects. For small m, the two curves are significantly different which supports the multiscaling in the return intervals, due to the nonlinear correlations in the original volatility. The inset demonstrates $\langle\alpha\rangle$ averaged over 500 stretched exponential distributed i.i.d. simulations with $\gamma=0.3$. Three sizes, 2×10^6, 2×10^5, and 2×10^4 are displayed, which clearly shows the finite size effect.

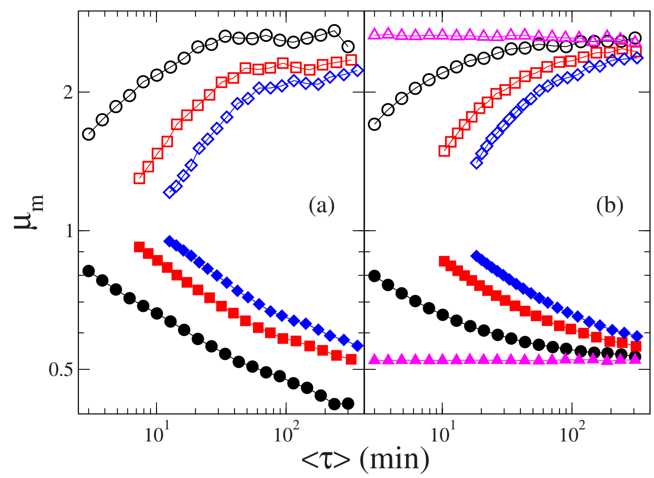

FIG. 5. (Color online) Discreteness effect in the moments. (a) Moments for stock GE with three resolutions, 1 (circles), 5 (squares), and 10 min (diamonds). Filled symbols are for $m=0.5$ while empty symbols are for $m=2$. (b) Moments of artificial records averaged over 100 simulations. For each trial, we simulate the return intervals which follow a stretched exponential distribution with $\gamma=0.3$ and the length of 200 000 points. Symbols are similar to that in (a). Three resolutions, 1, 5, and 10 time units are displayed. To show the disappearing of the discreteness, we also plot simulation results for continuous return intervals (triangles), which exhibit a constant moment, independent of $\langle\tau\rangle$.

this effect. Due to the limits in recording, we cannot have a continuous but discrete record. In our study the volatility is recorded in 1 min. The relative errors in moments will be considerably large for small $\langle\tau\rangle$ close to 1 min. By starting from $\langle\tau\rangle=3$, we only partially avoid the discreteness in the moments. To test the discreteness effects, since we cannot increase the resolution, we reduce it and compare the moments μ_m with three resolutions, 1, 5, and 10 min. Figure 5(a) shows this comparison for GE for $m=0.5$ and $m=2$. The three resolutions have a similar trend, showing that the curves become flatter for the higher resolution. For other stocks, we find similar behavior. This systematic tendency suggests that the recording limit (1 min) strongly affects the moments at short $\langle\tau\rangle$. To reduce it, we should raise the recording precision or study the moments of larger $\langle\tau\rangle$.

To further test this result, we simulate artificial return intervals with an i.i.d. process from stretched exponential distribution with $\gamma=0.3$, the same as the empirical γ for GE of $q=2$. We examine moments of $m=0.5$ and $m=2$ with the same three resolutions (1, 5, and 10 time units) as in the empirical test done above. The simulated size is 200 000 points for each trial and we use the average over 100 trials for each resolution. Figure 5(b) shows curves similar to that of empirical [Fig. 5(a)]. For the higher resolution, the curve is closer to the horizontal line and finally may reach the line when we raise the resolution high enough. To show this, we also simulate continuous return intervals and find constant moments, as expected [Fig. 5(b)]. Therefore the discreteness effects can be overcome if the resolution is improved enough. Note that for the empirical data, we expect the moment not to be constant for small $\langle\tau\rangle$ even if we have a much better resolution, since the return intervals has the multiscaling behavior, as shown for larger $\langle\tau\rangle$ in the range $10<\langle\tau\rangle\leq100$ which is not affected by discreteness.

D. Moments vs order m

The moments μ_m have systematic dependence on m, as seen in Figs. 2 and 3 where the moments are plotted as the function of $\langle\tau\rangle$. It is of interest to explore the relation between the moments and m directly. For a fixed $\langle\tau\rangle$, representing a given threshold q, one can study the return intervals and their moments of various orders which exhibit information on different scales of τ. Moments of large m represent large τ and vice versa. If $\tau/\langle\tau\rangle$ follows a single distribution without corrections due to effects such as discreteness and finite size, curves of μ_m vs m for different $\langle\tau\rangle$ should collapse to a single one, which only depends on the scaling function $f(x)$ from Eq. (7). In Fig. 6 we plot μ_m vs m for both the original and surrogate records. We plot μ_m for m between 0.1 and 10 for three $\langle\tau\rangle$ values: 10, 80, and 400 min. For the original [Fig. 6(a)], there is substantial deviations from a single curve. This supports our suggestion that the return intervals has multiscaling behavior. Moments for the surrogate [Fig. 6(b)] converge to a single curve for $m\leq2$ but become diverse for high orders, which agrees with the strong influence of the finite size effects. As a reference, we also plot the analytical moments [Fig. 6(c)] from the stretched exponential distribution. Substituting Eq. (2) into Eq. (7), we obtain

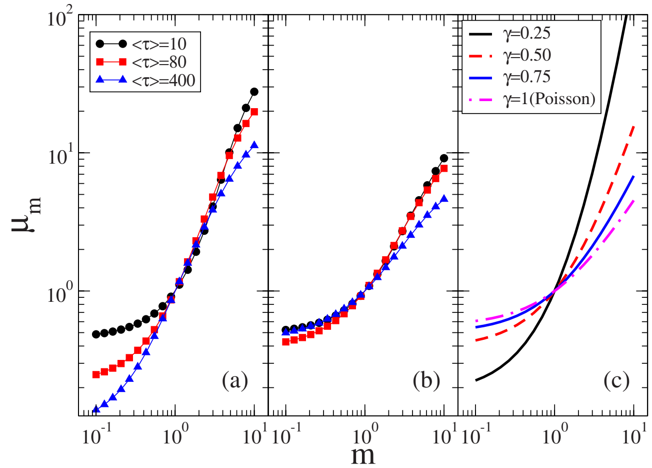

FIG. 6. (Color online) Dependence of the moment μ_m on the order m: (a) for the return intervals of the original volatility for stock GE; (b) for its surrogate. Three mean interval $\langle\tau\rangle=10$ (circles), 80 (squares), and 400 min (triangles) are demonstrated in (a) and (b). (c) Analytical moments from stretched exponential distributions, $\gamma=0.25$, 0.50, 0.75, and 1 (Poisson distribution) taken from Eq. (9). For large m, both (a) and (b) show discrepancies related to the finite size effect. For small m, the difference between (a) and (b) is due to the nonlinear correlations in the original volatility. Compared to the analytical curves in (c), the return intervals for the original records show multiscaling behavior.

$$\mu_m = \frac{1}{a}\left\{\frac{\Gamma((m+1)/\gamma)}{\Gamma(1/\gamma)}\right\}^{1/m}. \tag{9}$$

Figure 6(c) shows analytical curves for various correlation exponent γ.

V. DISCUSSION

We study the scaling properties of the distribution of the volatility return intervals for all S&P 500 constituents. We find small but systematic deviations from scaling assumption with the threshold q in the cumulative distribution. Compared to the good collapse for the surrogate records where nonlinearities are removed, this suggests that the origin of this trend is due to nonlinear correlations in the original volatility. The nonlinearities in the volatilities may originate from the nonlinear interactions between the many market participants [48,49]. Moreover, we find similar systematic deviations for the moments μ_m, which are also attributed to the nonlinear correlations in the volatility. We distinguish these deviations from the deviations due to the discreteness for small $\langle\tau\rangle$ and finite size effect for large $\langle\tau\rangle$. Further, we explore the dependence of the moment μ_m on its order m. When compared to surrogate records and to analytical curves, the results support the multiscaling hypothesis of the return intervals. Thus the scaling assumption in the return interval distributions, although it is a good approximation, cannot be exact. To further understand the volatility and its return intervals, we also tried to simulate them using theoretical models such as fractional Brownian motion and fractionally integrated generalized auto regressive conditional heteroscedasticity (FIGARCH) [50]. Also, the stretched exponential form of the scaling function can only be an approximation. Recently Eisler *et al.* [38,39] exhibited that the distribution of intertrade times has similar multiscaling behavior and the market activity depends on the company capitalization. It would be interesting to connect the intertrade times with the return intervals and test size dependence in the return intervals. The multiscaling behavior in the return intervals may be related to the different underlying mechanics on different time scales. The behavior on different stocks may reflect those mechanics differently therefore a comprehensive examination on many companies provides another view on the multiscaling.

ACKNOWLEDGMENTS

We thank Y. Ashkenazy for his kind help in the simulations, R. Mantegna, J. Kertész, and Z. Eisler for fruitful discussions, and the NSF and the Merck Foundation for financial support.

[1] B. B. Mandelbrot, J. Business **36**, 394 (1963).
[2] R. N. Mantegna and H. E. Stanley, Nature (London) **376**, 46 (1995).
[3] *Econophysics: An Emerging Science*, edited by I. Kondor and J. Kertész (Kluwer, Dordrecht, 1999).
[4] R. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, England, 2000).
[5] H. Takayasu, H. Miura, T. Hirabayashi, and K. Hamada, Physica A **184**, 127 (1992); H. Takayasu, A. H. Sato, and M. Takayasu, Phys. Rev. Lett. **79**, 966 (1997); H. Takayasu and K. Okuyama, Fractals **6**, 67 (1998).
[6] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Phys. Rev. E **60**, 1390 (1999); V. Plerou, P. Gopikrishnan, X. Gabaix, L. A. Nunes Amaral, and H. E. Stanley, Quant. Finance **1**, 262 (2001); V. Plerou, P. Gopikrishnan, and H. E. Stanley, Phys. Rev. E **71**, 046131 (2005).
[7] P. Weber, F. Wang, I. Vodenska-Chitkushev, S. Havlin, and H. E. Stanley, Phys. Rev. E **76**, 016109 (2007).
[8] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge University Press, Cambridge, England, 2003).
[9] H. E. Stanley, Physica A **387**, 537 (2008).
[10] N. F. Johnson, P. Jefferies, and P. M. Hui, *Financial Market Complexity* (Oxford University Press, New York, 2003).
[11] Z. Ding, C. W. J. Granger, and R. F. Engle, J. Empirical Finance **1**, 83 (1993).
[12] R. A. Wood, T. H. McInish, and J. K. Ord, J. Financ. **40**, 723 (1985).

[13] L. Harris, J. Financ. Econ. **16**, 99 (1986).
[14] A. Admati and P. Pfleiderer, Rev. Financ. Stud. **1**, 3 (1988).
[15] G. W. Schwert, J. Financ. **44**, 1115 (1989); K. Chan, K. C. Chan, and G. A. Karolyi, Rev. Financ. Stud. **4**, 657 (1991); T. Bollerslev, R. Y. Chou, and K. F. Kroner, J. Econometr. **52**, 5 (1992); A. R. Gallant, P. E. Rossi, and G. Tauchen, Rev. Financ. Stud. **5**, 199 (1992); B. Le Baron, J. Business **65**, 199 (1992).
[16] M. M. Dacorogna, U. A. Muller, R. J. Nagler, R. B. Olsen, and O. V. Pictet, J. Int. Money Finance **12**, 413 (1993).
[17] A. Pagan, J. Empirical Finance **3**, 15 (1996).
[18] C. W. J. Granger and Z. Ding, J. Econometr. **73**, 61 (1996).
[19] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Physica A **245**, 437 (1997).
[20] R. Cont, Ph.D. thesis, Université de Paris XI, Orsay, 1998; see also R. Cont, e-print arXiv:cond-mat/9705075.
[21] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley, Physica A **245**, 441 (1997).
[22] M. Pasquini and M. Serva, Econ. Lett. **65**, 275 (1999).
[23] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley, Proc. Natl. Acad. Sci. USA **102**, 9424 (2005).
[24] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, Phys. Rev. E **73**, 026117 (2006).
[25] F. Wang, P. Weber, K. Yamasaki, S. Havlin, and H. E. Stanley, Eur. Phys. J. B **55**, 123 (2007) (special issue on econophysics).
[26] E. G. Altmann and H. Kantz, Phys. Rev. E **71**, 056106 (2005).
[27] A. Bunde, J. F. Eichner, S. Havlin, and J. W. Kantelhardt, Physica A **342**, 308 (2004).
[28] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).
[29] V. N. Livina, S. Havlin, and A. Bunde, Phys. Rev. Lett. **95**, 208501 (2005).
[30] G. Bonanno, D. Valenti, and B. Spagnolo, Phys. Rev. E **75**, 016106 (2007).
[31] M. Montero, J. Perelló, J. Masoliver, F. Lillo, S. Miccichè, and R. N. Mantegna, Phys. Rev. E **72**, 056101 (2005).
[32] J. Masoliver, M. Montero, and J. Perelló, Phys. Rev. E **71**, 056130 (2005).
[33] J. Masoliver and J. Perelló, Phys. Rev. E **75**, 046110 (2007).
[34] G. R. Jafari, M. S. Movahed, S. M. Fazeli, M. Reza Rahimi Tabar, and S. F. Masoudi, J. Stat. Mech.: Theory Exp. 2006, P06008.
[35] T. Di Matteo, Quant. Finance **7**, 21 (2007).
[36] C. Q. Cao and R. S. Tsay, J. Appl. Econ. **7**, S165 (1992).
[37] P. Ch. Ivanov, A. Yuen, B. Podobnik, and Y. Lee, Phys. Rev. E **69**, 056107 (2004).
[38] Z. Eisler and J. Kertész, Phys. Rev. E **73**, 046109 (2006).
[39] Z. Eisler and J. Kertész, Eur. Phys. J. B **51**, 145 (2006).
[40] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).
[41] H. A. Makse, S. Havlin, M. Schwartz, and H. E. Stanley, Phys. Rev. E **53**, 5445 (1996).
[42] T. Schreiber and A. Schmitz, Physica D **142**, 346 (2000).
[43] Note that S&P 500 constituents changes a small fraction every year. To fix datasets, we choose 500 stocks which have the longest time in the index for the years 2001 and 2002.
[44] Here we assume that the scaled interval $\tau/\langle\tau\rangle$ is continuous. This assumption is not precisely accurate since the return intervals are discrete. However, the minimum scaled interval usually is not very large. For example, $\langle\tau\rangle=9.8$ for GE return interval of $q=2$, the corresponding minimum is $1/\langle\tau\rangle\simeq 0.1$. For higher q such as $q=6$, $\langle\tau\rangle=347$ and the minimum $1/\langle\tau\rangle \simeq 0.003$ is even smaller.
[45] In this paper the cumulative distribution of the return intervals is actually the complementary cumulative distribution. For simplicity we call it "cumulative distribution."
[46] Note that for very large scaled intervals, $\tau/\langle\tau\rangle \gg 1$, the curves have apparent fluctuations, which cannot be trusted as much as that of smaller scaled intervals, due to poor statistics.
[47] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (revised and enlarged edition) (North-Holland, Amsterdam, 1992).
[48] W. Enders, *Applied Econometric Time Series* (Wiley, New York, 1995).
[49] A. Gaunersdorfer and C. Hommes, in *Long Memory in Economics*, edited by G. Teyssiere and A. Kirman (Springer, Berlin, 2005), p. 265.
[50] I. Vodenska-Chitkushev, F. Wang, P. Weber, K. Yamasaki, S. Havlin, and H. E. Stanley, Eur. Phys. J. B (to be published).