Estimator of a non-Gaussian parameter in multiplicative log-normal models

Ken Kiyono,[1] Zbigniew R. Struzik,[2] and Yoshiharu Yamamoto[2]
[1]*College of Engineering, Nihon University, 1 Naka-gawara, Tokusada, Tamura-machi, Koriyama City, Fukushima 963-8642, Japan*
[2]*Educational Physiology Laboratory, Graduate School of Education, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received 29 May 2007; revised manuscript received 16 August 2007; published 8 October 2007)

We study non-Gaussian probability density functions (PDF's) of multiplicative log-normal models in which the multiplication of Gaussian and log-normally distributed random variables is considered. To describe the PDF of the velocity difference between two points in fully developed turbulent flows, the non-Gaussian PDF model was originally introduced by Castaing *et al.* [Physica D **46**, 177 (1990)]. In practical applications, an experimental PDF is approximated with Castaing's model by tuning a single non-Gaussian parameter, which corresponds to the logarithmic variance of the log-normally distributed variable in the model. In this paper, we propose an estimator of the non-Gaussian parameter based on the qth order absolute moments. To test the estimator, we introduce two types of stochastic processes within the framework of the multiplicative log-normal model. One is a sequence of independent and identically distributed random variables. The other is a log-normal cascade-type multiplicative process. By analyzing the numerically generated time series, we demonstrate that the estimator can reliably determine the theoretical value of the non-Gaussian parameter. Scale dependence of the non-Gaussian parameter in multiplicative log-normal models is also studied, both analytically and numerically. As an application of the estimator, we demonstrate that non-Gaussian PDF's observed in the S&P500 index fluctuations are well described by the multiplicative log-normal model.

DOI: 10.1103/PhysRevE.76.041113 PACS number(s): 05.40.−a, 02.50.−r, 87.15.Ya, 47.11.St

I. INTRODUCTION

To describe the probability density function (PDF) of the velocity difference between two points in fully developed turbulent flows, Castaing *et al.* [1] introduced the following equation based on a log-normal cascade model (paradigm):

$$P_{\lambda,\sigma_0}(x) = \int_0^\infty \frac{1}{\sqrt{2\pi}\lambda} \exp\left(-\frac{\ln^2(\sigma/\sigma_0)}{2\lambda^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) d(\ln \sigma), \quad (1)$$

where λ and σ_0 are positive parameters. By taking the limit $\lambda \to 0$ in Eq. (1), a Gaussian distribution is obtained. On the other hand, the larger value of λ results in fatter non-Gaussian tails of the PDF [Fig. 2(a)]. The other parameter σ_0 only affects the standard deviation of $P_{\lambda,\sigma_0}(x)$. If $\sigma_0=1$ and $\lambda>0$, the variance of $P_{\lambda,1}(x)$ is greater than one. Therefore, to describe a standardized PDF with zero mean and unit variance, we have to adjust the value of σ_0 in Eq. (1).

It is important to note that the PDF described by Eq. (1) does not always imply the existence of a cascade process such as that used in modeling turbulence. If a stochastic process $\{x_i\}$ can be described as

$$x_i = \xi_i e^{\omega_i}, \quad (2)$$

where ξ is a Gaussian variable with zero mean and ω is also a Gaussian independent of ξ, the PDF of $\{x_i\}$ has the same functional form as Eq. (1). In this case, it is possible to assume that $\{\xi_i\}$ and $\{\omega_i\}$ are both uncorrelated random variables, although the long range correlation of the ω_i plays an important role in generating intermittency [2,3]. In this paper, we refer to a stochastic process exhibiting the non-Gaussian PDF described by Eq. (1) as a multiplicative log-normal model in a broad sense.

Although Castaing's equation [Eq. (1)] was originally introduced to study fully developed turbulence [1,4], it has been demonstrated that this equation provides a good approximation of non-Gaussian PDF's observed not only in hydrodynamic turbulence, but also in diverse fields such as solar wind [5], foreign exchange rate [6], stock index [7], and human heartbeat [8,9] fluctuations. Hence, its increasingly widespread application is foreseen.

In practical applications of Eq. (1), the non-Gaussian parameter λ has been estimated from a fitting algorithm to minimize the χ^2 statistic between the observed and numerically integrated PDF's. In this approach, it is very difficult to estimate the small value of λ (<0.1), because multiplications of extremely large and almost zero values are computed in the numerical integration of Eq. (1). In this case, the accuracy of the estimation strongly depends on the algorithm of the numerical integration. In addition, we have to repeat the numerical integration until the χ^2 statistic reaches the minimum value, which makes the computation time very long. To avoid such problems of the fitting procedure, we here propose an estimator of λ based on qth order absolute moments.

II. RELATION BETWEEN THE NON-GAUSSIAN PARAMETER λ^2 AND qTH ORDER ABSOLUTE MOMENTS

We can calculate the standard deviation σ_1 of $P_{\lambda,\sigma_0}(x)$ [Eq. (1)] with $\sigma_0=1$ as

$$\sigma_1 = \exp(\lambda^2).$$

Thus, the standardized PDF of $P_{\lambda,\sigma_0}(x)$ [Eq. (1)] is given by

 ©2007 The American Physical Society

$$f_\lambda(x) = \int_0^\infty \frac{1}{\sqrt{2\pi}\lambda} \exp\left(-\frac{(\ln\sigma + \lambda^2)^2}{2\lambda^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) d(\ln\sigma), \quad (3)$$

which depends only on a single parameter λ. In this case, the relation between σ_0 and λ is given by

$$\sigma_0 = \exp(-\lambda^2).$$

Using Eq. (3), we can calculate the qth order absolute moments

$$\langle |x|^q \rangle = \int_{-\infty}^{\infty} |x|^q f_\lambda(x) dx = \frac{2^{q/2}\Gamma\left(\frac{q+1}{2}\right)}{\sqrt{\pi}} \exp\left(\frac{q(q-2)\lambda^2}{2}\right), \quad (4)$$

where the angular brackets mean the statistical average, and Γ is the gamma function and $q > -1$. From Eq. (4), we obtain the estimator $\hat{\lambda}_q^2$ of the non-Gaussian parameter λ^2 as

$$\hat{\lambda}_q^2 = \frac{2}{q(q-2)}\left[\ln\left(\frac{\sqrt{\pi}\langle |x|^q\rangle}{2^{q/2}}\right) - \ln\Gamma\left(\frac{q+1}{2}\right)\right], \quad (5)$$

where $q \neq 0, 2$. In practical time series analysis, the $\langle |x|^q \rangle$ is estimated from a time series $\{x_i\}$ with zero mean and unit variance.

III. ILLUSTRATIVE EXAMPLES OF MULTIPLICATIVE LOG-NORMAL MODELS

To test the estimator $\hat{\lambda}_q^2$ [Eq. (5)], we introduce stochastic processes within the framework of Eq. (2). By analyzing the numerically generated data set of the stochastic processes, we demonstrate that the estimator $\hat{\lambda}_q^2$ [Eq. (5)] can reliably determine the theoretical value of the non-Gaussian parameter.

A. Independent and identically distributed non-Gaussian variables

A stochastic process $\{x_i\}$ $(i=1,2,3,\ldots)$ exhibiting the non-Gaussian PDF of $f_\lambda(x)$ [Eq. (3)] is described by

$$x_i = \xi_i e^{\lambda\omega_i - \lambda^2}, \quad (6)$$

where $\{\xi_i\}$ and $\{\omega_i\}$ are both sequences of Gaussian white noise with zero mean and unit variance, and independent of each other; and λ is the non-Gaussian parameter in Eq. (3). Equation (6) is derived by the standardization of Eq. (2). Note that the x_i is a Gaussian variable when $\lambda = 0$. Similarly, Eq. (3) converges to a Gaussian in the limit $\lambda \to 0$.

Example time series are shown in Fig. 1. As the non-Gaussian parameter λ increases, the occurrence of large deviations from the average value becomes more pronounced, and the tails of the corresponding PDF become increasingly stretched (Fig. 2).

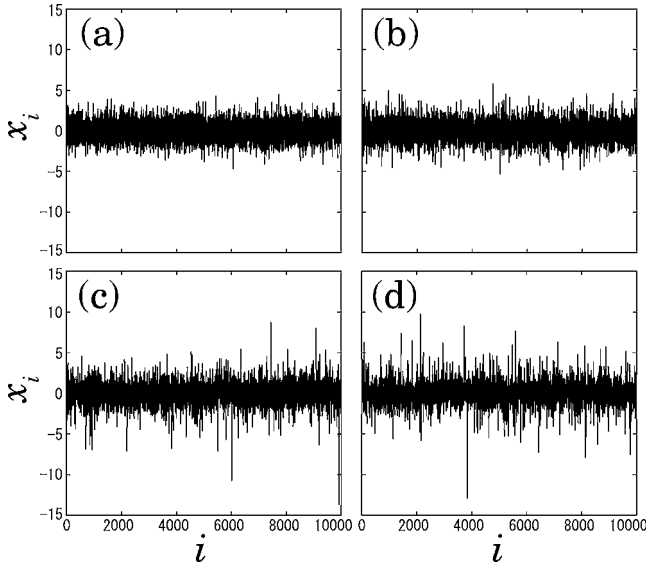

FIG. 1. Time series of independent and identically distributed non-Gaussian variables with zero mean and unit variance [Eq. (6)]. (a) $\lambda = 0.2$. (b) $\lambda = 0.4$. (c) $\lambda = 0.6$. (d) $\lambda = 0.8$.

To test the estimator $\hat{\lambda}_q^2$ [Eq. (5)], we generate non-Gaussian data sets using Eq. (6), and then compute the value of the estimator $\hat{\lambda}_q^2$. Figure 3 shows the results of the estimation averaged over 200 samples. In this case, a constant value of $\hat{\lambda}_q^2$ independent of q is expected from Eq. (5). If the data length n is sufficiently long [Figs. 3(a) and 3(b)], the estimator $\hat{\lambda}_q^2$ provides a good estimation until higher order q. Figure 3(c) demonstrates that the $\hat{\lambda}_q^2$ in the range $0 < q < 1$ still provides a good estimation when the data length $n \sim 10^4$.

To obtain the optimal estimation of the non-Gaussian parameter λ^2, we have to choose the order q of $\hat{\lambda}_q^2$. To do this, we study numerically the error and asymptotic properties of the estimator $\hat{\lambda}_q^2$. In statistics, to judge the quality of an esti-

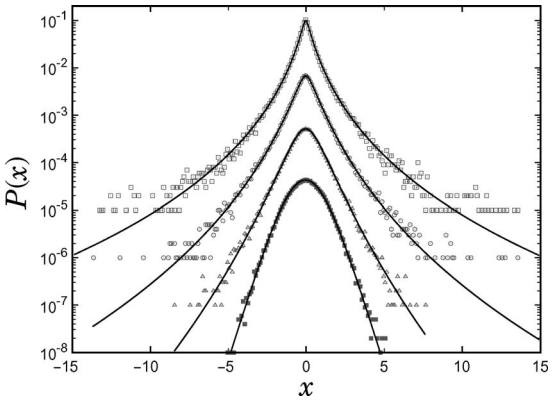

FIG. 2. Probability density functions (PDF's) of Eq. (3) with the non-Gaussian parameters $\lambda = 0.2$, $\lambda = 0.4$, $\lambda = 0.6$, $\lambda = 0.8$ (from bottom to top). Solid lines: numerical integration of Eq. (3). Symbols: estimated PDF's from the time series shown in Fig. 1. The PDF's are shifted in vertical directions for convenience of presentation; thus the vertical axis is given in arbitrary units.

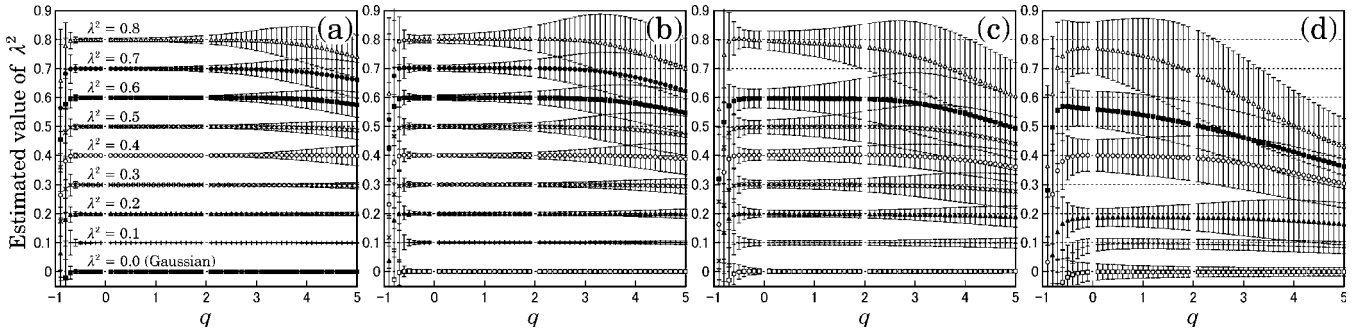

FIG. 3. The sample mean of the estimator $\hat{\lambda}_q^2$, where the observed time series $\{x_i\}$ is standardized before computing $\hat{\lambda}_q^2$. The sample mean was estimated from 200 samples. The theoretical values of λ^2 are shown in panel (a). The error bars indicate the sample standard deviation. (a) Data length $n=10^6$; (b) $n=10^5$; (c) $n=10^4$; and (d) $n=10^3$.

mator, the mean squared error (MSE) has generally been used. For the estimator $\hat{\lambda}_q^2$ of the parameter λ^2, the MSE of $\hat{\lambda}_q^2$ is defined as

$$\epsilon(\hat{\lambda}_q^2) = \langle (\hat{\lambda}_q^2 - \lambda^2)^2 \rangle, \tag{7}$$

where the angular brackets stand for the ensemble average.

The numerically estimated MSE are shown as the square root of $\epsilon(\hat{\lambda}_q^2)$ in Fig. 4(a). As shown in Fig. 4(a), the optimal order q to provide minimum MSE depends on the value of the parameter λ^2. Numerical results show that an optimal value of q can be found in the range $0<q<1$ when $0.1<\lambda^2<0.4$, and in the range $-0.3<q<0$ when $0.4<\lambda^2<0.8$. In practical applications, for the purpose of guaranteeing numerical stability (and in order to avoid division by zero in computer arithmetic), it is better to use a positive value of q. The actual value of λ^2 observed experimentally is also unknown *a priori*, leaving one to guess the optimal q to be used. Therefore, we propose to use $\hat{\lambda}_{q=0.5}^2$ as a nearly optimal estimator for $\lambda^2<0.8$. In addition, our numerical study suggests that the $\hat{\lambda}_{0.5}^2$ for $\lambda^2<0.8$ is a consistent estimator. As we can see in Fig. 4(b), the MSE of $\hat{\lambda}_{0.5}^2$ is approximately proportional to n^{-1}. Extrapolation of this dependence to $n\to\infty$ implies statistical consistency: $\mathrm{plim}_{n\to\infty}\hat{\lambda}_q^2=\lambda^2$.

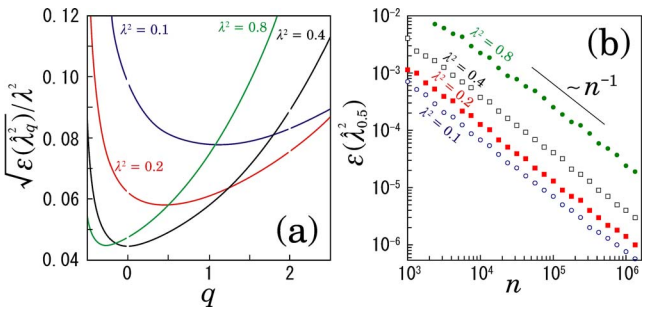

FIG. 4. (Color online) (a) Dependence of mean squared error $\epsilon(\hat{\lambda}_q^2)$ on the order q, where the data length $n=10^4$, and $q\neq 0,2$. (b) Asymptotic properties of $\epsilon(\hat{\lambda}_{0.5}^2)$, where n is the data length. The $\epsilon(\hat{\lambda}_q^2)$ are numerically obtained from 1024 samples. The values of the parameter λ^2 are shown in each panel.

Based on the above facts, we use the estimator $\hat{\lambda}_{0.5}^2$ in the following.

B. Log-normal cascade-type multiplicative process

For the next example, we study a log-normal cascade-type multiplicative process inspired by the intermittency problem of hydrodynamic turbulence [10]. Although there are several ways to simulate intermittency properties of a turbulent velocity field [3,11–14], we here remain within the framework of a multiplicative log-normal model [Eq. (2)]. The important property of our model is that its PDF is described exactly by Eq. (1), and shows the deformation predicted by the log-normal model of Kolmogorov [15] and Obukhov [16].

The numerical procedure to generate a time series of our model is as follows. First we generate a time series $\{\xi_i\}$ of Gaussian white noise with zero mean, as illustrated in Fig. 5(a). In this case, the total number of data points is 2^m ($i=1,\ldots,2^m$), where m is the total number of cascade steps. In the first cascade step ($j=1$), we divide the whole interval into two equal subintervals, and then multiply ξ_i in each subinterval by random weights $\exp[\omega^{(1)}(k)](k=0,1)$ [Fig. 5(b)]. In our model, $\omega^{(j)}(k)$ are independent Gaussian random variables with zero mean and constant variance, $\langle \omega^{(j)}(k)^2\rangle=\lambda_0^2$. In the next cascade step ($j=2$), we further divide each subinterval into two equal subintervals, and apply the random weights $\exp[\omega^{(2)}(k)]$ $(k=0,1,2,3)$ [Fig. 5(c)]. The same procedure is repeated, and after m cascade steps, the time series $\{x_i\}$ of the cascade process is given by

$$x_i = \xi_i \exp\left[\sum_{j=1}^{m} \omega^{(j)}\left(\left\lfloor \frac{i-1}{2^{m-j}} \right\rfloor\right)\right], \tag{8}$$

where $\lfloor\cdot\rfloor$ is the floor function. Note that in order to simplify the notation, the random variable x_i is not standardized.

The sum of the Gaussian variables $\omega^{(j)}(k)$ in Eq. (8) is also a Gaussian variable. Thus, the time series $\{x_i\}$ can be described by the same form of Eq. (2), with the PDF given by Eq. (1) with $\lambda^2=m\lambda_0^2$. As we can see in Figs. 5(b)–5(d), heterogeneity of local variance generated by the cascade steps results in a non-Gaussian PDF. The variance heteroge-

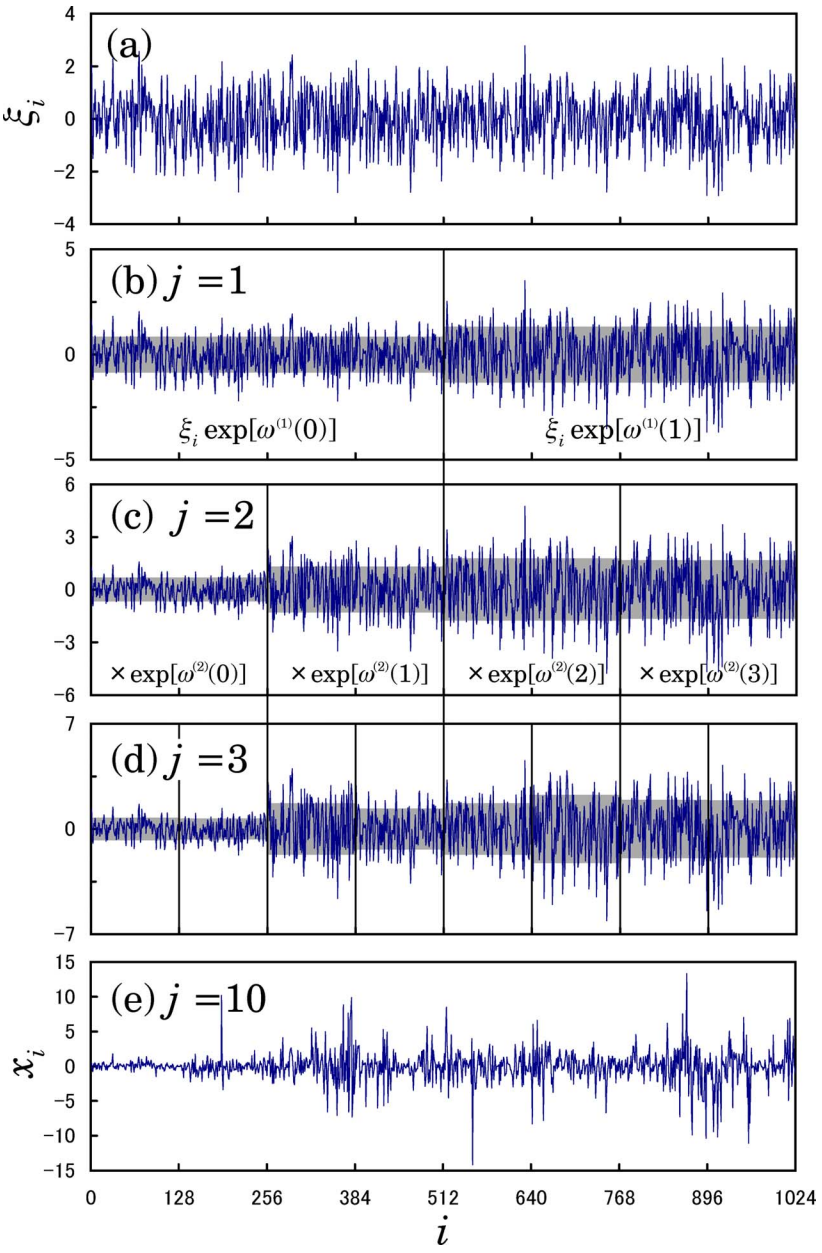

FIG. 5. (Color online) Illustration of data generation procedures of the multiplicative cascade process, where the total number of cascade steps is $m=10$ and $\langle \omega^{(j)}(k)^2 \rangle = 0.8^2/10$. (a) Time series $\{\xi_i\}$ of Gaussian white noise with zero mean and unit variance; (b) first cascade step; (c) second cascade step; (d) third cascade step; and (e) generated time series after 10 cascade steps ($\lambda^2 = 0.8^2$). The gray shading indicates the standard deviation in each subinterval. In this illustration, the variance of the time series is set to unity.

neity is an important difference from an independent and identically distributed (IID) random sequence, as shown in Fig. 1.

In the investigation of the intermittency of turbulence, the scale dependence of the PDF has been studied with great interest [1,4,5]. As pointed out by Castaing *et al.* [1], the original Kolmogorov-Obukhov theory [15,16] predicts that the non-Gaussian parameter λ^2 in the inertial range is proportional to $-\ln s$, where s is the spatial scale. This feature is simulated by our model. In our case, the logarithmic dependence of λ^2 is observed as the slow convergence to Gaussian behavior with an increasing coarse-grained level of the time series.

To demonstrate that our estimator $\hat{\lambda}_q^2$ [Eq. (5)] can also quantify the scale dependence of the non-Gaussian PDF with an increasing coarse-grained level of time series, we consider integrated time series $\{Z_n\}$ of a time series $\{x_i\}$, where

$$Z_n = x_1 + x_2 + \cdots + x_n.$$

We study the deformation process of the PDF of the difference $\Delta_s Z_i = Z_{i+s} - Z_i$, which mimics the velocity difference in turbulence statistics. If the time series $\{x_i\}$ is the IID sequence introduced in the previous subsection, the PDF of $\Delta_s Z_i$ (the sum of x_i) rapidly converges to a Gaussian as the scale s increases, as shown in Fig. 6(b). This example (or the Berry-Esséen theorem [17]) provides a benchmark for the speed of convergence to a Gaussian. Compared to the IID sequence, the time series of the cascade model shows extremely slow convergence to a Gaussian, as shown in Fig. 6(a). In addition, as we can see in Fig. 7(b), the scale dependence of the non-Gaussian parameter λ^2 clearly displays the difference between processes different in terms of their convergence to the Gaussian.

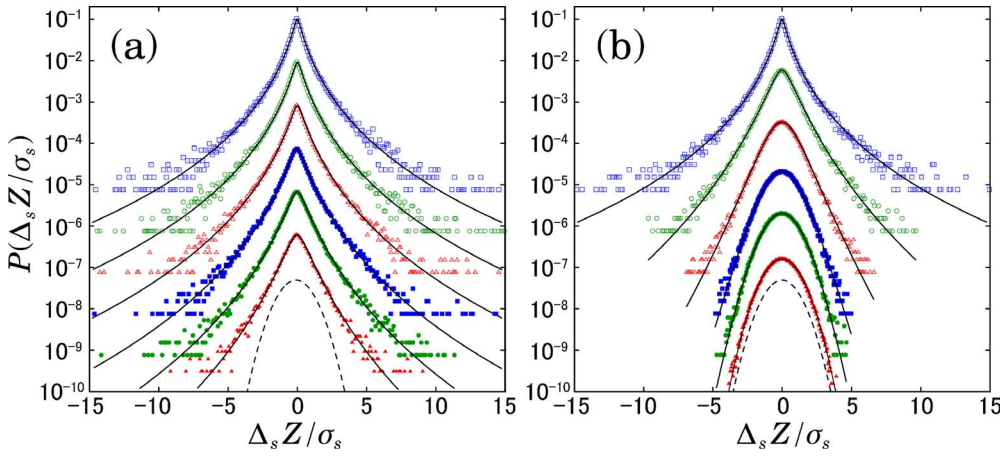

FIG. 6. (Color online) Deformation of PDF's across scales. Standardized PDF's (in logarithmic scale) of $\Delta_s Z$ for different scales are shown for (from top to bottom) $s=1,4,16,64,256,1024$. These PDF's are estimated from 256 samples. (a) Log-normal cascade process with $\lambda^2=0.8^2$ and $m=16$. (b) An IID sequence ($n=10^5$) of Eq. (6) with $\lambda^2=0.8^2$. The PDF's of $\Delta_1 Z$ in (a) and (b) have the same shape. In the solid line, we superimposed the PDF approximated by Eq. (3) with the estimated value of $\hat{\lambda}^2_{0.5}$. For comparison, the dashed line denotes a Gaussian distribution.

For the cascade model, if we approximate the local distribution of the sum of $\{x_i\}$ by a Gaussian, the PDF of $\Delta_s Z_i$ can be described by Eq. (3). Note that in an intermediate cascade step [Figs. 5(b)–5(d)], the distribution of x_i in each subinterval is a Gaussian. Therefore, a local distribution of $\{x_i\}$ is nearly Gaussian, while the convolution of the distributions becomes more Gaussian. Using the Gaussian approximation, the scale dependence of the non-Gaussian parameter λ^2 can be evaluated as

$$\lambda^2(s) \approx \lambda_0^2(m - \log_2 s) \sim -\ln s. \tag{9}$$

As we can see in Fig. 7, Eq. (9) provides good predictions. In other words, the estimator $\hat{\lambda}^2_{0.5}$ can properly characterize the intermittency properties.

As shown in Fig. 6(b), our numerical results demonstrate that for the IID sequence, the deformation of the PDF's of $\Delta_s Z_i$ is also well approximated by Eq. (3). Note that the PDF of $\Delta_s Z_i$ ($s>1$) of the IID sequence is not exactly the same as Eq. (3), although the PDF of x_i ($=\Delta_1 Z_{i-1}$) is given by Eq. (3) (see also the Appendix). As seen in Fig. 7 (filled triangles), the speed of convergence to a Gaussian is much faster than that of cascade processes.

IV. APPLICATION TO FINANCIAL TIME SERIES

Since Ghashghaie *et al.* [6] demonstrated that the PDF's of foreign exchange price changes at different time scales can be described in the same way as Castaing *et al.* [1] described the PDF's of velocity variations at different space separations in fully developed turbulence by Eq. (1), the

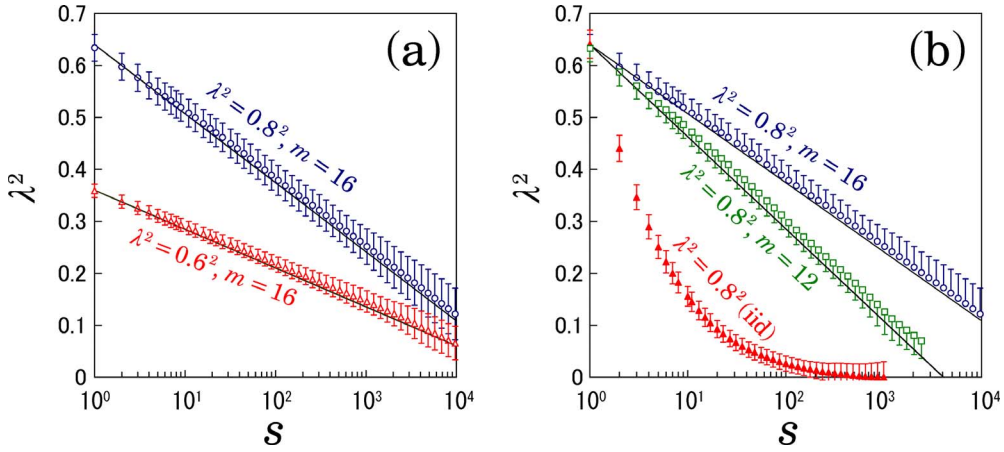

FIG. 7. (Color online) Scale dependence of the non-Gaussian parameter estimated by $\hat{\lambda}^2_{0.5}$. (a) Log-normal cascade process with $\lambda^2=0.8^2$ and $m=16$ (circles) and with $\lambda^2=0.6^2$ and $m=16$ (triangles). (b) Log-normal cascade process with $\lambda^2=0.8^2$ and $m=16$ (circles) and with $\lambda^2=0.8^2$ and $m=12$ (squares); IID random variables of Eq. (6) with $\lambda^2=0.8^2$ (filled triangles). The sample means of the non-Gaussian parameter estimated from 4096 samples are plotted. The error bars indicate the sample standard deviation. The solid lines indicate the theoretical prediction.

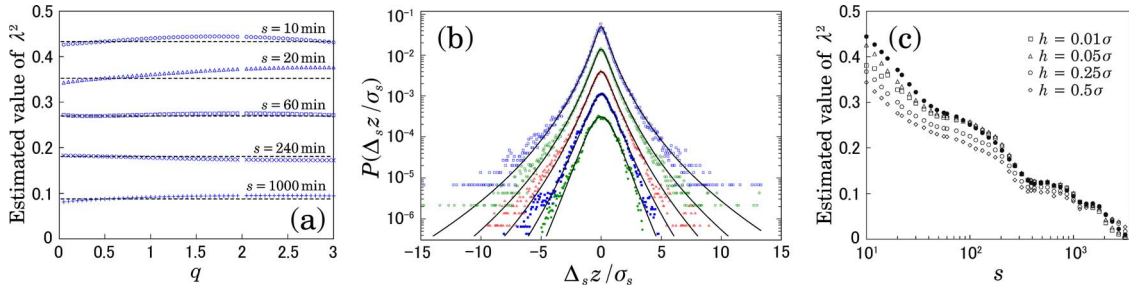

FIG. 8. (Color online) Estimation of the non-Gaussian parameter λ^2 for log-return fluctuations $\Delta_s z$ of the S&P500 index data from Jan. 2, 1992 to Dec. 30, 1994. (a) $\hat{\lambda}_q^2$ vs q for different time scales are shown for (from top to bottom) $s=10$, 20, 60, 240, and 1000 min. The dashed horizontal lines show the values of the $\hat{\lambda}_{0.5}^2$. (b) Standardized PDF's (in logarithmic scale) of $\Delta_s z$ for different time scales are shown for (from top to bottom) $s = 10$, 20, 60, 240, and 1000 min, where σ_s is the standard deviation of $\Delta_s z$. In solid lines, we superimpose the numerically integrated PDF's of Eq. (3) with $\hat{\lambda}_{0.5}^2$. The PDF's are shifted in vertical directions for convenience of presentation; thus the vertical axis is given in arbitrary units. (c) Scale dependence of the non-Gaussian parameter λ^2 estimated from $\hat{\lambda}_{0.5}^2$ (filled circles) and from a fitting algorithm based on the histogram estimator with a different bin width h. The values of the h are shown in the panel.

analogy between turbulence and finance has been studied by some physicists [18–21]. Recently, we also demonstrated that the PDF's of fluctuations of the U.S. S&P500 stock index are well approximated by Eq. (1) [7]. In addition, we revealed the empirical fact that the temporal dependence of the PDF shows a gradual, systematic increase in the non-Gaussian parameter λ^2 on approaching the Black Monday crash of October 1987. Because the marked fat tails characterized by the large value of λ^2 imply a high probability of a large price change, the quantitative characterization of the non-Gaussian behavior is potentially of great importance for risk analysis in the financial markets.

In this section, we demonstrate that our estimator $\hat{\lambda}_q^2$ [Eq. (5)] is applicable to the non-Gaussian behavior of the financial time series. As an example, we analyze the U.S. S&P500 historical data for the three-year period, from January 2, 1992 to December 30, 1994, with a sampling frequency of 2 min intervals. The total number of data points is 1.48×10^5. We define the S&P500 index fluctuations $\{\Delta_s z(t)\}$ as the log return, $\Delta_s z(t) = \ln y(t+s) - \ln y(t)$, where $y(t)$ denotes the S&P500 index at time t, and estimate the non-Gaussian parameter λ^2 of the log returns on different time scales s.

Figure 8(a) shows the $\hat{\lambda}_q^2$ vs q for different time scales. The nearly constant values of $\hat{\lambda}_q^2$ independent of q indicate that the standardized PDF of the log return can be well described by a multiplicative log-normal model [Eq. (3)] with a single non-Gaussian parameter λ^2. As shown in Fig. 8(b), we can see the good fit of the PDF of the multiplicative log-normal model [Eq. (3)] to the actual data.

The scale dependence of the $\hat{\lambda}_{0.5}^2$ is shown in Fig. 8(c) by filled circles. Compared to the numerical examples shown in Fig. 7(b), we can see the slow convergence to a Gaussian similar to the cascade models. This fact implies the existence of heterogeneity of local variance. In the field of financial time series analysis, quantification of the variance heterogeneity has been an important issue [22]. Our approach, inspired by turbulence statistics, may serve as a good characterization of the financial time series for this purpose.

To date, the non-Gaussian parameter λ^2 has been estimated from a fitting algorithm to minimize the χ^2 statistic between the experimental and numerically integrated PDF's. In this method, the construction of the experimental PDF has been based on a nonparametric probability density estimation, such as a conventional histogram estimator. It is important to note that the global consistency of the histogram estimator can be guaranteed only if, as the number of data points n goes to infinity, the histogram bin width h goes to zero, while ensuring $nh \to \infty$ [23]. Therefore, the point estimation of the λ^2 using the histogram estimator with a finite bin width is always biased, depending on the number of data points n, the bin width h, and the shape of the true density function. For instance, the scale dependence of λ^2 estimated from the histogram estimator with the different bin width h is shown in Fig. 8(c), which demonstrates that the estimated value of the λ^2 strongly depends on the bin width h, especially for the large value of λ^2. Hence the arbitrary choice of the bin width h makes the quantitative estimation of λ^2 unreliable. On the other hand, our estimator $\hat{\lambda}_q^2$ [Eq. (5)] is straightforward and without such arbitrariness.

V. SUMMARY AND DISCUSSION

In this paper, we derive the equations of a standardized PDF [Eq. (3)] and its qth order absolute moments [Eq. (4)] of multiplicative log-normal models [Eq. (2)]. Using the relation between a qth order absolute moment and the non-Gaussian parameter of a multiplicative log-normal model, an estimator [Eq. (5)] of the non-Gaussian parameter is introduced. By applying the estimator to a time series of multiplicative log-normal models with IID and cascade processes, it is demonstrated that the estimator can reliably determine the theoretical value of the non-Gaussian parameter. In addition, we demonstrate that non-Gaussian PDF's observed in the S&P500 index fluctuations are well described by the multiplicative log-normal model [Eq. (3)].

In the study of fully developed turbulence, one of the main tools to characterize the intermittency has been the multiscaling analysis of velocity structure functions. Moreover, the multiscaling technique has been applied to other

systems [24–26]. In our notations, the scaling of the structure functions $S_q(s)$ of time series $\{x_i\}$ is described as

$$S_q(s) = \langle |Z_{i+s} - Z_i|^q \rangle = \langle |\Delta_s Z_i|^q \rangle \sim s^{\zeta_q}, \tag{10}$$

where $Z_i = \Sigma_{j=1}^{i} x_i$. Within the framework of the multiplicative log-normal model, the relation between the $\langle |\Delta_s Z_i|^q \rangle$ and the non-Gaussian parameter of $\Delta_s Z_i$ is given by Eq. (4). Thus, scaling exponents ζ_q $(q \neq 2)$ are fully determined by the scale dependence of the non-Gaussian parameter λ^2 and a scaling exponent ζ_2. For instance, using the scale dependence of λ^2 [Eq. (9)] of the cascade model [Eq. (8)] we can obtain the following ζ_q spectrum:

$$\zeta_q = \frac{q}{2} - \frac{q(q-2)}{2}\lambda_0^2. \tag{11}$$

In this case, logarithmic decay of the scale dependence of λ^2, $\lambda^2(s) \sim -\ln s$, is essential for the existence of the scaling range because of $\exp(\lambda^2)$ in Eq. (4). Therefore, for other types of scale dependence in multiplicative log-normal models, the multiscaling analysis based on Eq. (10) cannot provide valid characterizations.

Although in our examples of the multiplicative log-normal model we assume that $\{\xi_i\}$ is a sequence of Gaussian white noise, generalization to fractional Gaussian noise is possible. Further study on generalizations of multiplicative log-normal models will be reported elsewhere.

ACKNOWLEDGMENTS

The author (K.K.) would like to thank Professor Naoaki Bekki for stimulating discussions. This work was partly supported by a Research Grant of the College of Engineering, Nihon University for 2007.

APPENDIX: SUM OF TWO IID VARIABLES IN A MULTIPLICATIVE LOG-NORMAL MODEL

In this Appendix, we study the PDF of the sum of two IID variables in a multiplicative log-normal model. Let $\{x_1, x_2, x_3, \ldots\}$ be a sequence of IID variables described by

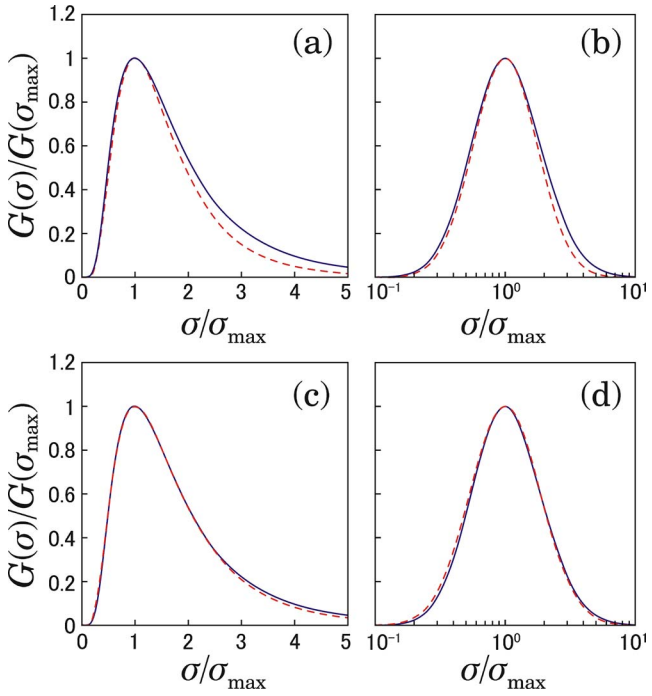

FIG. 9. (Color online) [(a) and (b)] Comparison between a log-normal kernel with $\lambda^2 = 0.8^2/2$ (dashed line) and the convoluted kernel $G(\sigma)$ [Eq. (A4)] with $\lambda^2 = 0.8^2$ (solid line), where $\sigma_{\max}$ gives the maximum value of each function. [(c) and (d)] Comparison between Eq. (A4) with $\lambda^2 = 0.8^2$ (solid line) and the approximation using a log-normal kernel (dashed line), where both distributions have the same logarithmic variance.

$$x_i = \xi_i e^{\omega_i}, \tag{A1}$$

where $\{\xi_i\}$ and $\{\omega_i\}$ are both sequences of Gaussian white noise with $\langle \xi \rangle = \langle \omega \rangle = 0$, $\langle \xi^2 \rangle = 1$, and $\langle \omega^2 \rangle = \lambda^2$. Under these assumptions, the PDF of x_i is given by

$$P_\lambda(x) = \int_0^\infty \frac{1}{\sqrt{2\pi}\lambda} \exp\left(-\frac{\ln^2 \sigma}{2\lambda^2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) d(\ln \sigma). \tag{A2}$$

To simplify the expression, we assume this PDF with the non-Gaussian parameter λ instead of the standardized PDF [Eq. (3)]. In this case, the PDF of the sum $x_{i+1} + x_{i+2} = \Delta_2 Z_i$ is given by the convolution of $P_\lambda(x)$ as follows:

$$P(\Delta_2 Z) = \int_{-\infty}^{\infty} P_\lambda(\Delta_2 Z - x) P_\lambda(x) dx = \int_0^\infty \frac{1}{\sqrt{2\pi}(\lambda/\sqrt{2})} \exp\left(-\frac{(\ln \sigma)^2}{2(\lambda/\sqrt{2})^2}\right) R_\lambda(\ln \sigma) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\Delta_2 Z)^2}{2\sigma^2}\right) d(\ln \sigma), \tag{A3}$$

where

$$R_\lambda(\ln \sigma) = \int_0^{\pi/2} \frac{1}{2\sqrt{\pi}\lambda} \exp\left(-\frac{2 \ln \sigma \{\ln(\sin\theta \cos\theta)\} + \ln^2(\sin\theta) + \ln^2(\cos\theta)}{2\lambda^2}\right) \frac{d\theta}{\sin\theta \cos\theta}. \tag{A4}$$

If the $R_\lambda(\ln \sigma)$ is equal to one, Eq. (A3) can be described by the same form as Eq. (A2) with the non-Gaussian parameter $\lambda/\sqrt{2}$. In this case, the scale dependence of λ^2 would be proportional to s^{-1}. However, the fact is that this assumption provides a rough approximation.

To illustrate the effect of $R_\lambda(\ln \sigma)$, we numerically evaluate the kernel in Eq. (A3) as follows:

$$G(\sigma) = \frac{1}{\sqrt{2\pi}(\lambda/\sqrt{2})} \exp\left(-\frac{(\ln \sigma)^2}{2(\lambda/\sqrt{2})^2}\right) R_\lambda(\ln \sigma). \quad \text{(A5)}$$

As shown in Figs. 9(a) and 9(b), the log-normal kernel is deformed to a slightly wider distribution by $R_\lambda(\ln \sigma)$. Nevertheless, the multiplication of the log-normal kernel and $R_\lambda(\ln \sigma)$ is well approximated by a log-normal kernel with a larger parameter λ [Figs. 9(c) and 9(d)]. In addition, the difference between the log-normal kernels with and without $R_\lambda(\ln \sigma)$ gradually decreases through the convergence to a Gaussian.

In conclusion, the above analysis demonstrates that the PDF of the sum of IID variables in a multiplicative log-normal model is well approximated (but not exactly described) by Eq. (1).

[1] B. Castaing, Y. Gagne, and E. J. Hopfinger, Physica D **46**, 177 (1990).
[2] E. Bacry, J. Delour, and J. F. Muzy, Phys. Rev. E **64**, 026103 (2001).
[3] A. Arneodo, E. Bacry, and J. F. Muzy, J. Math. Phys. **39**, 4142 (1998).
[4] B. Chabaud, A. Naert, J. Peinke, F. Chillà, B. Castaing, and B. Hébral, Phys. Rev. Lett. **73**, 3227 (1994).
[5] L. Sorriso-Valvo, V. Carbone, P. Veltri, G. Consolini, and R. Bruno, Geophys. Res. Lett. **26**, 1801 (1999).
[6] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature (London) **381**, 767 (1996).
[7] K. Kiyono, Z. R. Struzik, and Y. Yamamoto, Phys. Rev. Lett. **96**, 068701 (2006).
[8] K. Kiyono, Z. R. Struzik, N. Aoyagi, S. Sakata, J. Hayano, and Y. Yamamoto, Phys. Rev. Lett. **93**, 178103 (2004).
[9] K. Kiyono, Z. R. Struzik, N. Aoyagi, F. Togo, and Y. Yamamoto, Phys. Rev. Lett. **95**, 058101 (2005).
[10] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics* (MIT Press, Cambridge, MA, 1975).
[11] J. Eggers and S. Grossmann, Phys. Rev. A **45**, 2360 (1992).
[12] J. A. C. Humphrey, C. A. Schuler, and B. Rubinsky, Fluid Dyn. Res. **9**, 81 (1992).
[13] R. Benzi, L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, and A. Vulpiani, Physica D **65**, 352 (1993).
[14] A. Marshak, A. Davis, R. Cahalan, and W. Wiscombe, Phys. Rev. E **49**, 55 (1994).
[15] A. N. Kolmogorov, J. Fluid Mech. **13**, 82 (1962).
[16] A. M. Obukhov, J. Fluid Mech. **13**, 77 (1962).
[17] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. (Wiley, New York, 1971), Vol. 2.
[18] A. Arnéodo, J.-F. Muzy, and D. Sornette, Eur. Phys. J. B **2**, 277 (1998).
[19] J. F. Muzy, J. Delour, and E. Bacry, Eur. Phys. J. B **17**, 537 (2000).
[20] M. Ausloos and K. Ivanova, Phys. Rev. E **68**, 046122 (2003).
[21] N. Kozuki and N. Fuchikami, Physica A **329**, 222 (2003).
[22] J. Johnston and J. DiNardo, *Econometric Methods*, 4th ed. (McGraw Hill, New York, 1998).
[23] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization* (Wiley, New York, 1992).
[24] K. Ivanova and M. Ausloos, Eur. Phys. J. B **8**, 665 (1999).
[25] H. Nakao, Phys. Lett. A **266**, 282 (2000).
[26] D. C. Lin and R. L. Hughson, Phys. Rev. Lett. **86**, 1650 (2001).