PHYSICAL REVIEW E **71**, 025104(R) (2005)

Power-law correlated processes with asymmetric distributions

Boris Podobnik,[1,2,3] Plamen Ch. Ivanov,[2] Vojko Jazbinsek,[4] Zvonko Trontelj,[4] H. Eugene Stanley,[2] and Ivo Grosse[3]
[1]*Faculty of Civil Engineering, University of Rijeka, Rijeka, Croatia*
[2]*Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA*
[3]*Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany*
[4]*Institute of Mathematics, Physics and Mechanics, University of Ljubljana, Ljubljana, Slovenia*
(Received 8 November 2004; published 28 February 2005)

Motivated by the fact that many physical systems display (i) power-law correlations together with (ii) an asymmetry in the probability distribution, we propose a stochastic process that can model both properties. The process depends on only two parameters, where one controls the scaling exponent of the power-law correlations, and the other controls the degree of asymmetry in the distributions leaving the correlations unaffected. We apply the process to air humidity data and find that the statistical properties of the process are in a good agreement with those observed in the data.

DOI: 10.1103/PhysRevE.71.025104 PACS number(s): 02.50.−r, 05.40.−a, 87.19.−j

Many physical phenomena exhibit temporal or spatial correlations that can be approximated by power laws. For example, long-range power-law correlations have been found in physical, biological, and social systems [1–8], and various stochastic processes [9–12] have been developed to model these power-law scaling properties. Recent studies have shown that in addition to power-law correlations empirical data often exhibit a significant skewness or asymmetry in their distributions. Asymmetric distributions have been found in astrophysical data [13], genome sequences [14], respiratory dynamics [15], brain dynamics [16], heartbeat dynamics [17], turbulence [18], physical activities, and finance [19].

With the goal of constructing a stochastic process that can generate time series with both power-law correlated and asymmetrically distributed variables x_i, we define the process $\mathcal{A}(\rho,\lambda)$ by

$$x_i = \lambda|x_{i-1}| + \sum_{n=1}^{\infty} a_n(\rho)(x_{i-n} - \lambda|x_{i-n-1}|) + \eta_i, \qquad (1)$$

where $\rho \in (0,0.5)$ and $\lambda \in (-1,1)$ are free parameters, $a_n(\rho)$ are weights defined by $a_n(\rho)=\Gamma(n-\rho)/[\Gamma(-\rho)\Gamma(1+n)]$, Γ denotes the Gamma function, and η_i denotes independent and identically distributed Gaussian variables with expectation value $\langle\eta_i\rangle=0$ and variance $\langle\eta_i^2\rangle=1$. The parameter ρ controls the length of the memory, i.e., how rapidly the influence of past values x_{i-n} and $|x_{i-n-1}|$ on x_i decays in time, and the parameter λ controls the relative influence of x_{i-n} on x_i compared to the influence of $|x_{i-n-1}|$ on x_i.

$\mathcal{A}(\rho,\lambda)$ can be understood as a generalization [20,21] of the fractionally integrated process proposed in Refs. [9,10], to which $\mathcal{A}(\rho,\lambda)$ reduces for $\lambda=0$. While the fractionally integrated process $\mathcal{A}(\rho,0)$ is known to generate power-law correlated and symmetrically distributed time series [9,10], we will show in the following that, for $\lambda\neq 0$, $\mathcal{A}(\rho,\lambda)$ generates power-law correlations with an asymmetric distribution. Specifically, we will show that the parameter ρ controls the scaling exponent of the power-law correlations, and that the parameter λ controls the degree of asymmetry in the distributions, leaving the correlations almost unaffected.

Before studying the autocorrelation function of x_i, $C(n) \equiv \langle x_{i+n}x_i\rangle - \langle x_i\rangle^2$, and the probability distribution, $P(x)$, for different values of ρ and λ, we note that process $\mathcal{A}(\rho,\lambda)$ exhibits two invariance properties. Under the transformations $x_i \to -x_i$, $\eta_i \to -\eta_i$, $\lambda \to -\lambda$, one can see that $C(n|\rho,\lambda) = C(n|\rho,-\lambda)$ and $P(x|\rho,\lambda)=P(-x|\rho,-\lambda)$. That is, the autocorrelation functions calculated for opposite values of λ are identical, and the distributions for opposite values of λ are mirror images of each other. Hence, we focus on values of $\lambda \geqslant 0$ in the following study.

To quantify the autocorrelations in x_i generated by $\mathcal{A}(\rho,\lambda)$, we employ the method of detrended fluctuation analysis (DFA) [22]. In the DFA method one measures the standard deviation $F(n)$ of the detrended fluctuations as a function of the scale n. If $C(n)$ can be approximated by a power law with exponent γ, i.e., if $C(n) \propto n^{-\gamma}$, then $F(n)$ can also be approximated by a power law with exponent α, i.e., $F(n) \propto n^{\alpha}$, with $\alpha \approx 1+\gamma/2$ [22]. Hence, the value of α represents the degree of autocorrelations in the time series: if $\alpha > 0.5$, the time series is power-law correlated; if $\alpha=0.5$, the time series is uncorrelated or short-range correlated; and if $\alpha < 0.5$, the time series is power-law anticorrelated.

In order to study the influence of the parameter λ on autocorrelations and the degree of asymmetry in the distribution, we perform numerical simulations [23] of $\mathcal{A}(\rho,\lambda)$ with $\rho=0.3$ and varying λ. Figure 1(a) shows that, for $\lambda=0$, $F(n)$ can be approximated by a power law with scaling exponent α, i.e., $F(n) \propto n^{\alpha}$, where $\alpha \approx 0.5+\rho=0.8$, as expected from Refs. [9,10,24]. Figure 1(a) shows that, also for $\lambda \neq 0$, $F(n)$ can be approximated by a power law with scaling exponent α, where $\alpha \approx 0.5+\rho=0.8$, i.e., the value of λ has no visible effect on autocorrelations of x_i for asymptotically large values of n. We also find from Fig. 1(a) that, for $\lambda \neq 0$, the $F(n)$ curves exhibit a crossover at small time scales n, which becomes more pronounced and shifts to larger scales of n with increasing $|\lambda|$.

In Fig. 1(b) we see that, for $\lambda=0$, $P(x)$ is symmetric, as expected for the process of Refs. [9,10]. For $\lambda=0.6$ and $\lambda=0.9$, we find that $P(x)$ is asymmetric with positive skew-

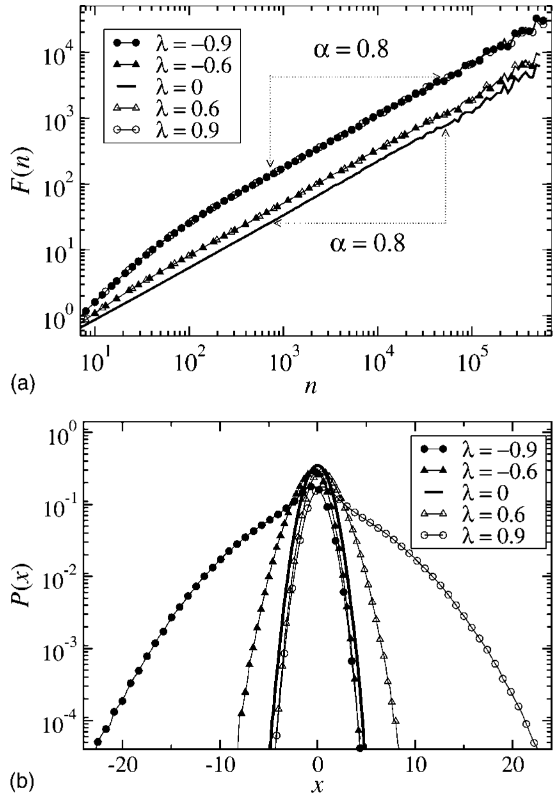

FIG. 1. Correlations and probability distributions obtained from numerical simulations of process $\mathcal{A}(\rho,\lambda)$ with ρ=0.3 and λ=0, ±0.6, and ±0.9. (a) Detrended fluctuation function $F(n)$. We see that the $F(n)$ curves for opposite values of λ are identical, and we find that, for all values of $\lambda, F(n)$ can be approximated by a power law for asymptotically large n, and the scaling exponent α in $F(n) \propto n^{\alpha}$ is virtually the same for all values of λ. (b) Probability distribution $P(x)$. We see that for opposite values of λ the distributions are mirror images of each other, and we find that $P(x)$ is symmetric for λ=0, $P(x)$ is asymmetric with positive skewness for $\lambda > 0$, and the degree of asymmetry increases with increasing $|\lambda|$.

ness, where the left tail is almost identical to the left tail of the symmetric distribution, and the right tail is broader than the right tail of the symmetric distribution. Due to the invariance $P(x|\rho,\lambda)=P(-x|\rho,-\lambda)$, we find that the distributions for positive and negative values of λ are mirror images of each other for opposite values of λ.

In order to investigate how the correlation properties of $\mathcal{A}(\rho,\lambda)$ depend on ρ, we perform numerical simulations to obtain time series for λ=0.6 and ρ ranging from 0 to 0.4. We find from Fig. 2 that the $F(n)$ curves can be approximated by power laws with a scaling exponent of $\alpha \approx 0.5+\rho$. This states that we obtain the same scaling law for the process $\mathcal{A}(\rho,\lambda)$ generating asymmetrical distributions as for the process $\mathcal{A}(\rho,0)$ generating symmetrical distributions. Numerically we find that, independently of λ, the relation $\alpha \approx 0.5+\rho$ holds for all values of ρ and λ where $\rho \in (0,0.5)$ and $\lambda \in (-1,1)$ [24].

In order to model probability distributions with a different shape, particularly with tails broader than those generated by process $\mathcal{A}(\rho,\lambda)$, we propose the process $\mathcal{B}(\rho,\lambda)$ by substituting the term η_i in Eq. (1) by the term $\sigma_i \eta_i$, where the

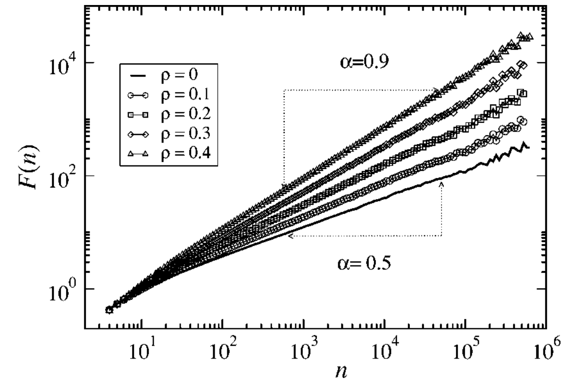

FIG. 2. Detrended fluctuation function $F(n)$ obtained from numerical simulations of process $\mathcal{A}(\rho,\lambda)$ with λ=0.6 and ρ=0, 0.1, 0.2, 0.3, and 0.4. For asymptotically large values of n, each of the $F(n)$ curves can be approximated by a power law $F(n) \propto n^{\alpha}$ with the scaling exponent $\alpha \approx 0.5+\rho$.

time-dependent standard deviation σ_i is defined by [25]

$$\sigma_i = \sum_{n=1}^{\infty} a_n(\rho) \frac{|x_{i-n}|}{\langle |x_{i-n}| \rangle}. \qquad (2)$$

$\mathcal{B}(\rho,\lambda)$ generates not only long-range autocorrelations in x_i, but also autocorrelations in the magnitudes $|x_i|$, and processes with autocorrelations in the magnitudes have been introduced to model broader tails in the distributions [26].

Figure 3(a) shows that, for asymptotically large n, each $F(n)$ curve can be approximated by a power law with scaling exponent $\alpha \approx 0.5+\rho$. This states that the time-dependent standard deviation σ_i does not affect the relation between α and ρ observed for process $\mathcal{A}(\rho,\lambda)$ [24]. Figure 3(b) shows the distribution of x_i generated by $\mathcal{B}(\rho,\lambda)$ for ρ=0.3 and λ ranging from 0 to 0.3. As expected, the asymmetry vanishes for λ=0 even in the presence of the term $\sigma_i \eta_i$, meaning that this term alone does not create an asymmetry in the distribution of x_i but only broadens its tails [26]. For $\lambda > 0$, we find that, as λ increases, the right tails of the distributions become broader, the left tails of the distributions become thinner, and thus the asymmetry becomes more pronounced. Comparing Figs. 3(b) and 1(b) we find that the time-dependent standard deviation σ_i broadens the tails and increases the skewness of $P(x)$.

To exemplify the utility of process $\mathcal{B}(\rho,\lambda)$ for modeling real-world data, we study air humidity data, which can be considered an output of a complex geophysical system. We analyze the relative air humidity recorded in 10-min intervals at the Institute of Plant Genetics and Crop Plant Research in Gatersleben [27]. We denote the differences of successive relative air humidity by $\tilde{x}_i$, and we show in Fig. 4(a) the time series $\tilde{x}_i$. Figures 4(b) and 4(c) show that the time series $\tilde{x}_i$ exhibits both power-law autocorrelations with a scaling exponent of $\alpha \approx 0.87$ and an asymmetric distribution.

In order to investigate to which degree process $\mathcal{B}(\rho,\lambda)$ can approximate the statistical properties of the empirical time series $\tilde{x}_i$, we generate time series by numerical simulations of process $\mathcal{B}(\rho,\lambda)$ with ρ=0.37 and λ=0.15, where we

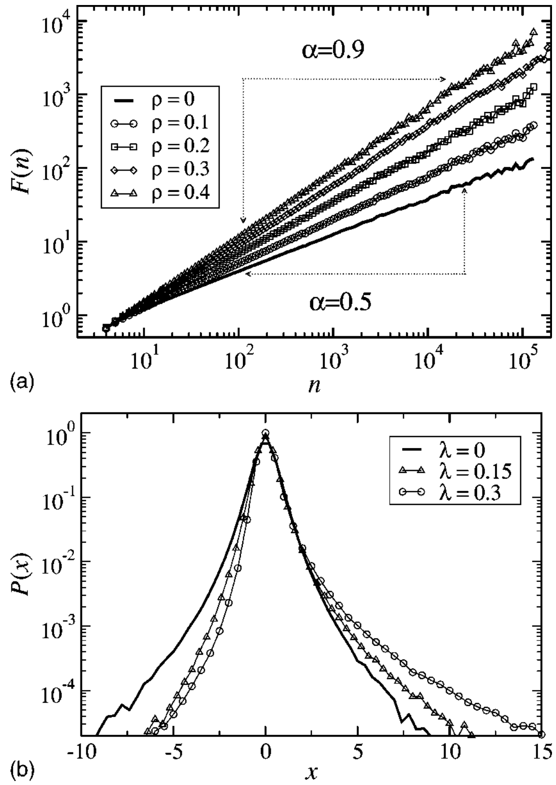

FIG. 3. Correlations and probability distributions of process $\mathcal{B}(\rho,\lambda)$. (a) Detrended fluctuation function $F(n)$ obtained from numerical simulations of process $\mathcal{B}(\rho,\lambda)$ with $\lambda=0.15$ and varying values of $\rho=0$, 0.1, 0.2, 0.3, and 0.4. We find that, independently of λ, $F(n)$ can be approximated by a power law $F(n)\propto n^{\alpha}$ with the scaling exponent $\alpha\approx 0.5+\rho$. (b) Probability distributions $P(x)$ obtained from numerical simulations with $\rho=0.3$ and $\lambda=0$, 0.15, and 0.3. We see that $P(x)$ is symmetric for $\lambda=0$, $P(x)$ is asymmetric with positive skewness for $\lambda>0$, and the skewness increases with increasing λ.

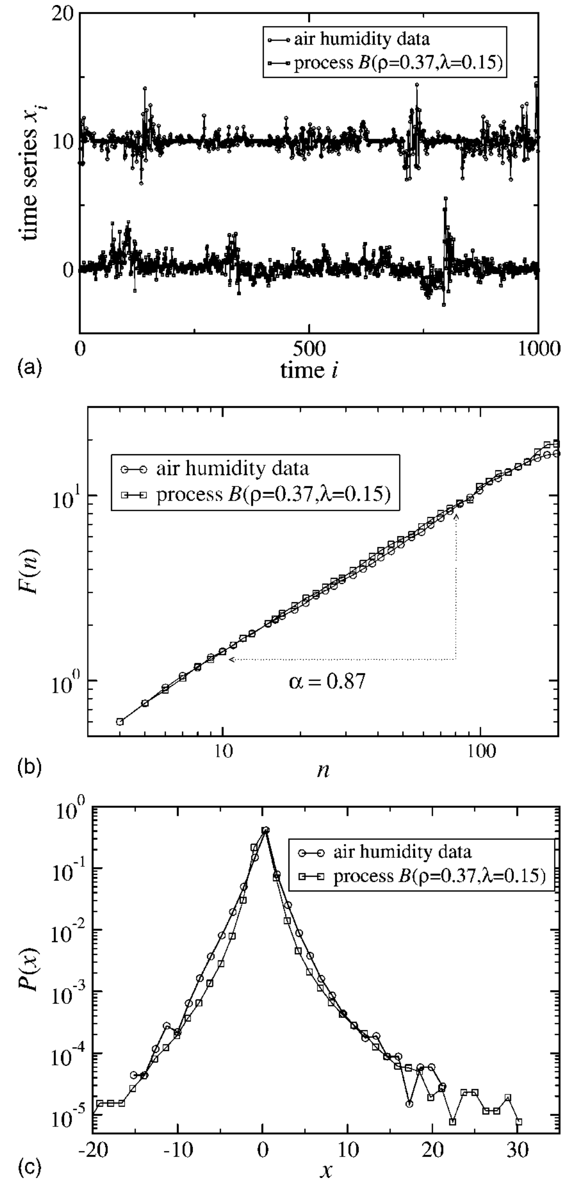

FIG. 4. Comparison of the changes of relative air humidity $\tilde{x}_i$ with the time series x_i generated by process $\mathcal{B}(\rho,\lambda)$ with $\rho=0.37$ and $\lambda=0.15$. (a) Time series $\tilde{x}_i$ and x_i. We find that both time series show sudden bursts of large fluctuations predominantly in the positive direction. (b) Detrended fluctuation functions $F(n)$. We find that autocorrelations of $\tilde{x}_i$ and x_i are very similar, and consistent with a power-law scaling of $F(n)\propto n^{\alpha}$ with the scaling exponent $\alpha\approx 0.87$. (c) Probability distributions $P(\tilde{x})$ and $P(x)$. We find that both distributions are asymmetric with positive skewness. Moreover, we find that even the shapes of both distributions are similar.

set ρ by using the relation $\alpha\approx 0.5+\rho$, and where we find λ based on a numerical least-square minimization. In Fig. 4, we present (a) the time series $\tilde{x}_i$ and x_i, (b) their detrended fluctuation functions $F(n)$, and (c) their distributions $P(\tilde{x})$ and $P(x)$. Figures 4(a)–4(c) show that the time series of $\tilde{x}_i$ and x_i look similar, that the autocorrelation behavior of the simulated time series x_i is in good agreement with that of the air humidity time series $\tilde{x}_i$, and that the distributions $P(\tilde{x})$ and $P(x)$ are asymmetric with positive skewness. Moreover, we find that even the shapes of both distributions are similar, which is surprising because the shape of $P(x)$ is not fitted to the shape of $P(\tilde{x})$, but the shape of $P(x)$ is entirely given by the values of ρ and λ.

One possible explanation for the positive skewness in the data is that it is very easy to increase the humidity rapidly, by rain for example, but it is hard to dry it rapidly. This simple physical fact could be one of the origins of the asymmetry observed in the distribution of $\tilde{x}_i$. The agreement of the statistical properties of $\tilde{x}_i$ and x_i observed in Fig. 4 might indicate that humidity changes at time i depend not only on past humidity changes x_{i-n} but also on their magnitudes $|x_{i-n-1}|$. The degree of asymmetry in the distribution of x_i reproduced by process $\mathcal{B}(\rho,\lambda)$ with a small value of $\lambda\approx 0.15$ suggests that the influence of the past magnitudes $|x_{i-n-1}|$ on x_i is significantly smaller than the influence of the past humidity changes x_{i-n}. Specifically, we might speculate that the influence of the past humidity changes x_{i-n} on x_i is approximately seven times greater than the influence of their magnitudes $|x_{i-n-1}|$ on x_i. Even though both processes $\mathcal{A}(\rho,\lambda)$ and $\mathcal{B}(\rho,\lambda)$ can generate asymmetric distributions, we find that the empirical distribution cannot be reproduced by process

$\mathcal{A}(\rho,\lambda)$, but it can be almost perfectly reproduced by process $\mathcal{B}(\rho,\lambda)$. This surprising observation indicates that the environmental factors η_i at time i might be amplified by a multiplicative factor σ_i, which does not depend on past humidity changes x_{i-n} but on their magnitudes $|x_{i-n-1}|$.

In conclusion, we propose two stochastic processes, $\mathcal{A}(\rho,\lambda)$ and $\mathcal{B}(\rho,\lambda)$, that generate simultaneously power-law autocorrelations and asymmetric probability distributions. Both processes depend on only two parameters, ρ and λ, where ρ controls the scaling exponent of the power-law autocorrelations and λ controls the degree of asymmetry. We study air humidity time series, and we find that they display both power-law autocorrelations and asymmetric distributions. We find that process $\mathcal{B}(\rho,\lambda)$ is capable of reproducing—qualitatively and quantitatively—the autocorrelations and the distribution of the data. The quantitative agreement of the shape of the distribution generated by process $\mathcal{B}(\rho,\lambda)$ with the shape of the distribution of the air humidity changes is surprising, because the shape of the distribution is not fitted, but fully determined by the parameters ρ and λ controlling the scaling exponent of the power-law autocorrelations and the skewness of the distribution, respectively. The surprising agreement of the shapes of the distributions might suggest that air humidity changes at time i are possibly driven by (i) past air humidity changes at times $i-n$, (ii) their magnitudes at times $i-n$, and (iii) environmental factors at time i amplified by a multiplicative factor that itself depends on past magnitudes at times $i-n$. It is clear that processes $\mathcal{A}(\rho,\lambda)$ and $\mathcal{B}(\rho,\lambda)$ lack many important details necessary for realistic weather models, but the simplicity and generality of processes $\mathcal{A}(\rho,\lambda)$ and $\mathcal{B}(\rho,\lambda)$ might possibly make them useful for modeling diverse physical systems exhibiting both power-law correlations and asymmetric distributions.

We thank Stephan Posch, Roland Schnee, and Stephan Weise for valuable discussions and the Ministry of Science of Croatia (Grant No. 0119623), the NIH (Grant No. HL071972), and the German Federal Ministry of Education and Research (Grant No. 0312706A) for financial support.

[1] H. E. Hurst, Proc.-Inst. Civ. Eng. **1**, 519 (1951).
[2] M. Casandro and G. Jona-Lasinio, Adv. Phys. **27**, 913 (1978).
[3] *Correlations and Connectivity: Geometric Aspects of Physics, Chemisty and Biology*, edited by H. E. Stanley and N. Ostrowsky (Kluwer, Dordrecht, 1990).
[4] T. Vicsek, *Fractal Growth Phenomenon*, 2nd ed. (World Scientific, Singapore, 1993).
[5] *Fractals in Science*, edited by A. Bunde and S. Havlin (Springer, Berlin, 1994).
[6] J. B. Bassingthwaighte, L. S. Liebovitch, and B. J. West, *Fractal Physiology* (Oxford University Press, New York, 1994).
[7] J. Beran, *Statistics for Long-Memory Processes* (Chapman & Hall, New York, 1994).
[8] H. Takayasu, *Fractals in the Physical Sciences* (Manchester University Press, Manchester, 1997).
[9] C. W. J. Granger and R. Joyeux, J. Time Ser. Anal. **1**, 15 (1980).
[10] J. Hosking, Biometrika **68**, 165 (1981).
[11] C.-K. Peng *et al.*, Phys. Rev. A **44**, R2239 (1991).
[12] H. A. Makse *et al.*, Phys. Rev. E **53**, 5445 (1995).
[13] L. Hui and E. Gaztanaga, Astrophys. J. **519**, 622 (1999).
[14] International Human Genome Sequence Consortium, Nature (London) **409**, 860 (2001); J. C. Venter *et al.*, Science **291**, 1304 (2001).
[15] M. Samon and F. Curley, J. Appl. Physiol. **83**, 975 (1997).
[16] C. L. Ehlers *et al.*, J. Neurosci. **18**, 7474 (1998).
[17] C. Braun *et al.*, Am. J. Physiol. Heart Circ. Physiol. **275**, H1577 (1998).
[18] G. Boffetta, A. Celani, and M. Vergassola, Phys. Rev. E **61**, R29 (2000).
[19] K. Ohashi, L. A. N. Amaral, B. H. Natelson, and Y. Yamamoto, Phys. Rev. E **68**, 065204(R) (2003).
[20] J. D. Hamilton, *Time Series Analysis* (Princeton, New Jersey, 1994).
[21] To eliminate trends in empirical data, one commonly takes first-order differences $x_i - x_{i-1}$, or higher-order integer differences. This differencing procedure is accomplished through the linear operator $1-L$, defined by $(1-L)x_i = x_i - x_{i-1}$ [20], where L is the linear backward-shift operator defined by $L^n x_i = x_{i-n}$. Fractional processes [9,10] are obtained by allowing the order ρ in $(1-L)^\rho$ to take fractional values. After expanding the fractionally differencing operator $(1-L)^\rho$ as an infinite binomial series in powers of L, one can see that Eq. (1), including the definition of $a_n(\rho)$, can be expressed by $(1-L)^\rho(x_i - \lambda|x_{i-1}|) = \eta_i$.
[22] C.-K. Peng *et al.*, Phys. Rev. E **49**, 1685 (1994).
[23] We introduce a cutoff length $\ell = 30\,000$ in our numerical simulations, and we let the sum in Eq. (1) run from 1 to ℓ, i.e., we set $a_n = 0$ for $n \geqslant \ell$.
[24] Based on the relation $\alpha \approx 1 - \gamma/2$ [22], and based on the fact that the process $\mathcal{A}(\rho,0)$ is known to generate power-law correlated variables with exponent $\gamma = 1 - 2\rho$, we obtain the relation $\alpha \approx 0.5 + \rho$.
[25] C. W. J. Granger and Z. Ding, J. Econometr. **73**, 61 (1996).
[26] R. F. Engle, Econometrica **50**, 987 (1982).
[27] The data are publicly available at http://portal.bic-gh.de