Q-Gaussian diffusion in stock markets

Fernando Alonso-Marroquin,[1] Karina Arias-Calluari,[1,*] Michael Harré,[1] Morteza N. Najafi,[2] and Hans J. Herrmann[3]
[1]*School of Civil Engineering, The University of Sydney, NSW 2006, Australia*
[2]*Department of Physics, University of Mohaghegh Ardabili, Ardabil, Iran*
[3]*PMMH, ESPCI, 7 Quai St. Bernard, 75005 Paris, France*

(Received 11 February 2019; published 28 June 2019)

We analyze the Standard & Poor's 500 stock market index from the past 22 years. The probability density function of price returns exhibits two well-distinguished regimes with self-similar structure: the first one displays strong superdiffusion together with short-time correlations and the second one corresponds to weak superdiffusion with weak time correlations. Both regimes are well described by q-Gaussian distributions. The porous media equation—a special case of the Tsallis-Bukman equation—is used to derive the governing equation for these regimes and the Black-Scholes diffusion coefficient is explicitly obtained from the governing equation.

Price fluctuations in stock markets exhibit remarkable features such as strong short-time correlations, weak long-time correlations, power law tails, and slow convergence to the normal distribution [1]. In the earliest stock market model, Bachelier proposed classical Brownian motion to represent price fluctuations whose description has been a landmark of research [2,3]. This model was the cornerstone for the well-established Black-Scholes equations for stock markets. Mandelbrot suggested that classical diffusion was not appropriate for modeling real stock markets [4]. His conclusion was based on the analysis of price variations of the cotton index whose probability density function (PDF) was better described by a Lévy distribution. Later, Mantegna and Stanley proposed that this Lévy distribution should be truncated to achieve consistency with the slow convergence to normality, and to guarantee that the standard deviation of price variations remains finite [5]. Most recent developments display correlations and PDFs obeying q-Gaussian distribution on price increments of the NASDAQ [6] and Dow Jones indexes [7]. Consequently, the q-Gaussian distribution—an extension of the Gaussian distribution for correlated fluctuations—is more appropriate to account for the correlations of simple price increments [8] or log price increments [6,9–11]. The time correlations lead to anomalous diffusion, a phenomenon that is pervasive in strongly correlated classical systems, such as silo discharge [12] and velocity fluctuation in granulence of sheared granular flow [13]. Modifications to classical models have been proposed by the maximization of Tsallis entropy [6,10] to capture the features of price fluctuations. Examples are the stochastic processes with statistical feedback [6], the generalized autoregressive conditional heteroskedasticity (GARCH) algorithm [8], and the superstatistical approach with time scales for short and large return periods [9,11]. These approaches are based on q-Gaussian distribution to describe price return behavior.

In this paper we have analyzed the Standard & Poor's 500 (S&P500) stock market data during the 22 year period from January 1996 to May 2018, with an interval span of 1 min. The stock market index at time t is denoted by $I(t)$. The price return in a time interval from t_0 to t is defined by

$$X(t, t_0) = I(t_0 + t) - I(t_0). \tag{1}$$

The stock market index fluctuates over time in a random fashion. The main interest of economists is to predict the price return at any future time $t_0 + t$. Here we adopt the probabilistic approach: we assume that the price return X is a random variable with probability density function (PDF) $P_X(x, t)$. Then we formulate the governing equation for this distribution. The standard diffusion process is an oversimplification, as the price fluctuations strongly correlate for times of the order of minutes, and they weakly correlate for longer times [14]. A candidate for such a distribution function is the q-Gaussian distribution. The q-Gaussian is a generalization of the Gaussian distribution and is defined as [15]

$$g_q(x, \beta) = \frac{\sqrt{\beta}}{C_q} e_q(-\beta x^2), \tag{2}$$

where $e_q(x) = [1 + (1-q)x]^{\frac{1}{1-q}}$ is the q-exponential function. For $q > 1$, the q-Gaussian has asymptotic heavy-tail power law given by $g_q \sim 1/x^{2/(q-1)}$. The "q-Gauss" is a special case of q-Gaussian distribution defined as $g_q(x) = g_q(x, \beta = 1)$. The exponential and Gaussian functions can be recovered by taking the limit $q \to 1$. For $1 < q < 3$, the normalizing constant C_q is given by

$$C_q = \sqrt{\frac{\pi}{q-1}} \frac{\Gamma\left(\frac{3-q}{2(q-1)}\right)}{\Gamma\left(\frac{1}{q-1}\right)}. \tag{3}$$

The so-called q-central limit theorem states that the q-Gaussian is the limit of the distributions of specially correlated random processes [16]. Thus it is reasonable to propose the q-Gaussian as a candidate to fit the PDF distribution of stock markets. With this aim, we construct the PDF of

*kari0293@uni.sydney.edu.au

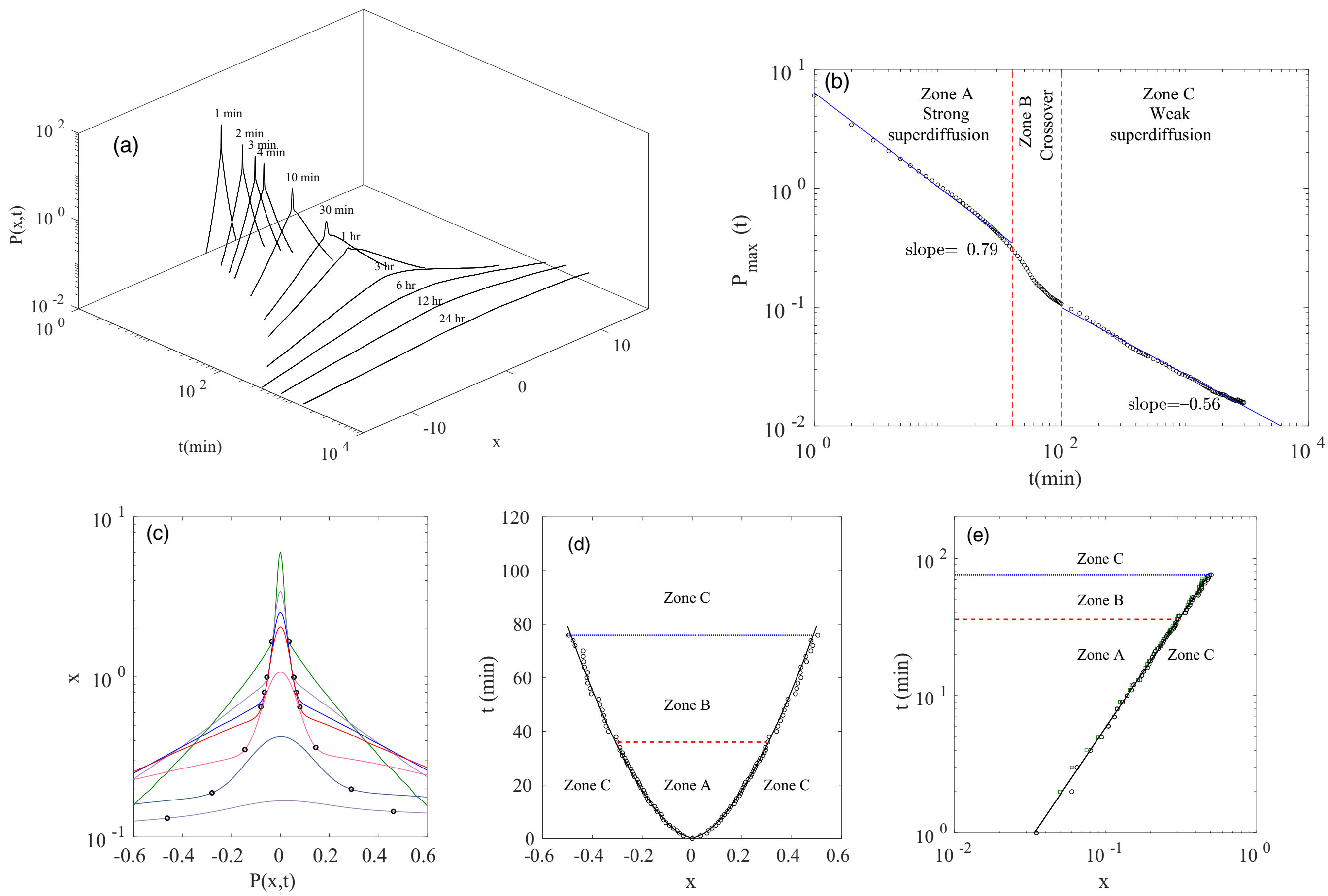

FIG. 1. (a) Time evolution of the PDF of price return. Initially the PDF has a pronounced bump in the center that fully disappears close to 78 min. (b) The time evolution of the height of the PDF. Two well-defined power laws are observed. (c) From (a) for the first hour analysis, time increases from top to bottom. The ends of the bump are obtained from the two points at the PDF with abrupt change of slope. These points correspond to a transition from strong to weak superdiffusion. (d) The circles represent the end points plotted against time that are fitted in (e) by the power law $x = \pm a(t/t_o)^\nu$, with $a = (3.39 \pm 0.01) \times 10^{-2}$, $t_o = 1$ min, and $\nu = 0.62 \pm 0.05$. This curve and the line $t = 35$ min (red dotted line) define the strong superdiffusion regime (zone A). The bump disappears completely at $t = 78$ min (blue dashed line). The remaining area corresponds to the weak superdiffusion regime (zone C). The crossover regime (zone B) is limited by the curve and 35 min $< t <$ 78 min. In this regime the bump still has not dissipated but experiences a transition from strong to weak superdiffusion.

the S&P500 stock market index using the kernel density estimator. The bandwidth of the kernel is set to $h = 0.005$ which is small enough to capture the nontrivial structure of the PDFs. Figures 1(a) and 1(b) show the time evolution of the PDF and its height from 1 min to 24 h of active market time. The initial distribution at $t = 1$ min consists of heavy tails and a pronounced bump at the center. This bump is easily distinguished from the rest of the distribution by an abrupt change of the slope of the distribution; see Fig. 1(c). The points where the abrupt change of the slope occurs are plotted against the time in Figs. 1(d) and 1(e). These points define the top and bottom boundary of what we called the domain of the bump. As time evolves, the bump diffuses and completely disappears after 78 min.

As shown in Fig. 1(b), the time variation of the height of the bump obeys a power law with exponent $P_{\max} \sim t^{-1/\alpha}$ with $\alpha = 1.26 \pm 0.04$ in the strong superdiffusion regime. This is different from the exponent $\alpha = 2$ expected in classical diffusion processes. Between $t = 38$ min and $t = 78$ min, we observe a crossover region. The end of the crossover corresponds to the region where the bump fully disappears, which is shown in Figs. 1(d) and 1(e). After the end of the crossover, the new height of the distribution obeys a different power law with exponent $\alpha = 1.79 \pm 0.01$, which is closer to the exponent of classical diffusion. Mantegna and Stanley's analysis on a more limited data set of the S&P500 index led to the exponent $\alpha = 1.40 \pm 0.05$ [17]. This is in reasonable agreement with our exponents, considering that they used a single exponent to fit the entire time evolution of the height.

Based on the domain of the bump and the time evolution of the height of the PDF, we partition the two-dimensional space (price and time) into three zones as shown in Fig. 1(d): zone A is the domain of the bump where the power law holds, zone B is the area of the bump's domain where the power law smoothly changes to another power law, and zone C is the remaining space. In zones A and C we propose a self-similar distribution given by

$$P(x,t) = \frac{1}{(Dt)^{\frac{1}{\alpha}}} f\left(\frac{x}{(Dt)^{\frac{1}{\alpha}}}\right), \qquad (4)$$

where $f(x)$ is a normalized distribution. The exponent α defines the diffusion process as follows. The second moment

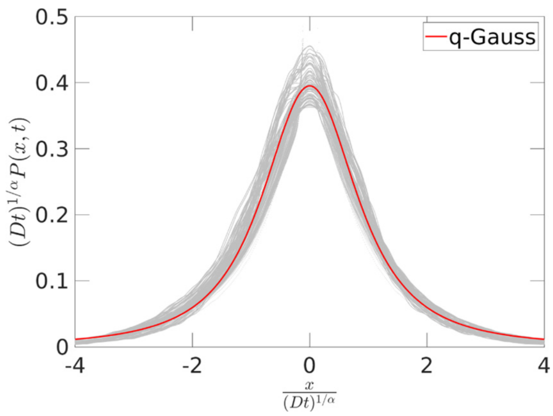

FIG. 2. Collapse of the PDFs of the weak superdiffusion regime (zone C). The exponent used for the collapse is $\alpha = 1.79 \pm 0.01$. The collapse data is fitted by a q-Gauss (red line) with exponent $q = 1.71 \pm 0.01$ and $D = 0.1118 \pm 0.005$ min^{-1}. Both exponents are related by Eq. (10).

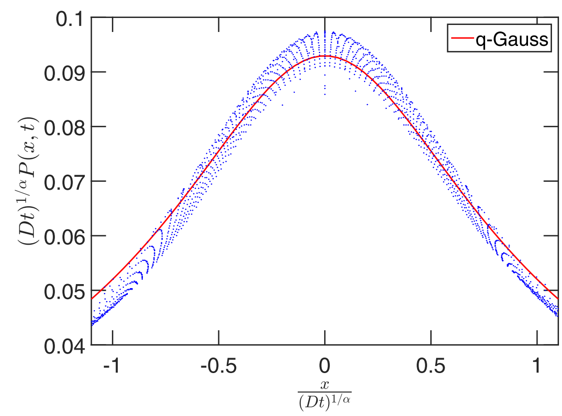

FIG. 3. Collapse of the PDFs of the strong superdiffusion regime (zone A). The exponent used for the collapse is $\alpha = 1.26 \pm 0.04$ and $D = (4.8 \pm 0.2) \times 10^{-3}$ min^{-1}. The collapse data is fitted by a q-Gauss (red line) with exponent $q = 2.73 \pm 0.005$.

of the distribution in Eq. (4) is $\langle x^2 \rangle \sim t^{\mu}$, where $\mu = 2/\alpha$ and it is known as the Tsallis diffusion exponent [18]. Then $\alpha < 2$ corresponds to superdiffusion, whereas $\alpha > 2$ leads to subdiffusion. The exponent α scales the height of the distribution as $P_{\max} \sim t^{-1/\alpha}$ as the previous fitting in Fig. 1(b). The present favorite alternative is $f(x) = L_\alpha(x)$ the Lévy distribution, as proposed by Mandelbrot [4], and Mantegna and Stanley [17]. They obtained $\alpha = 1.4$ as a fitted parameter, indicating superdiffusion. Here we propose a different approach by seeking a self-similar fitting for both the weak and strong superdiffusion regime using the q-Gauss function $f(x) = g_q(x)$ in Eq. (4). In both models the classical diffusion can be recovered by taking $\alpha = 2$ and $g(x) = g_1(x) = L_2(x)$. This limit corresponds to the self-similar solution of the diffusion equation that does not fit well to the stock market data.

The self-similar fitting in zones A and C is performed as follows. First we fit each PDF to Eq. (2) using q and β as fitted parameters. We evaluate the time dependence of the fitting parameters. For each zone, we found that q is approximately constant while β follows a power law relation that is written as $\beta = (Dt)^{-2/\alpha}$, where D and α are fitting parameters of this power law. Then, we collapse the PDFs for both weak and strong superdiffusion, as shown in Figs. 2 and 3.

To collapse the PDFs in the weak superdiffusion regime—zone C in Fig. 1(d)—we use the data from $t = 1$ min to $t = 3000$ min. The data is detrended by subtracting from the time series the average value within a time window of one month. This removes the effect of the drift on the PDF. We obtain an excellent agreement for the collapsed data with the q-Gauss distribution [Eq. (2) with $\beta = 1$]. The q-exponent in this regime is $q = 1.72 \pm 0.03$, which is larger than the value $q = 1$ expected for uncorrelated random processes. This is consistent with the weak correlation of the price fluctuations in this regime that is given by an autocorrelation of price fluctuating around 0.1%. The exponent $\alpha = 1.79 \pm 0.01$ is the same as the one calculated in Fig. 1(b) and is lower than the value of 2 expected for classic diffusion. The collapse of the PDFs in the strong superdiffusion regime—zone A in Fig. 1(d)—is shown in Fig. 3 like in the weak superdiffusion case; the collapse fits well to the q-Gauss distribution. However, in this case the exponent $q = 2.73 \pm 0.005$ is larger and $\alpha = 1.26 \pm 0.04$ is lower than the exponents in the weak superdiffusion regime. This indicates stronger deviation from classical diffusion. As in the previous case, these exponents should be considered as independent exponents. Note that in Fig. 3 we collapse only the bump of the distribution, while its tail is fitted in Fig. 2 because it belongs to the weak superdiffusion regime (zone C).

Next we construct the evolution equation of the PDF based on the q-Gaussian fitting. It is natural to relate the superdiffusion processes with a well-known anomalous diffusion model, which is given by a nonlinear Fokker-Planck equation [6], also known as the porous media equation [18]:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u^m}{\partial x^2}. \tag{5}$$

The Barenblatt solution of Eq. (5) for $m < 2$ and $t > 0$ is given by [19]

$$u_m(x,t) = \frac{1}{t^{\frac{1}{m+1}}}\left(C + \frac{m-1}{2m(m+1)}\frac{x^2}{t^{\frac{2}{m+1}}}\right)^{\frac{1}{m-1}}, \tag{6}$$

where C is an integration constant. There is only one exponent in Eq. (5), while we have two independent exponents α and q in the fitting of the collapsed data (see Figs. 2 and 3). An additional exponent ξ is introduced and proposing $m = 2 - q$ we write

$$P(x,t) = u_{2-q}(x,\tau), \quad \tau = (Bt)^{\xi}. \tag{7}$$

The scale parameter B is obtained as follows. By placing Eq. (7) into Eq. (6) we obtain

$$P(x,t) = \frac{1}{(Bt)^{\frac{\xi}{3-q}}}\left(C - \frac{1-q}{2(2-q)(3-q)}\frac{x^2}{(Bt)^{\frac{2\xi}{3-q}}}\right)^{\frac{1}{1-q}}.$$

This equation can be written as

$$P(x,t) = \frac{1}{C_q(Dt)^{\frac{1}{\alpha}}}\left(1 - (1-q)\frac{x^2}{(Dt)^{\frac{2}{\alpha}}}\right)^{\frac{1}{1-q}}, \tag{8}$$

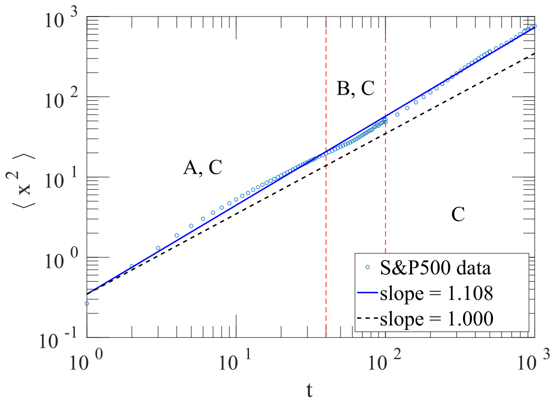

FIG. 4. Time evolution of the second moment of the full PDF of the S&P500 index. The data is compared to the best fitting and the linear fitting that corresponds to classical diffusion. The zones where the integration is performed are also shown.

in which C_q is the normalization constant in Eq. (3). The B and C values are related to D and C_q by

$$C_q = B^{\frac{1}{\alpha}} C^{\frac{1}{q-1}} D^{-\frac{1}{\alpha}}, \quad D = B[2C(2-q)(3-q)]^{\alpha/2}, \qquad (9)$$

and ξ is calculated in terms of α and q by

$$\xi = \frac{3-q}{\alpha}. \qquad (10)$$

The parameters α, q, and D are derived by fitting the collapse data. Equation (8) corresponds to Eq. (4) with $f(x) = g_q(x)$ defined by Eq. (2). Finally, the governing equation of $P(x,t)$ is obtained by first taking the partial time derivative in Eq. (7)

$$\frac{\partial P}{\partial t} = \frac{\partial P}{\partial \tau}\frac{\partial \tau}{\partial t} = \frac{\partial u_{2-q}}{\partial t}(Bt)^{\xi-1}\xi B. \qquad (11)$$

Placing Eq. (11) into Eq. (5) the governing equation is obtained:

$$t^{1-\xi}\frac{\partial P}{\partial t} = \xi D^{\xi}\frac{\partial^2 P^{2-q}}{\partial x^2}. \qquad (12)$$

Equation (12) is of great importance as it is the evolution equation of the distribution of price returns of the S&P500 index. The parameters q, ξ, and D take different values depending on whether we are in the weak or strong superdiffusion regime. Equation (12) constitutes a particular case of the Tsallis-Bukman scaling law when $\xi = 1$ [20] (corresponding to the case $\mu = 1$ and $\nu = 2 - q$ for the symbols used in the original paper of Tsallis and Bukman). The variable x scales like $t^{\xi/(3-q)}$ from Eq. (8). Analogous to the Tsallis-Bukman law, the second moment is $\langle x\rangle^2 \propto t^{\frac{2\xi}{3-q}}$ for the anomalous diffusion case [13]. By using Eq. (10) we obtain $\langle x\rangle \propto t^{\frac{2}{\alpha}}$, which matches the second moment display on Fig. 4. Also, Eq. (12) reflects a distinct feature of stock markets, namely that the diffusion coefficient depends on the diffused quantity itself, a feature that is observed also in many biological and physical systems [21,22]. Then, the Black-Scholes coefficient of diffusion can be calculated by comparing Eq. (12) to the linear Fokker-Planck equation $\frac{\partial P}{\partial t} = \frac{\partial^2 (D_2 P)}{\partial x^2}$ [1]. The direct comparison leads to $D_2 = \xi D^{\xi} P^{1-q} t^{\xi-1}$. Replacing Eqs. (8) and (10) into this equation we obtain an explicit expression for the coefficient of diffusion:

$$D_2(x,t) = \frac{(3-q)D^{\frac{2}{\alpha}}}{\alpha C_q^{1-q} t^{\frac{\alpha-2}{\alpha}}}\left(1-(1-q)\frac{x^2}{(Dt)^{\frac{2}{\alpha}}}\right). \qquad (13)$$

For a fixed time and large price fluctuations, the scaling $D_2 \sim x^2$ of the Black-Scholes equation for geometric Brownian motion is recovered [1]. Our extension introduces the power law dependency with time into this relation that accounts for superdiffusion.

The strong superdiffusion regime occurs in a very small zone in the phase space, so that it only contributes to the moments of the PDF distributions to a small extent. We characterize the *global diffusion* by plotting the second moment of the full PDF as a function of time, and fitting it to the best power law. The result is shown in Fig. 4. The second moment is almost linear with time, suggesting that the global diffusion is weakly superdiffusive.

We have identified two regimes in the time evolution of the PDFs of the price return of the S&P500 index that account for strong and weak superdiffusion. Both regimes are described by the family of self-similar q-Gaussian distributions that accurately capture the tail distributions, needed for financial risk estimates, by accounting for the strong superdiffusion in the center of the distribution. The time evolution of the PDF is consistent with the central limit theorem for correlated fluctuations. The strong correlations in the short-time regime are reflected in a strong superdiffusive q-Gaussian regime that has not previously been reported and the weak correlation in the long-time regime corresponds to the weak superdiffusive q-Gaussian regime. These regimes become evident only after the analysis of high-frequency data. We demonstrated that these regimes are well described by a nonlinear Fokker Plank equation that is used to obtain an explicit formula for the Black-Scholes diffusion coefficient.

H.J.H. thanks CAPES and FUNCAP for support. M.H. acknowledges Australian Research Council Grant No. DP170102927. K.A.C. thanks The Sydney Informatics Hub at The University of Sydney for providing access to HPC-Artemis for financial data processing. F.A.M. thanks Sornette and Tsallis for inspiring discussions.

[1] G. L. Vasconcelos, Braz. J. Phys. **34**, 1039 (2004).
[2] Z.-Q. Jiang, W.-X. Zhou *et al.*, Phys. Rev. E **84**, 016106 (2011).
[3] K. Kanazawa, T. Sueshige, H. Takayasu, and M. Takayasu, Phys. Rev. Lett. **120**, 138301 (2018).
[4] B. Mandelbrot, J. Business **36**, 394 (1963).
[5] R. N. Mantegna and H. E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
[6] L. Borland, Phys. Rev. Lett. **89**, 098701 (2002).
[7] J. Ludescher and A. Bunde, Phys. Rev. E **90**, 062809 (2014).
[8] S. D. Queirós and C. Tsallis, Eur. Phys. J. B **48**, 139 (2005).

[9] E. Van der Straeten and C. Beck, Phys. Rev. E **80**, 036108 (2009).

[10] Y. A. Katz and L. Tian, Physica A **392**, 4989 (2013).

[11] D. Xu and C. Beck, Physica A **453**, 173 (2016).

[12] R. Arévalo, A. Garcimartin, and D. Maza, Eur. Phys. J. E **23**, 191 (2007).

[13] G. Combe, V. Richefeu, M. Stasiak, and A. P. Atman, Phys. Rev. Lett. **115**, 238301 (2015); D. Tsuji and H. Katsuragi, Phys. Rev. E **92**, 042201 (2015); G. Combe, V. Richefeu, G. Viggiani, S. A. Hall, A. Tengattini, and A. P. F. Atman, in *Powders and Grains 2013: Proceedings of the 7th International Conference on Micromechanics of Granular Media*, edited by A. Yu, K. Dong, R. Yang, and S. Luding, AIP Conf. Proc. 1542 (AIP, New York, 2013), pp. 453–456; V. Richefeu, G. Combe, and G. Viggiani, Géotech. Lett. **2**, 113 (2012).

[14] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, London, 1999).

[15] S. Umarov, C. Tsallis, and S. Steinberg, Milan J. Math. **76**, 307 (2008).

[16] U. Tirnakli, C. Beck, and C. Tsallis, Phys. Rev. E **75**, 040106 (2007).

[17] R. N. Mantegna and H. E. Stanley, Nature (London) **376**, 46 (1995).

[18] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Springer Science & Business Media, New York, 2009); A. Pereira, J. Fernandes, A. Atman, and J. Acebal, Physica A **509**, 369 (2018).

[19] J. L. Vazquez, *The Porous Medium Equation: Mathematical Theory* (Clarendon, Oxford, 2007).

[20] C. Tsallis and D. J. Bukman, Phys. Rev. E **54**, R2197 (1996).

[21] V. I. Bogachev, N. V. Krylov, M. Rockner, and S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov Equations* (American Mathematical Society, Providence, RI, 2015), Vol. 207.

[22] T. D. Frank, *Nonlinear Fokker-Planck Equations: Fundamentals and Applications* (Springer, Berlin, 2005).