Volume 384, issue 1 1 October 2007 ISSN 0378-4371

ELSEVIER

PHYSICA A
STATISTICAL MECHANICS AND ITS APPLICATIONS
Recognized by the European Physical Society

STATPHYS - KOLKATA VI

Proceedings of the International Conference on Statistical Physics, held at Raichak (Fort Radisson Resort) and Kolkata (Saha Institute of Nuclear Physics and University of Calcutta), India, January 5–9, 2007

Guest Editors:
Parongama Sen
Pradeep Kumar Mohanty

Available online at
ScienceDirect
www.sciencedirect.com

http://www.elsevier.com/locate/physa

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Physica A 384 (2007) 21–24
PHYSICA A
www.elsevier.com/locate/physa

Scaling properties and entropy of long-range correlated time series

Anna Carbone[a,*], H. Eugene Stanley[b]

[a]Dipartimento di Fisica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[b]Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA

Available online 8 May 2007

Abstract

We calculate the Shannon entropy of a time series by using the probability density functions of the characteristic sizes of the long-range correlated clusters introduced in [A. Carbone, G. Castelli, H.E. Stanley, Phys. Rev. E **69** (2004) 026105]. We define three different measures of the entropy related, respectively, to the length, the duration and the area of the clusters. For all the three cases, the entropy increases as the logarithm of a power of the size with exponents equal to the fractal dimension of the cluster length, duration and area, respectively.

Keywords: Time series analysis; Shannon entropy; Hurst exponent

The behavior of many complex systems is probed by detecting a variable over a certain temporal or spatial range obtaining a record of the relevant observable as a function of time or space (e.g., time series or character sequence). The challenge is to deduce detailed information from the data in order to gain deeper physical insight into the system's dynamics and the origin of complex behavior. In this regard, the concept of *self-organized criticality* (SOC) has been used to explain why many systems behave like nonequilibrium systems right at a phase transition temperature, spontaneously self-organizing themselves into states characterized by algebraic correlations—unlike systems in equilibrium for which tuning is essential [1–8].

The aim of this paper is to propose a method to calculate the entropy of a long-range correlated stochastic sequence. Methods of calculation of the entropy of sequences have been proposed in Refs. [12–18]. The entropy is usually calculated according to the well-known Shannon definition [9]:

$$S(x) \equiv -\sum_i P(x_i) \log P(x_i). \tag{1}$$

Here we calculate the entropy $S(x)$ of the sequence by using the probability distribution functions (pdfs) of the clusters introduced in Refs. [10,11].

*Corresponding author. Fax: +39 011 5647399.
E-mail address: anna.carbone@polito.it (A. Carbone).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2007.04.105

Consider a long-range correlated time series $y(t)$ with correlation exponent H (Hurst exponent) [19]. In Refs. [10,11] it was shown that when the function

$$\tilde{y}_n(t) \equiv \frac{1}{n}\int_0^n y(t-t_0)\,\mathrm{d}t_0 \tag{2}$$

intersects $y(t)$, the fractional Brownian walk is partitioned in a sequence of elementary *paths*, with n denoting the time window over which the integral (2) is calculated. Fig. 1 shows the function $\widetilde{y}_n(t)$ with $n = 1000$.

The length of the segment of the time series between two subsequent intersections between $y(t)$ and $\widetilde{y}_n(t)$ is defined by

$$\ell_{j,n} \equiv \int_{t_c(j)}^{t_c(j+1)} \delta\ell(t)\,\mathrm{d}t, \tag{3}$$

where $\delta\ell(t)$ is the length of each elementary step of the fractional walker. The pdf $P(\ell)$ has been obtained by counting the segments $\mathcal{N}(\ell_1), \mathcal{N}(\ell_2), \ldots, \mathcal{N}(\ell_j)$ having, respectively, lengths $\ell_1, \ell_2, \ldots, \ell_j$, up to a given value of the index j, where

$$\mathcal{N}(n) = \sum_j \mathcal{N}(\ell_j, n). \tag{4}$$

The pdfs of the random variables ℓ is a power law:

$$P(\ell) \sim \ell^{-(2-H)}, \tag{5}$$

where the exponent $2 - H$ coincides with the fractal dimension D of the time series. The intersection of the functions $y(t)$ and $\widetilde{y}_n(t)$ generates a sequence of *clusters* besides the elementary *paths* discussed above. The *clusters* correspond to the regions bounded by $y(t)$ and $\widetilde{y}_n(t)$ between two subsequent crossing points $t_c(j)$ and $t_c(j+1)$. The durations and areas of the clusters have been defined, respectively, by

$$\tau_j \equiv \int_{t_c(j)}^{t_c(j+1)} \mathrm{d}t \tag{6}$$

and

$$\mathcal{A}_j \equiv \int_{t_c(j)}^{t_c(j+1)} |y(t) - \widetilde{y}_n(t)|\,\mathrm{d}t. \tag{7}$$

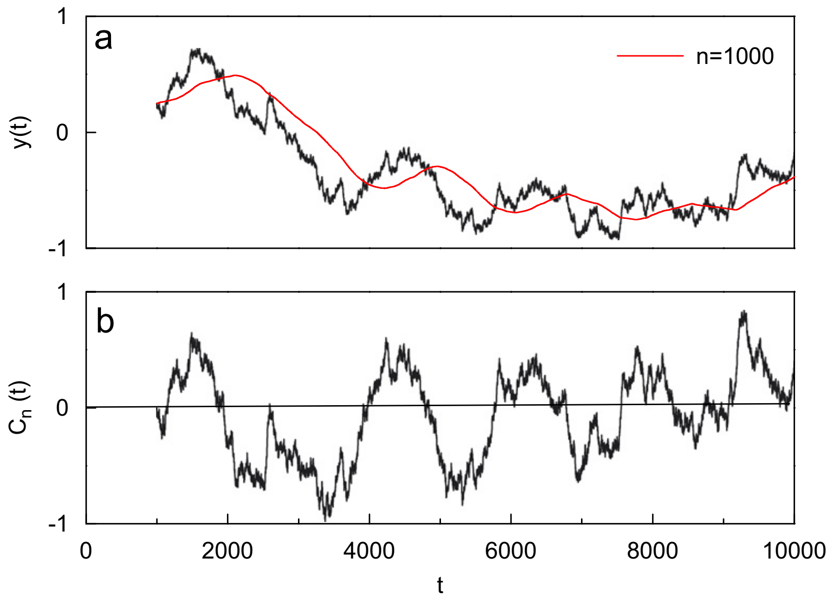

Fig. 1. (Color online) (a) A fractional Brownian path (black), with $H = 0.5$, and the functions $\widetilde{y}_n(t)$ for $n = 1000$ (red). (b) The function $C_n(t) \equiv y(t) - \widetilde{y}_n(t)$ for $n = 1000$.

The scaling relationships $\ell_\tau \sim \tau_n^{\psi_\ell}$ and $\mathscr{A}_{\tau,n} \sim \tau_n^{\psi_\mathscr{A}}$ hold, where $\ell_{\tau,n}$ and $\mathscr{A}_{\tau,n}$ are the values of the path length and cluster area, obtained by averaging $\ell_{j,n}$ and $\mathscr{A}_{j,n}$ over the subset of paths/clusters having the same average lifetime τ. The pdfs $P(\tau)$ and $P(\mathscr{A})$ can be obtained, similarly to $P(\ell)$, by counting all clusters having, respectively, duration $\tau_1, \tau_2, \ldots, \tau_j$ and areas $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_j$ up to a given value of the index j. The pdfs $P(\tau)$ and $P(\mathscr{A})$ have been found to be, respectively [10,11]:

$$P(\tau) \sim \tau^{-(2-H)} \tag{8}$$

and

$$P(\mathscr{A}) \sim \mathscr{A}^{-2/(1+H)}. \tag{9}$$

The pdfs $P(\ell)$, $P(\tau)$ and $P(\mathscr{A})$ will be used to evaluate the entropy of the time series. In order to do so we consider the Shannon entropy expression (Eq. (1)) where the probability distribution function $P(\tau)$ is introduced:

$$S(\tau) \equiv -\sum_{\mu(\tau)} P(\tau) \log P(\tau) \tag{10}$$

and the sum is performed over elementary cells with size: $\mu(\tau) = \tau^{2-H}$. The quantity $\mu(\tau)$ represents the size of the elementary cell spanned by a particle experiencing an elementary Brownian path between two subsequent intersections $t_c(j)$ and $t_c(j+1)$. Note that the elementary cell coincides, respectively, with a one-dimensional line for $H = 1$ and with a square for $H = 0$. Using Eq. (8), Eq. (10) becomes

$$S(\tau) = S_{0,\tau} + (2 - H) \log \tau, \tag{11}$$

where $S_{0,\tau}$ represents the constant term of the sum. Eq. (11) refers to the entropy that would be produced by the Brownian walker in the absence of finite-size effects.

The entropies $S(\ell)$ and $S(\mathscr{A})$, related, respectively, to the distributions of the cluster lifetime and area, can be calculated as they were for the entropy $S(\tau)$. We obtain for the cluster length:

$$S(\ell) = S_{0,\ell} + (2 - H) \log \ell \tag{12}$$

and for the cluster area:

$$S(\mathscr{A}) = S_{0,\mathscr{A}} + \frac{2}{1+H} \log \mathscr{A}. \tag{13}$$

The probability distribution functions and the five exponents ψ_ℓ, $\psi_\mathscr{A}$, α, β and γ, are reported in Table 1. The third column contains the expression of the entropy calculated according to the proposed approach. Note that for the simple random walk, H is equal to 0.5 and the values of these exponents coincide with those of the Dhar–Ramaswamy model [2].

A plot of $S(\tau)$ is shown in Fig. 2 for different values of the fractal dimension $D = 2 - H$, respectively with Hurst exponent $H = 0.0$, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.

The entropy takes the minimum value for $\tau \to 1$, as expected for a system characterized by minimum disorder. Conversely, when τ increases, clusters of a huge range of sizes are generated, and $P(\tau)$ spreads resulting in the increase of entropy, i.e., in the increase of disorder. In the absence of finite-size effects, $S(\tau)$ behaves as $\log \tau$.

Table 1
Scaling relations, exponents and entropy relations

	Scaling relations	Exponents	Entropy
Length	$\ell_\tau \sim \tau^{\psi_\ell}$	$\psi_\ell = 1$	–
Area	$\mathscr{A}_\tau \sim \tau^{\psi_\mathscr{A}}$	$\psi_\mathscr{A} = 1 + H$	–
Length pdf	$P(\ell) \sim \ell^{-\alpha}$	$\alpha = 2 - H$	$S(\ell) = S_{0,\ell} + \alpha \log \ell$
Lifetime pdf	$P(\tau) \sim \tau^{-\beta}$	$\beta = 2 - H$	$S(\tau) = S_{0,\tau} + \beta \log \tau$
Area pdf	$P(\mathscr{A}) \sim \mathscr{A}^{-\gamma}$	$\gamma = 2/(1 + H)$	$S(\mathscr{A}) = S_{0,\mathscr{A}} + \gamma \log \mathscr{A}$

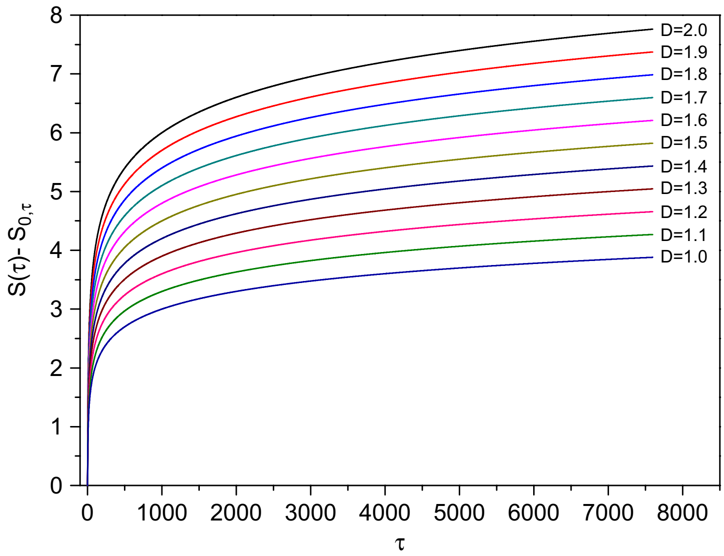

Fig. 2. (Color online) Plot of the entropy $S(\tau) - S_{0,\tau}$ for different values of the fractal dimension $D = 2 - H$. It can be observed that the entropy is larger for signals with larger fractal dimension, i.e., with smaller Hurst exponent.

It is worthy to point to the connection with the Boltzmann relation $S = S_0 + \log \Omega$ with Ω the phase space volume. Bearing in mind that $2 - H$ is equal to the fractal dimension D, the quantity ℓ^D corresponds to the volume occupied by the fractal of length ℓ. Therefore, the entropy $S(\ell)$, varying as the logarithm ℓ^D, expresses the dependence on the fractal volume ℓ^D. Similar considerations hold for the entropies $S(\tau)$ and $S(\mathscr{A})$. As expected, all the three entropies decrease when H increases, in agreement with the fact that larger values of H correspond to more "ordered" signals.

In summary, in this work we have put forward a link between the power-law scaling of a long-range correlated series and the Shannon entropy using the pdfs of the clusters introduced in Refs. [10,11]. We obtain a family of entropies, resp. Eqs. (11)–(13) increasing as the generalized volumes of the cluster sizes with exponent related to the fractal dimension of the series. In particular, we find that anticorrelated time series, with Hurst exponent $0 < H < 0.5$ are characterized by entropies greater than correlated time series having $0.5 < H < 1$.

References

[1] P. Bak, C. Tang, K. Wiesenfeld, Phys. Rev. Lett. 59 (1987) 381.
[2] D. Dhar, R. Ramaswamy, Phys. Rev. Lett. 63 (1989) 1659.
[3] S.S. Manna, J. Physica A 24 (1991) L363.
[4] A. Erzan, S. Sinha, Phys. Rev. Lett. 66 (1991) 2750.
[5] A. Vespignani, R. Dickman, M.A. Munoz, S. Zapperi, Phys. Rev. E 62 (2000) 4564.
[6] M.E.J. Newman, Contemp. Phys. 46 (2005) 323.
[7] L. de Arcangelis, C. Perrone-Capano, H.J. Herrmann, Phys. Rev. Lett. 96 (2006) 028107.
[8] H.J. Jensen, Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems, Cambridge Lecture Notes in Physics, 1998.
[9] C.E. Shannon, Bell Syst. Tech. J. 27 (1948) 379.
[10] A. Carbone, G. Castelli, H.E. Stanley, Phys. Rev. E 69 (2004) 026105.
[11] A. Carbone, H.E. Stanley, Physica A 340 (2004) 544.
[12] P. Grassberger, I. Procaccia, Phys. Rev. Lett. 50 (1983) 346.
[13] T. Schurmann, P. Grassberger, Chaos 69 (1996) 414.
[14] S.M. Pincus, Ann NY Acad. Sci. 954 (2001) 245.
[15] M. Costa, A.L. Goldberger, C.K. Peng, Phys. Rev. Lett. 89 (2002) 068102.
[16] N. Scafetta, P. Grigolini, Phys. Rev. E 66 (2002) 036130.
[17] J.P. Crutchfield, D.P. Feldman, Chaos 13 (2003) 25.
[18] D.G. Perez, L. Zunino, M. Garavaglia, O.A. Rosso, Physica A 365 (2006) 282.
[19] B.B. Mandelbrot, J.W. van Ness, SIAM Rev. 4 (1968) 422.