

## Preferential attachment growth model and nonextensive statistical mechanics

D. J. B. SOARES<sup>1,2</sup>, C. TSALLIS<sup>3,4</sup>, A. M. MARIZ<sup>1</sup> and L. R. DA SILVA<sup>1</sup>

<sup>1</sup> *Departamento de Física Teórica e Experimental*

*Universidade Federal do Rio Grande do Norte*

*Campus Universitario, 59072-970 Natal-RN, Brazil*

<sup>2</sup> *Departamento de Física, Universidade Federal do Ceará*

*60451-970, Fortaleza-Ce, Brazil*

<sup>3</sup> *Centro Brasileiro de Pesquisas Físicas - Rua Xavier Sigaud 150*

*22290-180 Rio de Janeiro-RJ, Brazil*

<sup>4</sup> *Santa Fe Institute - 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

received 27 October 2004; accepted in final form 8 February 2005

published online 9 March 2005

PACS. 05.70.Ln – Nonequilibrium and irreversible thermodynamics.

PACS. 89.75.Hc – Networks and genealogical trees.

PACS. 89.75.-k – Complex systems.

**Abstract.** – We introduce a two-dimensional growth model where every new site is located, at a distance  $r$  from the barycenter of the pre-existing graph, according to the probability law  $1/r^{2+\alpha_G}$  ( $\alpha_G > 0$ ), and is attached to (only) one pre-existing site with a probability  $\propto k_i/r_i^{\alpha_A}$  ( $\alpha_A \geq 0$ ;  $k_i$  is the number of links of the  $i$ -th site of the pre-existing graph, and  $r_i$  its distance to the new site). Then we numerically determine that the probability distribution for a site to have  $k$  links is asymptotically given, for all values of  $\alpha_G$ , by  $P(k) \propto e_q^{-k/\kappa}$ , where  $e_q^x \equiv [1 + (1 - q)x]^{1/(1-q)}$  is the function naturally emerging within nonextensive statistical mechanics. The entropic index is numerically given (at least for  $\alpha_A$  not too large) by  $q = 1 + (1/3)e^{-0.526 \alpha_A}$ , and the characteristic number of links by  $\kappa \simeq 0.1 + 0.08 \alpha_A$ . The  $\alpha_A = 0$  particular case belongs to the same universality class to which the Barabasi-Albert model belongs. In addition to this, we have numerically studied the rate at which the average number of links  $\langle k_i \rangle$  increases with the scaled time  $t/i$ ; asymptotically,  $\langle k_i \rangle \propto (t/i)^\beta$ , the exponent being close to  $\beta = \frac{1}{2}(1 - \alpha_A)$  for  $0 \leq \alpha_A \leq 1$ , and zero otherwise. The present results reinforce the conjecture that the microscopic dynamics of nonextensive systems typically build (for instance, in Gibbs  $\Gamma$ -space for Hamiltonian systems) a scale-free network.

*Introduction.* – Among the subjects that are being studied intensively nowadays in statistical physics, there are two, namely *nonextensive statistical mechanics* (see [1] for a review) and *networks* [2], in particular *scale-free* networks [3], which receive special attention in connection with *complex systems* [4–7]. Could these two topics be intimately related? This would not be so surprising after all, given the fact that both research lines frequently address similar types of natural and artificial systems, in physics, economics, chemistry, biology, linguistics, social sciences and others. In fact, such a connection has already been conjectured in several

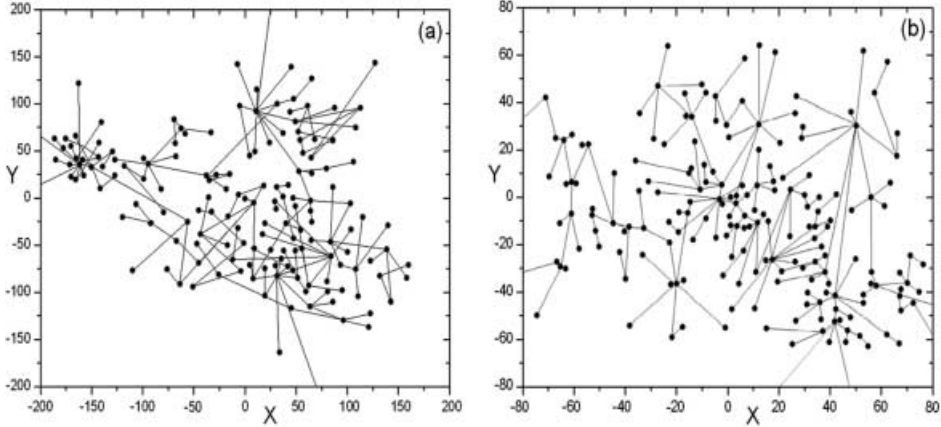


Fig. 1 – Typical  $N = 250$  network for (a)  $(\alpha_G, \alpha_A) = (1, 0)$  and (b)  $(\alpha_G, \alpha_A) = (1, 4)$ . The starting site is at  $(X, Y) = (0, 0)$ . Notice the spontaneous emergence of hubs.

occasions, *e.g.*, ref. [1], (Preface and Chapt. 1) and [8]. In the present paper we propose a growth model, on which we exhibit and quantitatively explore this connection.

*Simulation model.* – Let us consider a continuous plane. We shall construct a single connected network of *sites* (or *nodes* or *vertices*) and *links* (or *bonds* or *edges*) by gradually (sequentially) making it grow. We first fix one site ( $i = 1$ ) at some arbitrary origin of the plane. The second site ( $i = 2$ ) is randomly and isotropically chosen at a distance  $r$  distributed according to the probability law  $P_G(r) \propto 1/r^{2+\alpha_G}$  ( $\alpha_G > 0$ ;  $G$  stands for *growth*). This second site is then linked to the first one. To locate the third site ( $i = 3$ ), we move the origin to the barycenter of the two first sites, and apply again the distribution  $P_G(r)$  from this new origin. This third site is now going to be linked to only one of the pre-existing two sites. To do this, we use an attachment probability  $p_A \propto k_i/r_i^{\alpha_A}$  ( $\alpha_A \geq 0$ ;  $A$  stands for *attachment*), where  $r_i$  is the distance of the newly arrived site to the  $i$ -th site of the pre-existing cluster, and the *connectivity*  $k_i$  is the number of links already arriving to the same  $i$ -th site (at the present stage,  $k_1 = k_2 = 1$ ). The growth-attachment process is sequentially repeated as long as we want; the sites have unit radius, and no new site is admitted which overlaps with a previous one. If we denote with  $N$  the total number of sites of the cluster, it immediately follows that the linking of the newly arrived site ( $i = N$ ) is done with the probability

$$p_A = \frac{k_i/r_i^{\alpha_A}}{\sum_{j=1}^{N-1} k_j/r_j^{\alpha_A}}. \quad (1)$$

It is clear that the dynamics of this model makes the arriving sites to have preferential attachment to the previous sites that already have many links (*hubs*), *as long as they are not too far*. This competition (already explored in [7] for the particular case of a fractal distribution of sites) between connectivity and (Euclidean) proximity is less pronounced when  $\alpha_A$  is close to zero, and completely disappears only at  $\alpha_A = 0$ . For this particular case, one expects behaviors consistent with the Barabasi-Albert model [5], which has topology but no metrics. In the present paper, we focus on two main aspects: i) the stationary-state connectivity distribution  $P(k)$  associated with the number of sites that have  $k$  links in the  $N \rightarrow \infty$  limit; ii) the time dependence of the average number  $\langle k_i \rangle$  of links, more precisely how  $\langle k_i \rangle$  grows with the scaled time  $t/i$  ( $t \geq i$ ), particularly in the limit  $t/i \rightarrow \infty$  (see [5]).

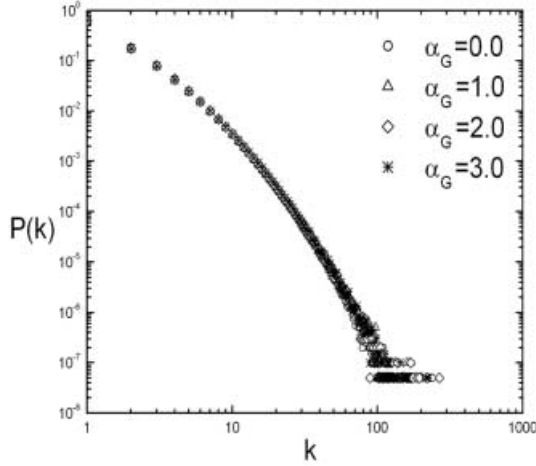


Fig. 2 – Connectivity distribution for  $\alpha_A = 1$  and typical values of  $\alpha_G$ ; 2000 realizations of  $N = 10000$  networks.

Typical networks obtained with this model are indicated in fig. 1. Our numerical results for  $P(k)$  are indicated in figs. 2 and 3. We illustrate in fig. 2 the fact that  $P(k)$  does not depend on  $\alpha_G$  for any given value of  $\alpha_A$ . The exponent  $\alpha_G$  controls the metrics of the emerging cluster, but has no influence on the connectivity distribution (due, at least in part, to the fact that the barycenter of the graph asymptotically converges onto a fixed point in space). In contrast, this distribution is greatly influenced by the exponent  $\alpha_A$ , as illustrated in fig. 3. We can verify that all our examples are very well fitted with the form

$$P(k) = P(0) e_q^{-k/\kappa}, \quad (2)$$

where the  $q$ -exponential function is defined as follows [1,9]:

$$e_q^x \equiv [1 + (1 - q)x]^{1/(1-q)} \quad (e_1^x = e^x); \quad (3)$$

$\kappa > 0$  is a characteristic number of links. In fig. 4 we show the  $\alpha_A$  dependences of  $q$  and  $\kappa$ . In particular,  $q(\alpha_A)$  does not exhibit a critical value of  $\alpha_A$  above which a different regime could emerge. However, unless a detailed analysis (out of the aim of the present letter) is done of the finite-size effects, such a possibility should not be excluded.

*Results and discussions.* – We must now address a relevant point: Why have we fitted our curves with  $q$ -exponentials? The reason lies on the conjectured connection with nonextensive statistical mechanics [10] (see [1,11] for reviews), a theory which we briefly illustrate now. Let us consider the following generalization of the Boltzmann-Gibbs entropy  $S_{BG} = - \int dk P(k) \ln P(k)$ :

$$S_q = \frac{1 - \int dk [P(k)]^q}{q - 1} \quad (q \in \mathcal{R}; S_1 = S_{BG}). \quad (4)$$

If we optimize this entropy with the constraints  $\int dk P(k) = 1$  and  $\{\int dk k [P(k)]^q\} / \{\int dk [P(k)]^q\} = K$ , we obtain straightforwardly [12]

$$P(k) = \frac{e_q^{-\beta_q(k-K)}}{\int dk' e_q^{-\beta_q(k'-K)}}, \quad (5)$$

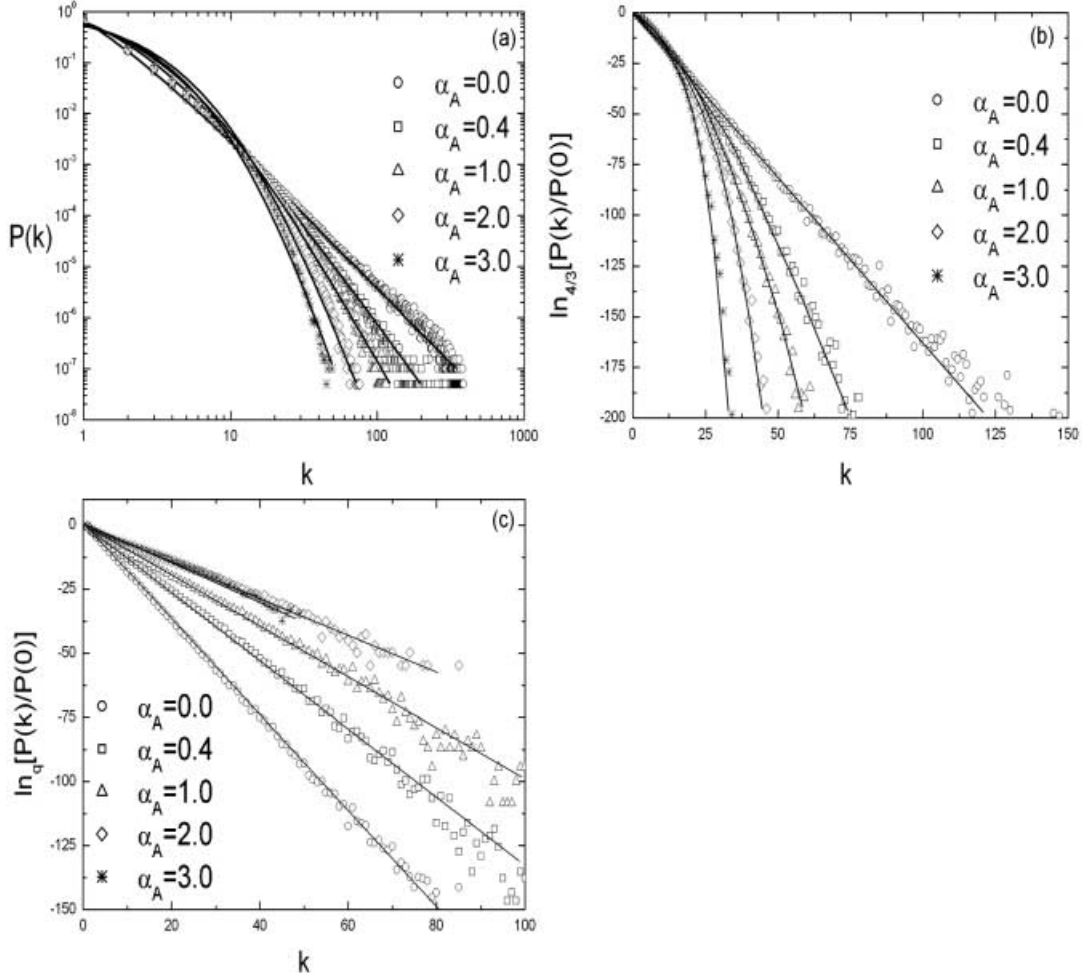


Fig. 3 – Connectivity distribution for typical values of  $\alpha_A$  (we have used  $\alpha_G = 2$ , but we recall that this value is irrelevant). Points are our computer simulation results; continuous lines are the best fits with the  $q$ -exponential function indicated in eq. (2). (a) log-log representation; (b)  $\ln_{4/3}$  —linear representation, with  $\ln_q x \equiv \frac{x^{1-q}-1}{1-q}$ ; (c)  $\ln_q$  —linear representation, where, for each value of  $\alpha_A$ , we have used its corresponding value of  $q$ . We have used 3 different representations to improve comprehension.

where  $\beta_q$  is simply related with a Lagrange parameter. This expression precisely coincides with eq. (2) through the identification

$$P(0) \equiv \frac{e_q^{\beta_q K}}{\int dk' e_q^{-\beta_q(k'-K)}}, \quad (6)$$

$$\kappa \equiv \frac{1 + (1-q)\beta_q K}{\beta_q}. \quad (7)$$

Before proceeding with our numerical results, let us briefly mention what indications make us believe that a close connection, and not just a mere functional coincidence, might exist between the present preferential-attachment growth model and the thermostistical systems

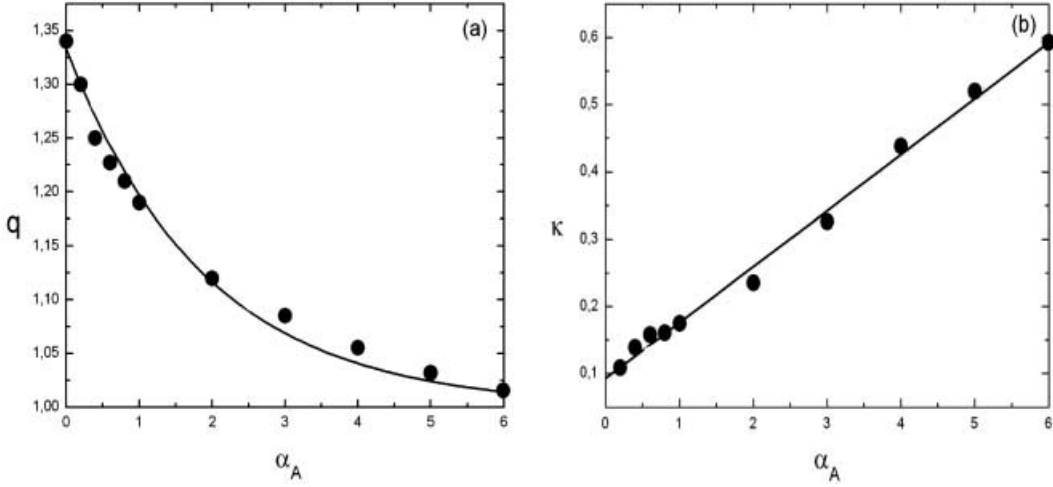


Fig. 4 – Values of  $q$  and  $\kappa$  used in the best fits indicated in fig. 3. The solid curves are: (a)  $q = 1 + (1/3)e^{-0.526 \alpha_A}$  ( $\forall \alpha_G$ ); (b)  $\kappa \simeq 0.083 + 0.092 \alpha_A$  ( $\forall \alpha_G$ ).

addressed by the entropy (4). First, a growth model involving, like the present one, preferential attachment has been proposed and both analytically and numerically discussed in [13]. In this model the connectivity distribution is analytically shown to be precisely of the form (2) (although written in a slightly different manner). The corresponding entropic index is given (see [1], Chapt. 1) by  $q = [2m(2 - r) + 1 - p - r] / [m(3 - 2r) + 1 - p - r]$ , where  $(m, p, r)$  are parameters of the model. Second, Lennard-Jones small clusters (with up to 14 atoms) have been numerically studied recently [14]. The distributions of the number of local minima of the potential energy with  $k$  neighboring saddle points in the configurational phase space can,

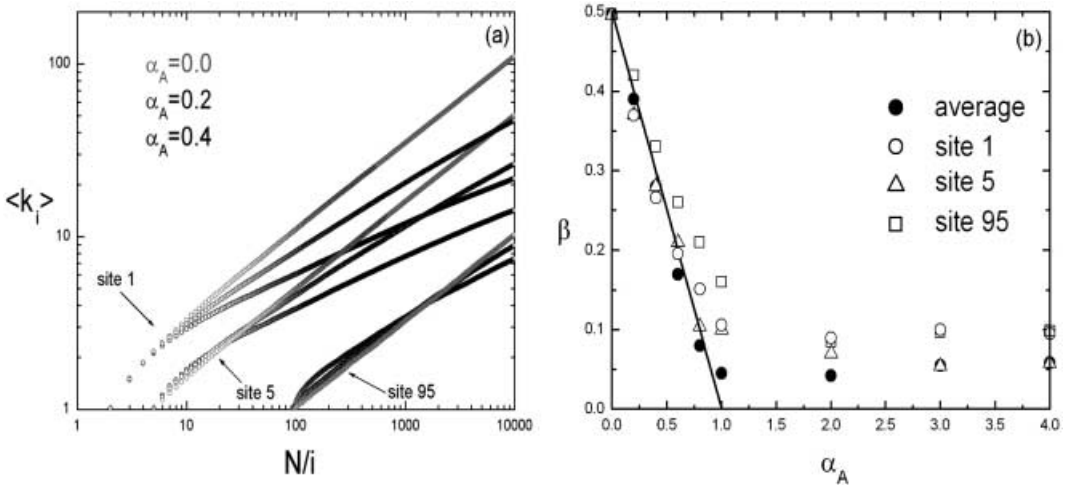


Fig. 5 – (a) Time dependence ( $t = N$ ) of the average number (over 2000 realizations) of links for typical values of  $\alpha_A$  for sites  $i = 1$ ,  $i = 5$  and  $i = 95$  (compare with fig. 2(c) of [5]). We have used  $\alpha_G = 2$ ; (b)  $\alpha_A$ -dependence of  $\beta$  (the straight line  $\beta = \frac{1}{2}(1 - \alpha_A)$  could be the exact answer).

although not mentioned in the original paper [14], be quite well fitted with  $q$ -exponentials with  $q = 2$  [8]. Third, the present model generates structures that are scale-free (see, for instance, fig. 1). Consistently, there is plethoric evidence of the connections of the entropy  $S_q$ , and of its associated statistics, with hierarchical and (multi)fractal structures (see, for instance, [15–18]). It is clear that none of these three features constitutes a proof; however, the set of them suggests the quite plausible scenario that the present growth model basically satisfies the dynamical requirements for the nonextensive concepts to be applicable.

Let us now present our results concerning the rate at which the number of links of a given site increases with time. A basic quantity is the average (over 2000 realizations) number  $\langle k_i \rangle$  of links at a given time, more precisely as a function of the *scaled time*  $t/i$  [5]. Our results are illustrated in fig. 5. We numerically find that, for all values of  $(\alpha_A, \alpha_G)$ ,  $\langle k_i \rangle \propto (t/i)^{\beta(\alpha_A)}$ , with  $\beta(0) = 1/2$ .

Let us summarize the present paper. We have introduced a growth model which has both preferential attachment and metrics. Every new site of the cluster has to “decide” to which one of the pre-existing sites it will link itself. In this stochastic choice, there might be competition between the “popular” sites (those that already have many links) and the nearby sites. The nature of this competition differs from that introduced in [7], the degree distribution of which does not always appear to be  $q$ -exponential. In any case, the present competition is more pronounced for increasing value of the exponent  $\alpha_A$ ; it disappears for  $\alpha_A = 0$ . The connectivity distribution and the rate of increase of links with time are substantially influenced by  $\alpha_A$ . The  $\alpha_A = 0$  model belongs to the same universality class to which the well-known Barabasi-Albert model belongs. In addition to these results, we have shown that the connectivity distribution is (numerically) given by the  $q$ -exponential function that emerges naturally in the frame of nonextensive statistical mechanics. This fact supports the conjecture [1] that the typical occupation of the accessible phase space of many nonextensive dynamical systems might be, in relevant stationary states, scale-free. This is in notorious variance with the occupation of standard, extensive, isolated dynamical systems, which tends to be uniform, and whose equilibrium state is that prescribed within Boltzmann-Gibbs statistical mechanics.

\* \* \*

We acknowledge useful remarks from A.-L. BARABASI, S. HAVLIN, S. S. MANNA, L. G. MOYANO and P. C. DA SILVA, as well as partial financial support by Pronex/MCT, Faperj, Capes and CNPq (Brazilian agencies).

## REFERENCES

- [1] GELL-MANN M. and TSALLIS C. (Editors), *Nonextensive Entropy - Interdisciplinary Applications* (Oxford University Press, New York) 2004.
- [2] WATTS D. J. and STROGATZ S. H., *Nature*, **393** (1998) 440.
- [3] ALBERT R. and BARABASI A.-L., *Rev. Mod. Phys.*, **74** (2002) 47.
- [4] GELL-MANN M., *The Quark and the Jaguar - Adventures in the Simple and the Complex* (W. H. Freeman, New York) 1994.
- [5] BARABASI A. L. and ALBERT R., *Science*, **286** (1999) 509.
- [6] ALBERT R., JEONG H. and BARABASI A. L., *Nature*, **401** (1999) 130; *Physica A*, **281** (2000) 69; JEONG H., MASON S. P., OLTAVAI Z. N. and BARABASI A. L., *Nature*, **411** (2001) 41; ROZENFELD A. F., COHEN R., BEN-AVRAHAM D. and HAVLIN S., *Phys. Rev. Lett.*, **89** (2002) 218701.
- [7] YOOK S. H., JEONG H. and BARABASI A.-L., *Proc. Natl Acad. Sci. U.S.A.*, **99** (2002) 13382; MANNA S. S. and SEN PARONGAMA, *Phys. Rev. E*, **66** (2002) 066114.

- [8] TSALLIS C., in *Anomalous Distributions, Nonlinear Dynamics and Nonextensivity*, edited by SWINNEY H. L. and TSALLIS C., *Physica D*, **193** (2004) 3.
- [9] TSALLIS C., *Quim. Nova*, **17** (1994) 468.
- [10] TSALLIS C., *J. Stat. Phys.*, **52** (1988) 479; CURADO E. M. F. and TSALLIS C., *J. Phys. A*, **24** (1991) L69; **24** (1991) 3187(E); **25** (1992) 1019(E).
- [11] ABE S. and OKAMOTO Y. (Editors), *Nonextensive Statistical Mechanics and Its Applications*, *Lect. Notes Phys.* (Springer-Verlag, Heidelberg) 2001; KANIADAKIS G., LISSIA M. and RAPISARDA A. (Editors), *Non Extensive Statistical Mechanics and Physical Applications*, *Physica A*, **305** (2002) 27; HERRMANN H. J., CURADO E. M. F. and BARBOSA M. (Editors), *Physica A*, **344**, Issue 3/4 (Elsevier, Amsterdam) 2004. For full bibliography see <http://tsallis.cat.cbpf.br/biblio.htm>.
- [12] TSALLIS C., MENDES R. S. and PLASTINO A. R., *Physica A*, **261** (1998) 534.
- [13] ALBERT R. and BARABASI A.-L., *Phys. Rev. Lett.*, **85** (2000) 5234.
- [14] DOYE J. P. K., *Phys. Rev. Lett.*, **88** (2002) 238701.
- [15] TSALLIS C., PLASTINO A. R. and ZHENG W.-M., *Chaos, Solitons Fractals*, **8** (1997) 885; COSTA U. M. S., LYRA M. L., PLASTINO A. R. and TSALLIS C., *Phys. Rev. E*, **56** (1997) 245; LYRA M. L. and TSALLIS C., *Phys. Rev. Lett.*, **80** (1998) 53; LATORA V., BARANGER M., RAPISARDA A. and TSALLIS C., *Phys. Lett. A*, **273** (2000) 97; BORGES E. P., TSALLIS C., ANANOS G. F. J. and DE OLIVEIRA P. M. C., *Phys. Rev. Lett.*, **89** (2002) 25; ANANOS G. F. J. and TSALLIS C., *Phys. Rev. Lett.*, **93** (2004) 020601.
- [16] BALDOVIN F. and ROBLEDO A., *Europhys. Lett.*, **60** (2002) 518; *Phys. Rev. E*, **66** (2002) 045104(R); *Phys. Rev. E*, **69** (2004) 045202(R).
- [17] BALDOVIN F., BRIGATTI E. and TSALLIS C., *Phys. Lett. A*, **320** (2004) 254; BALDOVIN F., *Physica A*, **342** (2004) 119; BALDOVIN F., MOYANO L. G., MAJTEY A. P., ROBLEDO A. and TSALLIS C., *Physica A*, **340** (2004) 205.
- [18] BOGHOSIAN B. M., LOVE P. J., COVENEY P. V., KARLIN I. V., SUCCI S. and YEPEZ J., *Phys. Rev. E*, **68** (2003) 025103(R); BOGHOSIAN B. M., LOVE P., YEPEZ J. and COVENEY P. V., in *Anomalous Distributions, Nonlinear Dynamics and Nonextensivity*, edited by SWINNEY H. L. and TSALLIS C., *Physica D*, **193** (2004) 169; TSEKOURAS G. A., PROVATA A. and TSALLIS C., *Phys. Rev. E*, **69** (2004) 016120; ANTENEODO C., *Eur. Phys. J. B*, **42** (2004) 271.