IOP Publishing Physica Scripta
Phys. Scr. **83** (2011) 045007 (8pp) doi:10.1088/0031-8949/83/04/045007

Anomalous diffusion in a symbolic model

H V Ribeiro[1,2], E K Lenzi[1,2], R S Mendes[1,2] and P A Santoro[1]

[1] Departamento de Física, Universidade Estadual de Maringá, Av. Colombo 5790, 87020-900, Maringá, PR, Brazil
[2] National Institute of Science and Technology for Complex Systems, CNPq, Rua Xavier Sigaud 150, 22290-180, Rio de Janeiro, RJ, Brazil

E-mail: hvr@dfi.uem.br, eklenzi@dfi.uem.br, rsmendes@dfi.uem.br and psantoro@dfi.uem.br

Received 26 August 2010
Accepted for publication 21 February 2011
Published 18 March 2011
Online at stacks.iop.org/PhysScr/83/045007

Abstract
In this work, we investigate some statistical properties of symbolic sequences generated by a numerical procedure in which the symbols are repeated following the power-law probability density. In this analysis, we consider that the sum of n symbols represents the position of a particle in erratic movement. This approach reveals a rich diffusive scenario characterized by non-Gaussian distribution and, depending on the power-law exponent or the procedure used to build the walker, we may have superdiffusion, subdiffusion or usual diffusion. Additionally, we use the continuous-time random walk framework to compare the analytic results with the numerical data, thereby finding good agreement. Because of its simplicity and flexibility, this model can be a candidate for describing real systems governed by power-law probability densities.

PACS numbers: 05.40.Fb, 02.50.−r, 05.45.Tp

(Some figures in this article are in colour only in the electronic version.)

1. Introduction

Studies of complex systems are widespread in the physics community [1–4] and a large number of these investigations deal with records of real numbers ordered in time or space. Based on these time series, the aim of these works is to extract some features, patterns or laws that govern a given system. There is an extensive body of literature on statistical tools devoted to analyzing time series. For instance, detrended fluctuation analysis (DFA) [5] can be used to examine the presence of correlations in the data. However, many of these analyses are not focused on the original data but on sub-series such as the absolute value, the return value or the volatility series [6].

In particular, a time series can be converted into a symbolic sequence by using a discrete partition in the data domain and assigning a symbol to each site of partition, a technique that is known as symbol statistics [7]. *A priori*, any dataset can be mapped into a symbolic sequence by using a specific rule (see for instance [8]). A typical analysis performed for this symbolic sequence is to evaluate its block entropy. This approach measures the amount of information contained in the block or the average information necessary to predict subsequent symbols. This analysis was applied to a wide range of topics, including DNA sequences [9]. In this context, Buiatti *et al* [10] introduced a numerical model that generates long-range correlations among the symbols of a symbolic sequence, leading to a slow growth of the usual block entropy.

Motivated by this anomalous behavior in the block entropy, our main goal is to construct a diffusive process based on these sequences. The diffusive processes generated with these sequences are expected to be Markovian or non-Markovian depending on the conditions imposed on them. For Markovian processes or short-term memory systems, the mean square displacement grows linearly in time. On the other hand, non-Markovian processes or long-term memory systems often display deviations of this linear behavior, being better described by a power law on time with the exponent α. This is the fingerprint of anomalous diffusion and depending on the α value we may have superdiffusion ($\alpha > 1$) or subdiffusion ($\alpha < 1$) and for $\alpha = 1$ the usual spreading is recovered. Several physical systems exhibit this power-law pattern. For instance, porous substrate [11], diffusion of high-molecular-weight polyisopropylacrylamide in nanopores [12], highly confined hard disk fluid mixture [13], fluctuating particle fluxes [14],

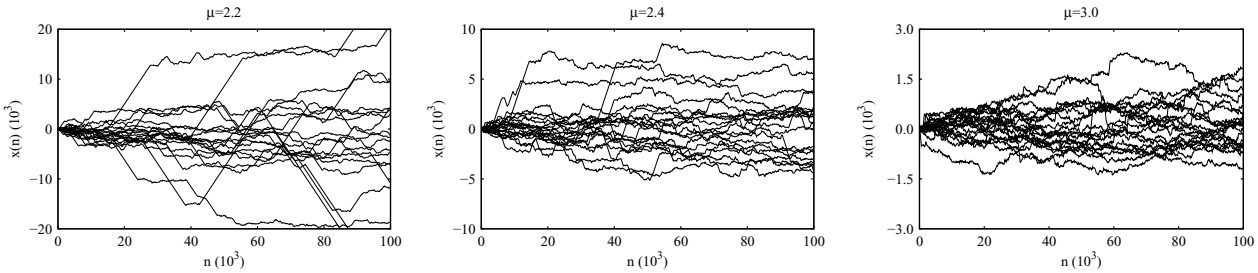

Figure 1. Erratic trajectories $x(n)$ for three values of μ when considering the two-symbol alphabet $\mathcal{A} = \{-1, 1\}$ and $A = 1$. Note that depending on the values of μ the trajectories are very different.

diffusion on fractals [15], ferrofluids [16], nanoporous materials [17] and colloids [18].

In this context, the model proposed by Buiatti *et al* has an essential ingredient leading to anomalous diffusion: the long-term memory present in their symbolic sequences. We will show that a diffusive process based on these sequences leads to a rich diffusive picture, where ballistic diffusion, superdiffusion, subdiffusion or also usual diffusion can emerge, depending on the model parameters and the mode of construction of the process. In addition, we compare our numerical results with analytic models on the basis of continuous-time random walk (CTRW) [19–28].

This paper is organized as follows. In section 2, we present and review some properties of the model. Section 3 is devoted to defining erratic trajectories from the sequences as well as to investigating their diffusive behavior. A comparative analysis of this diffusive aspect with the CTRW viewpoint is performed in section 4. Finally, we end with a summary and some concluding comments.

2. The model

The original model [10] is a numerical experiment that generates equally likely symbols that are repeated among the sequence following a power-law probability density. In order to describe the model, let $\mathcal{A} = \{a_1, \ldots, a_n\}$ be the set of symbols and $Q = \{Q_1, Q_2, \ldots, Q_N\}$ represent a sequence where $Q_i \in \mathcal{A}$. To specify each Q_i, we initially select randomly one of the symbols of the alphabet $\mathcal{A}$ and repeat it N_y times inside the sequence in such a way that $Q_i = Q_{i+1} = \ldots = Q_{i+N_y-1}$. The number N_y is obtained from

$$N_y = [y] + 1 \quad \text{with} \quad y = A\left[\frac{1}{(1-\eta)^{1/(\mu-1)}} - 1\right], \qquad (1)$$

where $A > 0$ and $\mu > 1$ are **real** parameters, η is a random variable uniformly distributed in the interval [0,1] and $[y]$ denotes the integer part of y. By using this procedure, a typical symbolic sequence with $N = 10$ and $\mathcal{A} = \{-1, 1\}$ is

$$Q = \left\{ \underbrace{-1, -1, -1,}_{[y]+1=3} \underbrace{+1, +1, +1, +1, +1}_{[y]+1=5} \underbrace{-1, -1,}_{[y]+1=2} \right\}.$$

Since η is a random number uniformly distributed, y will be a positive random number. Moreover, $p(y)$ can be calculated in a straightforward manner, leading to

$$p(y) = (\mu - 1)\frac{A^{\mu-1}}{(A+y)^{\mu}}. \qquad (2)$$

Therefore, the model basically consists of the repetitions of N_y blocks of symbols with N_y distributed according to a power law of exponent μ in the asymptotic limit. Furthermore, the first and second moments of $p(y)$ are given by

$$\langle y \rangle = \int_0^{\infty} y\, p(y)\, \mathrm{d}y = \frac{A}{(\mu-2)} \quad (\text{for} \quad \mu > 2) \qquad (3)$$

and

$$\langle y^2 \rangle = \int_0^{\infty} y^2\, p(y)\, \mathrm{d}y = \frac{2A^2}{(\mu-2)(\mu-3)} \quad (\text{for} \quad \mu > 3). \qquad (4)$$

Note that when $\mu \leqslant 2$ both moments diverge and when $\mu \leqslant 3$ the second moment diverges, while the first one remains finite. Thus, when μ is close to 2, N_y can be very large, filling a significant part of the sequence Q with the same symbol. On the other hand, very large values of N_y become rare for $\mu > 3$, which makes the sequence highly alternating.

It is well established that this method of building symbolic sequences generates long-range correlations between elements of the sequence characterized by a power-law correlation function (see the analytical development of Buiatti *et al* [10]). It was also studied whether correlations lead to a nonlinear growth of the usual block entropy, i.e. the usual entropy is not extensive for $\mu < 3$. In such a context, these sequences were investigated in the framework of the so-called non-extensive Tsallis statistical mechanics. In particular, it was shown that the Tsallis block entropy S_q can recover extensivity for a specific choice of the entropic index q [10, 29].

3. Diffusive process

As we mentioned in the introduction, long-term correlations or non-Markovian processes frequently present anomalous properties when investigated in the context of diffusion. In this direction, to construct erratic trajectories from these sequences may yield a rich diffusive scenario based on a simple model. This point has been noted by Buiatti *et al* and was also briefly discussed by us in [29].

A simple and direct way to obtain the trajectories is to consider that each symbol in a sequence represents the length of the jump of a particle in erratic movement. The position of

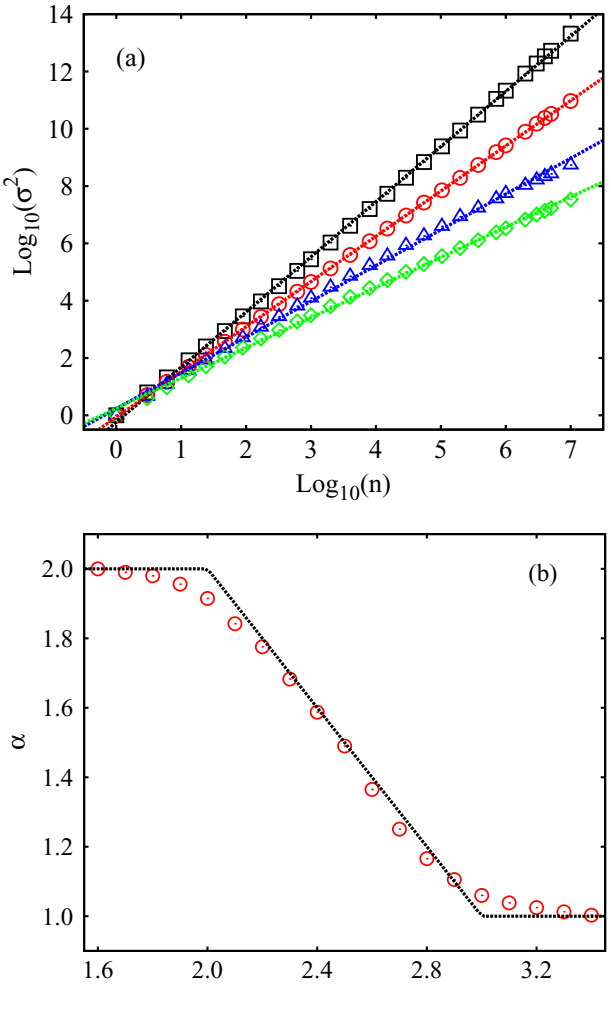

Figure 2. Results concerning the alphabet $\mathcal{A} = \{-1, 1\}$ for $A = 1$. (a) Logarithm of the variance $\sigma^2(n)$ versus logarithm of n for $\mu = 1.8$ (squares), $\mu = 2.4$ (circles), $\mu = 2.8$ (triangles) and $\mu = 3.4$ (diamonds). In this figure, the slopes of the curves are numerically equal to the exponents α and the straight lines are linear fit to the data used to obtain the values of α. (b) The dependence of the exponent α on μ. From this figure, we have basically three diffusion regimes: a ballistic ($\mu \lesssim 2$), a superdiffusive ($2 \lesssim \mu \lesssim 3$) and the usual Brownian motion ($\mu \gtrsim 3$) regime. The straight lines are the predictions of the CTRW model related to equation (9).

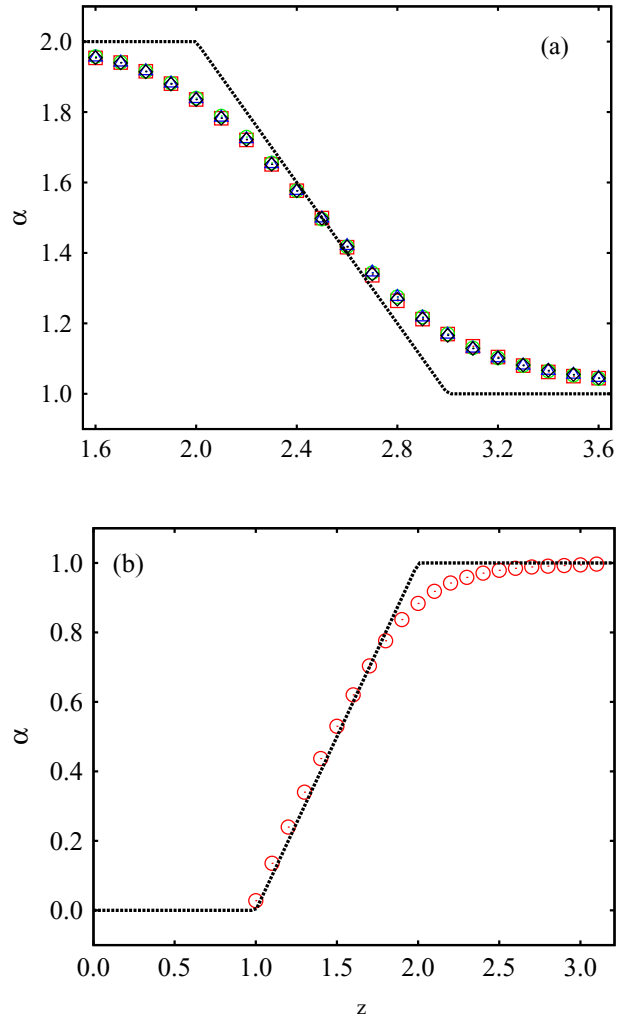

Figure 3. (a) The dependence of the exponent α on μ when considering the alphabet $\mathcal{A} = \{-1, 0, 1\}$ and varying the probability of the zero symbol within the sequence. For the squares, the symbols are equiprobable and for the circles the zero symbol is 20 times more likely to occur ($\mathcal{A} = \{-1, 0^{20}, 1\}$). The triangles and the diamonds are the results concerning the larger alphabets $\mathcal{A} = \{-2, -1, 0, 1, 2\}$ and $\mathcal{A} = \{-20, \ldots, -2, -1, 0, 1, 2, \ldots, 20\}$, respectively. (b) The dependence of the exponent α on μ_z when considering the equiprobable alphabet $\mathcal{A} = \{-1, 0, 1\}$ and $\mu_j = 6$. Note the presence of a subdiffusive regime for $1 \lesssim \mu_z \lesssim 2$. All results were obtained by using sequences of length 10^7 averaged over 5×10^5 realizations with $A = 1$. The straight lines are the predictions of the CTRW model related to equations (9) and (10), respectively.

the symbol plays the role of time. In this manner, the variable

$$x(n) = \sum_{i=1}^{n} Q_i \tag{5}$$

represents the position of the particle after a time n, which is an integer because of the construction.

Let us start our investigation by considering the simplest symmetric case, i.e. the two symbol alphabet $\mathcal{A} = \{-1, 1\}$. Thus, the particle is equally likely to jump to the right or to the left according to whether $Q_i = 1$ or $Q_i = -1$ and the variable $x(n)$ represents a random walk-like process where $x(1) = \pm 1$, $x(2) = 0, \pm 2$, $x(3) = \pm 3, \pm 1$ and so on, with n playing the role of *time*. Figure 1 illustrates $x(n)$ for three values of μ with $A = 1$. Note that the trajectories are remarkably different depending on the values of μ. For small values of μ ($\mu < 3$), we can see that the trajectories are governed by two mechanisms: spatial localization and large jumps. When $\mu \leqslant 2$, the jump is larger, reflecting the fact that all moments of $p(y)$ diverge for $\mu \leqslant 2$. On the other hand, the second moment of $p(y)$ is finite for $\mu > 3$ and the trajectories are very similar to usual random walks.

As pointed out in the introduction, when dealing with the diffusive process, it is very common to investigate how the particles are spreading by evaluating the variance $\sigma^2(n) = \langle (x(n) - \langle x(n) \rangle)^2 \rangle$, where the angular brackets denote an ensemble average. The usual Brownian motion [30, 31] is characterized by $\sigma^2(n) \sim n$ and by a Gaussian propagator $p(x, n) \sim n^{-1/2} \exp(-x^2/2n)$, which is a direct consequence of the central limit theorem and the Markovian nature of

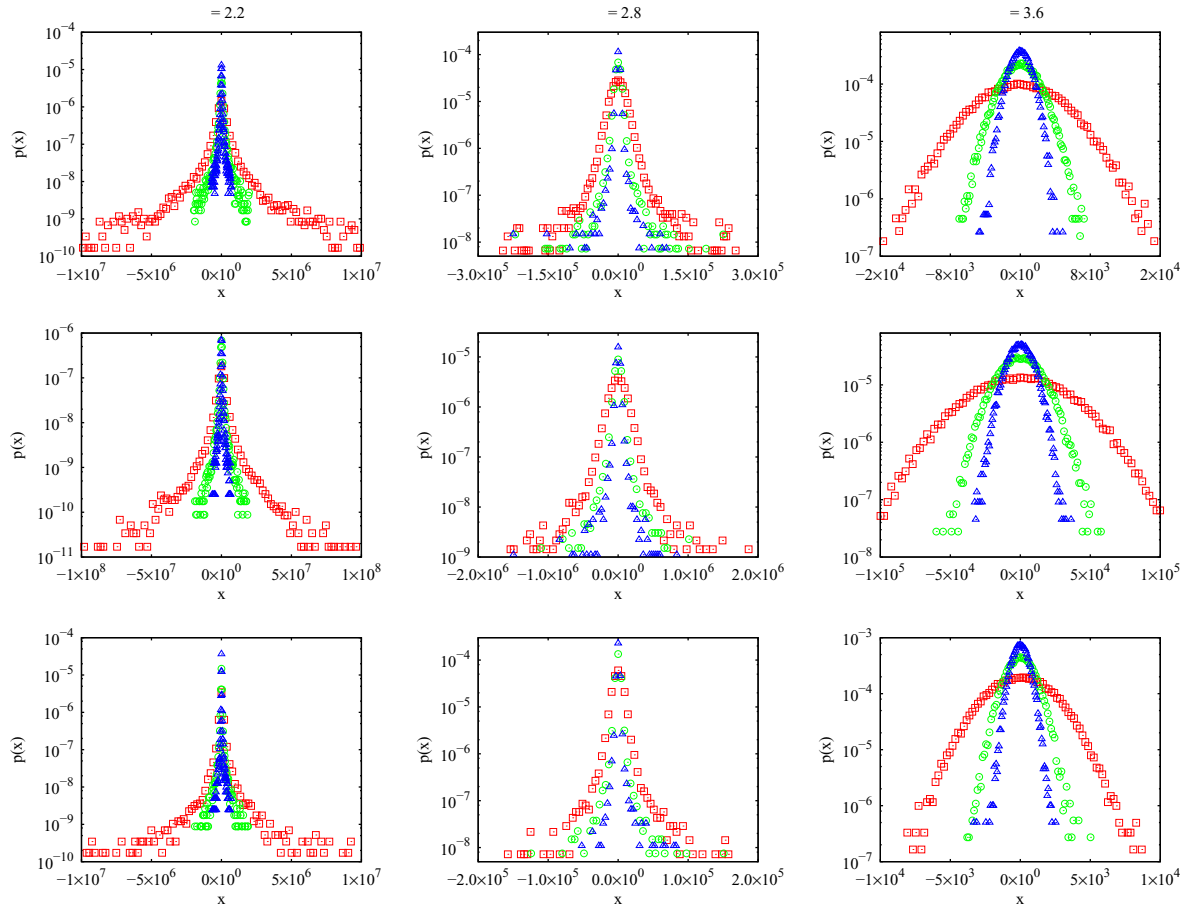

Figure 4. Probability density functions (PDF) of $x(n)$ for the values of μ (indicated in the figures) for three values of n: 10^7 (squares), 2×10^6 (circles) and 714×10^3 (triangles) when considering the equiprobable alphabets $\mathcal{A} = \{-1, 1\}$ (upper panels), $\mathcal{A} = \{-11, \ldots, 0, \ldots 11\}$ (middle panels) and the alphabet $\mathcal{A} = \{-1, 0^{10}, 1\}$ where the zero symbol is ten times more probable than the others (lower panels). The histograms were obtained by using sequences of length 10^7 with $A = 1$ and 5×10^5 realizations of the numerical experiment.

the underlying stochastic process. On the other hand, the anomalous diffusion behavior is usually distinguished by the value of the exponent α [32] in

$$\sigma^2(n) \propto n^{\alpha}. \tag{6}$$

We have subdiffusion when $0 < \alpha < 1$ and superdiffusion when $\alpha > 1$. The crossover between subdiffusion and superdiffusion corresponds to the usual Brownian motion and the case $\alpha = 2$ is the ballistic regime.

In this direction, we evaluate the variance for several values of μ over an ensemble average of 5×10^5 realizations as shown in figure 2(a). In a log–log plot, the slope of the curve $\sigma^2(n)$ versus n is numerically equal to the exponent α, which is visibly changing with the parameter μ. In figure 2(b), we quantify this dependence by plotting α versus μ. From this figure, we have basically three diffusion regimes depending on the existence of the first $\langle y \rangle$ and the second $\langle y^2 \rangle$ moments of $p(y)$: (i) a ballistic one for $\mu < 2$ ($\langle y \rangle \to \infty$ and $\langle y^2 \rangle \to \infty$), (ii) a superdiffusive one for $2 < \mu < 3$ ($\langle y \rangle$ finite and $\langle y^2 \rangle \to \infty$) and (iii) the usual Brownian motion for $\mu > 3$ ($\langle y \rangle$ and $\langle y^2 \rangle$ finite).

Next, we investigate the role of the size of the symbol space by considering that more symbols are present in the alphabet $\mathcal{A}$. We consider first the presence of zeros, i.e. $\mathcal{A} = \{-1, 0, 1\}$ where each symbol is equiprobable within the sequence. The zero symbol allows the particles to stay motionless for a certain time, which could be related to subdiffusion. However, as we show in figure 3(a) the presence of zeros in the sequence does not significantly change the profile of the relation μ versus α. Moreover, even if the zero symbol is becoming more probable within the sequence, i.e. $\mathcal{A} = \{-1, 0^T, 1\}$ where T is the number of zero symbols in the alphabet, this result remains valid, as we also show in figure 3(a). Secondly, we study larger alphabets from $\mathcal{A} = \{-2, -1, 0, 1, 2\}$ to $\mathcal{A} = \{-20, \ldots, -2, -1, 0, 1, 2, \ldots, 20\}$ and the results are shown in figure 3(a). We found that the relation μ versus α does not depend on the size of the symbol space. This relation is also robust for variations of the parameter A and for non-symmetric alphabets. In particular, the parameter A only produces a multiplicative effect in equation (6) and a non-symmetric alphabet produces a drift that does not affect the spreading of the system.

Until now we were not able to generate subdiffusion, even adding the zero symbol more likely to occur. This result suggests that only superdiffusion can emerge when considering the same value of μ for the symbols that lead to jumps and for the zero symbol that leads to the absence of motion. The reasons for this behavior are related to the fact

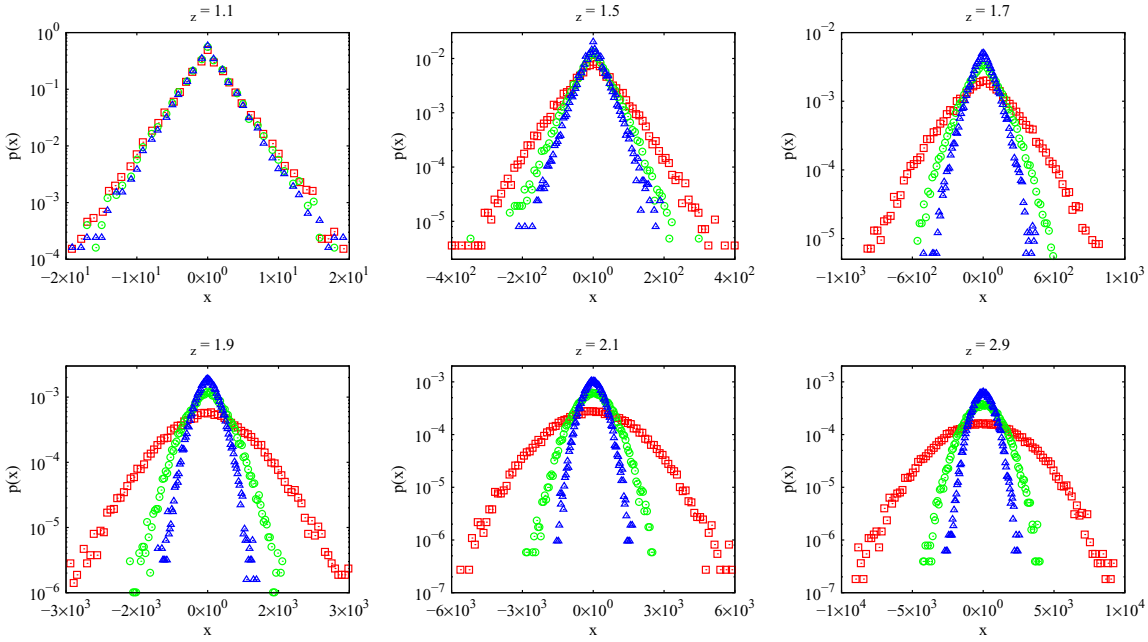

Figure 5. PDF of $x(n)$ for the values of μ_z (indicated in the figures) for three values of n: 10^7 (squares), 2×10^6 (circles) and 714×10^3 (triangles). The histograms were obtained by using sequences of length 10^7 with $A = 1$, $\mu_j = 6$ and 5×10^5 realizations of the numerical experiment with the alphabet $\mathcal{A} = \{-1, 0, 1\}$.

that even if the zero symbol is much more likely than other symbols, the number of repetitions N_y is independent of the symbol. Thus, the particles can remain at rest for a long time but the flights can be equally long, since for $\mu < 3$ there is no characteristic scale for N_y.

In this direction, let us consider a sequence where the jumping symbols are related to the μ_j value ($\mu_j > 3$) and the zero symbol is related to other μ_z values. In this manner, flights have a characteristic scale, whereas the rest periods may or may not have this characteristic scale (depending on the μ_z value). The results concerning this scenario are shown in figure 3(b), where we show the dependence of α on μ_z for a fixed value of $\mu_j = 6$ (different values of $\mu_j > 3$ do not affect this relation). From this figure, we can identify three diffusive regimes: no diffusion for $\mu_z \approx 1$ where the sequence is practically filled by zeros, subdiffusion for $1 \lesssim \mu_z \lesssim 2$ and usual diffusion for $\mu_z \gtrsim 2$.

We also evaluated the probability density functions (PDF) of $x(n)$ to investigate the shape of the distribution $p(x, n)$ for different values of μ as well as for the different constructions of the erratic trajectories. Figure 4 shows these distributions for the equiprobable alphabets $\mathcal{A} = \{-1, 1\}$ (upper panels), $\mathcal{A} = \{-11, \ldots, 0, \ldots 11\}$ (middle panels) and the alphabet $\mathcal{A} = \{-1, 0^{10}, 1\}$ with the zero symbol being ten times more probable than the others (lower panels). We can see that the distributions are characterized by non-Gaussian profiles with heavy tails when $\mu \lesssim 3$, recovering the Gaussian propagator when $\mu \gtrsim 3$. Further, a visual inspection suggests that the different constructions of the erratic trajectories only change the scale of these plots.

The situation is remarkably different when considering one value of μ for the jumping symbols and another for the zero symbol with the alphabet $\mathcal{A} = \{-1, 0, 1\}$. Figure 5 shows the distributions for this case. Note that the shape of distributions goes from a Laplace ($p(x) \sim \exp(-|x|)$) to Gaussian distribution, depending on the μ_z value.

4. Continuous-time random walk models

So far, we have empirically described the diffusive behavior of the symbolic sequences proposed by Buiatti *et al*. Let us compare these empirical findings with some analytical models based on CTRW.

In the CTRW model of Montroll [19] (see also [20]), the random walk process is fully specified by the function $\psi(x, t)$, the probability density to move a distance x in time t. We can distinguish three different ways to make the movement: the particle waits until it moves instantaneously to a new position (jump model) or the particle moves at constant velocity to a new position and chooses randomly a new direction (velocity model) or the particle moves at constant velocity between turning points that are chosen randomly [21]. There are two fundamental approaches to the CTRW: (i) the decoupled and (ii) the coupled formalisms. In (i), the function $\psi(x, t)$ is supposed to factor in the form $\psi(x, t) = w(t)\lambda(x)$, i.e. the jumping and the waiting time are independent random variables. For (ii), both processes are coupled, a jump of a certain length may involve a time cost or vice versa. This coupled form commonly leads to more cumbersome calculations.

Here, we note that because of the erratic trajectories construction, every continuous jump (without changing the sequence symbol) with length N_y occurs at constant velocity and costs the same N_y units of time to perform. This fact leads us to the velocity and coupled model when considering the equiprobable alphabet $\mathcal{A} = \{-1, 1\}$. When adding the zero symbol the resting times are decoupled from the jumps, but the jumping times stay coupled. In addition, remember that

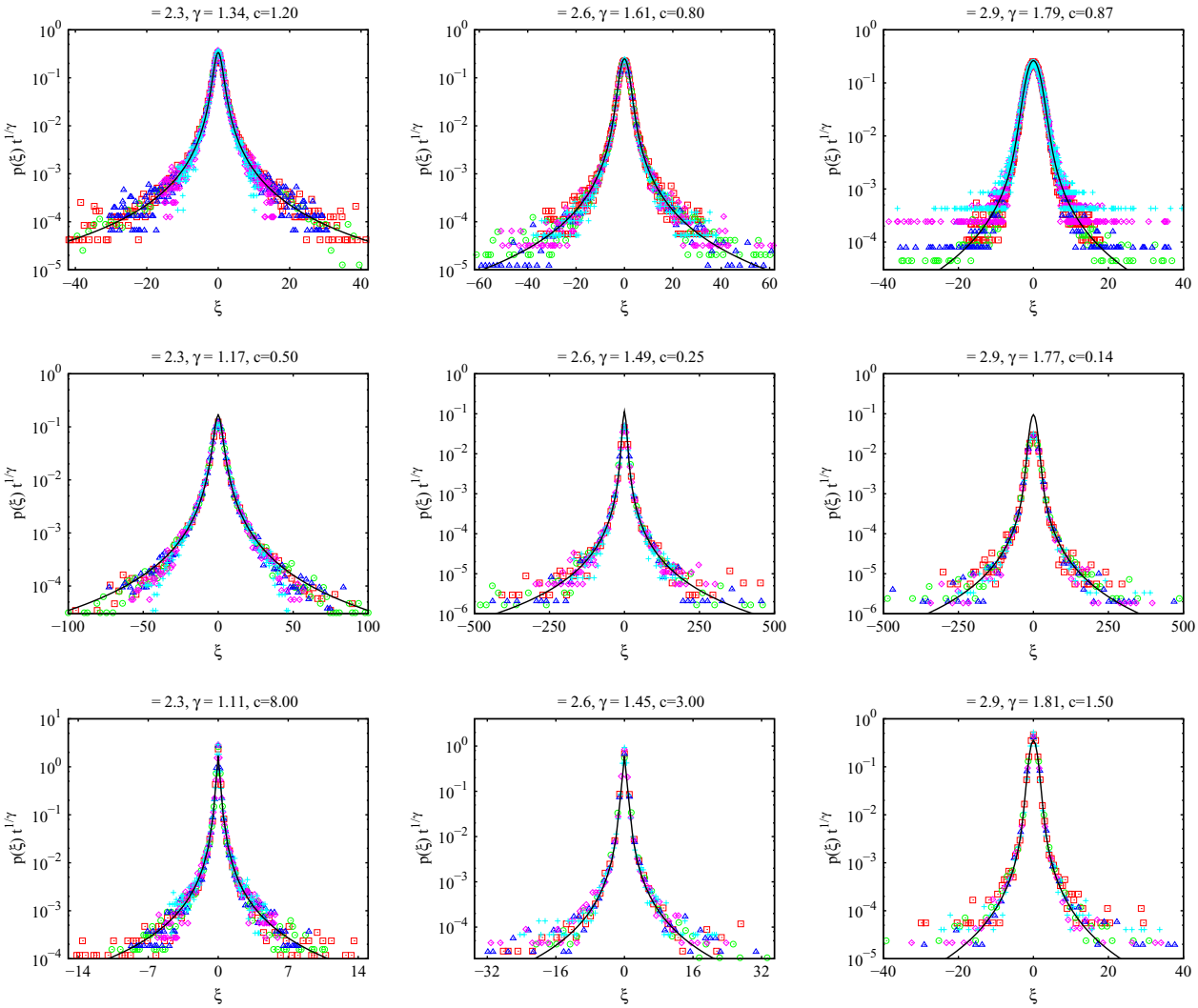

Figure 6. PDF of the scaling variable $\xi = c\,x/t^{\gamma}$ for some values of μ (indicated in the figure) for five values of n: 10^7 (squares), 2×10^6 (circles), 714×10^3 (triangles), 55×10^3 (diamonds) and 15×10^3 (crosses). The upper panels show the results for the equiprobable alphabet $\mathcal{A} = \{-1, 1\}$, the middle panels for $\mathcal{A} = \{-11, \ldots, 0, \ldots 11\}$ and the lower panels for $\mathcal{A} = \{-1, 0^{10}, 1\}$ with the zero symbol being ten times more probable than the others. The continuous lines are the predictions of the CTRW model, equation (11). The values of c and γ are indicated in the figure and the numerical data were obtained from 5×10^5 realizations of equation (5) with $A = 1$.

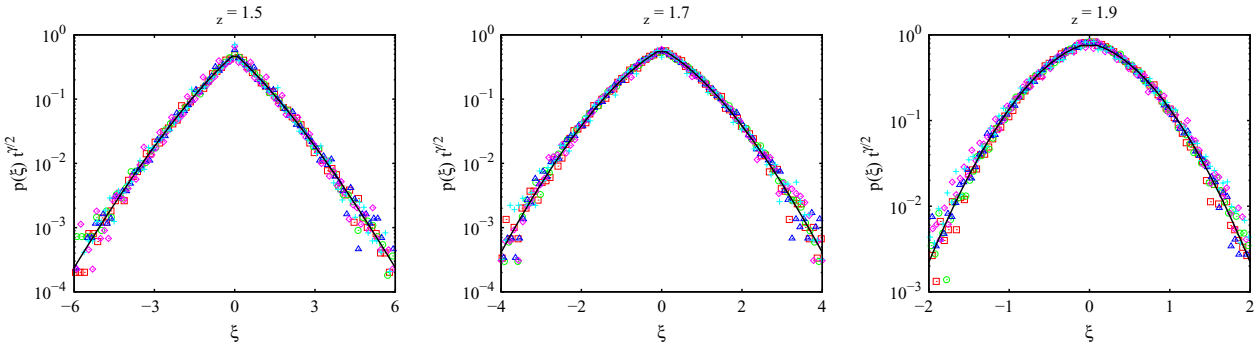

Figure 7. PDF of the scaling variable $\xi = cx/t^{\gamma/2}$ for some values of μ_z (indicated in the figure) for five values of n: 10^7 (squares), 2×10^6 (circles), 714×10^3 (triangles), 55×10^3 (diamonds) and 15×10^3 (crosses) when considering the equiprobable alphabet $\mathcal{A} = \{-1, 0, 1\}$ and $\mu_j = 6$. The continuous line is the prediction of the CTRW model, equation (12) with $\gamma = 0.47$ ($c = 0.65$) for $\mu_z = 1.5$, $\gamma = 0.64$ ($c = 0.37$) for $\mu_z = 1.7$ and $\gamma = 0.94$ ($c = 0.15$) for $\mu_z = 1.9$. The numerical data were obtained from 5×10^5 realizations of the numerical experiment with $A = 1$.

our formal time n is a discrete variable. Thus, comparisons with this formalism should be viewed as semi-quantitative. In this context, it is interesting to note that the work of Gorenflo *et al* [23, 24] extends Montroll's theory to the discrete domain considering the decoupled version of the CTRW, in contrast to the first approach used here.

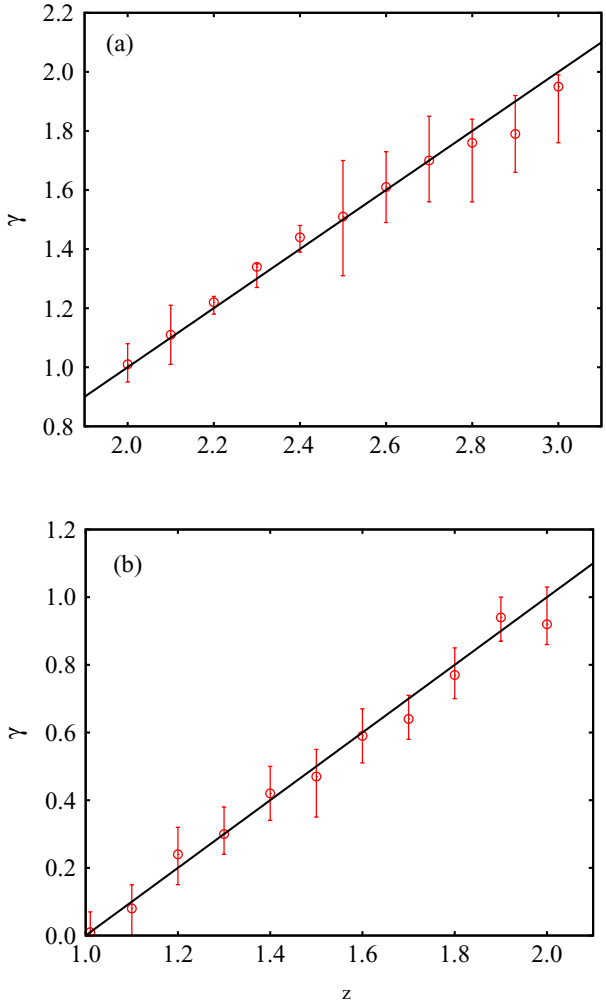

Figure 8. (a) The averaged value of γ versus μ obtained via the procedure described in the text for the equiprobable alphabet $\mathcal{A}=\{-1,1\}$. (b) The averaged value of γ versus μ_z when considering the equiprobable alphabet $\mathcal{A}=\{-1,0,1\}$ and $\mu_j=6$. Note that the relation $\gamma=\mu-1$ or $\gamma=\mu_z-1$ obtained by comparing $w(t)$ and $p(y)$ holds in general. The error bars are calculated via the bootstrap resampling method [35].

We start by considering the coupled velocity model of Zumofen and Klafter [22] for a comparison with our sequences. For this case

$$\psi(x,t)=\tfrac{1}{2}\delta(|x|-t)\,w(t), \tag{7}$$

where $w(t)\sim t^{-\gamma-1}$. Note that in this CTRW model long jumps are always penalized by long waiting times due to the presence of the δ function. Furthermore, because of the asymptotic behavior of $p(y)\sim y^{-\mu}$, γ should be related to μ via $\gamma=\mu-1$. Now, following the approach of Zumofen and Klafter, we can write the form of the distribution $p(x,t)$ in the Fourier–Laplace space as

$$p(k,u)=\frac{\Psi(k,u)}{1-\psi(k,u)}, \tag{8}$$

where

$$\Psi(x,t)=\frac{1}{2}\delta(|x|-t)\int_t^\infty w(t')\,\mathrm{d}t'$$

is the probability density to move a distance x in time t in a single event and not necessarily stop at x. By using $p(k,u)$, we can evaluate the variance

$$\sigma^2(t)=\mathcal{L}^{-1}\left[-\frac{\partial^2}{\partial k^2}p(k,u)\Big|_{k=0}\right]\sim\begin{cases}t^2, & 1<\mu<2,\\ t^{4-\mu}, & 2<\mu<3,\\ t, & \mu>3.\end{cases} \tag{9}$$

Figure 2(b) shows the comparison with numerical data for the alphabet $\mathcal{A}=\{-1,1\}$ and figure 3(a) makes this for the alphabets $\mathcal{A}=\{-1,0,1\}$ (with the zero symbol more probable) and also the larger alphabets $\mathcal{A}=\{-2,-1,0,1,2\}$ and $\mathcal{A}=\{-20,\ldots,-2,-1,0,1,2,\ldots,20\}$. Naturally, the agreement is better for the first case than for the others, since it fulfills the requirements of the model. In general, we can see that the presence of the zero symbol makes the convergence of α to the limiting regimes slower ($\alpha=2$ and $\alpha=1$).

We consider another CTRW model trying to reproduce the subdiffusive regime. Specifically, we employ the decoupled version proposed by Montroll [33], where $\lambda(x)\sim\exp(-x^2)$ and $w(t)\sim t^{-\gamma-1}$ ($0<\gamma<1$). Following Montroll [33], also Metzler and Klafter [20], we obtain

$$\sigma^2(t)\sim t^{\mu_z-1}, \tag{10}$$

where again we have used the relation $\gamma=\mu_z-1$. Figure 3(b) confronts the numerical data with this expression for which we can see a good agreement.

Additionally, we may also obtain the propagator from a small (k,u) expansion for both previous cases. For the first one $p(k,u)\sim 1/(u+c|k|^\mu)$, which for $2<\mu<3$ yields

$$p(x,t)\sim\begin{cases}t^{-1/\gamma}L_\gamma(\xi), & |x|<t,\\ 0, & |x|>t,\end{cases} \tag{11}$$

where $L_\gamma(\xi)$ is the Lévy stable distribution and $\xi=c\,x/t^\gamma$ is the scaling variable. For the second one, $p(k,u)\sim 1/(u+c\,k^2\,u^\gamma)$, leading to

$$p(x,t)\sim t^{-\gamma/2}H_{1,2}^{2,0}\left[\xi^2\Big|_{(0,1),(1/2,1)}^{(1-\gamma/2,\gamma)}\right], \tag{12}$$

where $H_{1,2}^{2,0}[\xi^2|_{(0,1),(1/2,1)}^{(1-\gamma/2,\gamma)}]$ is the Fox H function [34] and $\xi=cx/t^{\gamma/2}$ is the scaling variable.

Figure 6 shows the comparison for the first case and figure 7 for the second one. In both cases, we can see a good-quality data collapse when the scaling is performed. Moreover, these figures show that we have found good agreement between the numerical data and the CTRW models. The fitting parameter γ for each case was obtained by minimizing the difference between the numerical data and the analytic expressions using the nonlinear least squares method. In all these figures, we have employed the averaged value of γ obtained by applying the method for 17 values of n chosen logarithmically spaced from 10^3 to 10^7. In addition, figures 8(a) and (b) show the dependence of the averaged value γ on μ for both cases, showing that the relation $\gamma=\mu-1$ or $\gamma=\mu_z-1$ is consistent with the numerical data.

5. Summary

We verified that the symbolic model presented by Buiatti *et al* [10] gives rise to a rich diffusive scenario. Depending on the parameter μ (or μ_z), different anomalous diffusive regimes can emerge. Specifically, we have found subdiffusive, superdiffusive, ballistic and usual regimes, depending on the model parameters or the trajectories construction. We also investigated the probability distributions of these processes where non-Gaussians were observed. Our findings support the existence of self-similarity in the data, due to the good quality of the data collapse when the scaling is performed. In addition, the numerical data were compared with predictions of the CTRW framework, thereby finding good agreement. We believe that our empirical findings may help modeling systems for which power laws are present as well as to motivate other random walk constructions based on symbolic sequences.

Acknowledgments

We thank CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho—São Paulo) for computational support and CAPES/CNPq (Brazilian agencies) for financial support. HVR wishes acknowledge Eduardo G Altmann for helpful discussions at LAWNP'09.

References

[1] Auyang S Y 1998 *Foundations of Complex Systems* (Cambridge: Cambridge University Press)
[2] Jensen H J 1998 *Self-Organized Criticality* (Cambridge: Cambridge University Press)
[3] Albert R and Barabási A-L 2002 *Rev. Mod. Phys.* **74** 47
[4] Boccara N 2004 *Modeling Complex Systems* (New York: Springer)
[5] Peng C-K *et al* 1992 *Nature* **356** 168
[6] Mantegna R N and Stanley H E 1999 *An Introduction to Econophysics* (Cambridge: Cambridge University Press)
[7] Tang X Z, Tracy E R and Brow R 1997 *Physica* D **102** 253
[8] Steuer R, Molgedey L, Ebeling W and Jiénez-Montaño M A 2001 *Eur. Phys. J.* B **19** 265
[9] Voss R F 1992 *Phys. Rev. Lett.* **68** 3805
[10] Buiatti M, Grigolini P and Palatella L 1999 *Physica* A **268** 214
[11] Brault P *et al* 2009 *Phys. Rev. Lett.* **102** 045901
[12] Caspi Y, Zbaida D, Cohen H and Elbaum M 2009 *Macromolecules* **42** 760
[13] Ball C D, MacWilliam N D, Percus J K and Bowles R K 2009 *J. Chem. Phys.* **130** 054504
[14] Saenko V V 2009 *Plasma Phys. Rep.* **35** 1
[15] Metzler R, Glockle G and Nonnenmacher T F 1994 *Physica* A **211** 13
[16] Mertelj A, Cmok L and Copic M 2009 *Phys. Rev.* E **79** 041402
[17] Roque-Malherbe R M A 2007 *Adsorption and Diffusion in Nanoporous Materials* (New York: Taylor and Francis)
[18] Golestanian R 2009 *Phys. Rev. Lett.* **102** 188305
[19] Montroll E W and Weiss G H 1965 *J. Math. Phys.* **6** 167
[20] Metzler R and Klafter J 2000 *Phys. Rep.* **339** 1
[21] Zumofen G and Klafter J 1993 *Phys. Rev.* E **47** 851
[22] Zumofen G and Klafter J 1994 *Europhys. Lett.* **25** 565
[23] Gorenflo R and Vivoli A 2003 *Signal Process.* **83** 2411
[24] Gorenflo R, Vivoli A and Mainardi F 2004 *Nonlinear Dyn.* **38** 101
[25] Barkai E 2002 *Chem. Phys.* **284** 13
[26] Meerschaert M M, Benson D A, Scheffler H-P and Becker-Kern P 2002 *Phys. Rev.* E **66** 060102(R)
[27] Zaburdaev V Y and Chukbar K V 2002 *J. Exp. Theor. Phys.* **94** 252
[28] Sokolov I M and Metzler R 2003 *Phys. Rev.* E **67** 010101(R)
[29] Ribeiro H V, Lenzi E K, Mendes R S, Mendes G A and da Silva L R 2009 *Braz. J. Phys.* **39** 444
[30] Gardiner C W 2004 *Handbook of Stochastic Methods* (Berlin: Springer)
[31] Vlahos L, Isliker H, Kominis Y and Hizanidis K 2008 arXiv:0805.0419v1
[32] Dybiec B and Gudowska-Nowak E 2009 *Phys. Rev.* E **80** 061122
[33] Montrol E W 1956 *J. SIAM* **4** 241
[34] Mathai A M and Saxena R K 1978 *The H-Function with Application in Statistics and Other Disciplines* (New York: Wiley)
[35] Efron B and Tibshirani R 1993 *An Introduction to the Bootstrap* (London: Chapman and Hall)