Restricted random walk model as a new testing ground for the applicability of q-statistics

Ugur Tirnakli[1], Henrik Jeldtoft Jensen[2] and Constantino Tsallis[3]

[1] *Department of Physics, Faculty of Science, Ege University, 35100 Izmir, Turkey*
[2] *Complexity & Networks Group and Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK*
[3] *Centro Brasileiro de Pesquisas Físicas and National Institute of Science and Technology for Complex Systems, Rua Dr. Xavier Sigaud 150, 22290-180 Rio de Janeiro, Brazil and*
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Abstract. - We present exact results obtained from Master Equations for the probability function $P(y,T)$ of sums $y = \sum_{t=1}^{T} x_t$ of the positions x_t of a discrete random walker restricted to the set of integers between $-L$ and L. We study the asymptotic properties for large values of L and T. For a set of position dependent transition probabilities the functional form of $P(y,T)$ is with very high precision represented by q-Gaussians when T assumes a certain value $T^* \propto L^2$. The domain of y values for which the q-Gaussian apply diverges with L. The fit to a q-Gaussian remains of very high quality even when the exponent a of the transition probability $g(x) = |x/L|^a + p$ with $0 < p << 1$ is different from 1, although weak, but essential, deviation from the q-Gaussian does occur for $a \neq 1$. To assess the role of correlations we compare the T dependence of $P(y,T)$ for the restricted random walker case with the equivalent dependence for a sum y of uncorrelated variables x each distributed according to $1/g(x)$.

Introduction. – The central limit theorem states that appropriately scaled sums of independent random variables will be distributed according to a Gaussian [1, 2]. The random walker is the prototype example of a stochastic Gaussian process [3, 4]. The standard random walker is characterized by transition constant probabilities, which are independent of position and time. Here we point out that for a certain class of position dependent transition probabilities correlations arise, which leads to deviations away from Gaussian behavior. There exist already a large amount of evidence, which points to q-Gaussians as the relevant high quality approximates for the functional form for the distribution function in a range of cases where correlations play an essential rôle [5]. The evidence for the relevance of the q-Gaussian is however often derived from numerical experiments in which fluctuations limits the accuracy and therefore the precision of the fit to the q-Gaussian form (analytical exceptions to this frequent difficulty can be found in [6]). Moreover we believe the random walk example we discuss here to be highly generic. It is related to e.g. particles moving in a confining potential or to more branching processes subject to resource limitations.

Here we present an investigation of the sum $y = \sum_{t=1}^{T} x_t$ of positions x_t passed through

by a Restricted Random Walker (RRW). The underlying stochastic process is sufficiently simple to allow exact numerical solution of the Master Equation (ME) for the probability distribution $P(y,T)$. This ensures a very high precision fit to the q-Gaussian form and thereby a very accurate determination of the relevant parameters. We find that a broad range of *transition* probabilities for the random walker leads to q-Gaussians with q parameters depending on transition probabilities. Since the ME can be easily handled in exact numerically form, the RRW model is an excellent laboratory for understanding the conditions under which sums of correlated random variables are distributed as q-Gaussians:

$$P(y) = \begin{cases} P(0)\left[1 - \beta(1-q)y^2\right]^{\frac{1}{1-q}} & \text{for } \beta(1-q)y^2 < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $q < 3$ and $\beta > 0$ are parameters (for $q \geq 3$ normalizability is lost). As $q \to 1$ the function $P(y)$ approaches the Gaussian.

Restricted Random Walk Model. – We consider a one dimensional symmetric random walker confined to the integers between $-L$ and L. The motion of the walker is controled by the following time evolution

$$x_{t+1} = \begin{cases} x_t + 1 & \text{with probability } g(x)/2 \\ x_t - 1 & \text{with probability } g(x)/2 \\ x_t & \text{with probability } 1 - g(x). \end{cases} \tag{2}$$

We concentrate on the following form

$$g(x) = \min\left\{\left|\frac{x}{L}\right|^a + p, 1\right\}, \tag{3}$$

with reflective boundary conditions: If $x_{t+1} > L$ $(< -L)$ we let $x_{t+1} \mapsto x_{t+1} - 1$ $(+1)$. We find numerically that the first return time (defined as the time elapsed until the walker, who leaves its $x = 0$ position, returns to the zero position again, and note we do not include walkers who remain at $x = 0$ for all times) distribution for these RRW behaves asymptotically like $P(T) \sim T^{-\tau}$, with $\tau = 2$, *i.e.* *different* from the exponent $\tau = 3/2$ for ordinary RW. We study the sum $y = \sum_{t=1}^{T} x_t$ in the limit $p \to 0$ for values of the exponent $a = 0.75$, 1 and 1.25. For $p \to 1$ and $L \to \infty$ the process reduces to the ordinary random walk.

The highly restrictive nature of the RRW is clearly seen from the 6 trajectories shown in Fig.1 in the case $L = 120$ (with $a = 1$ and $p = 5 \cdot 10^{-6}$). For comparison we present the trajectories of an ordinary random walk on $x_{t+1} = x_t \pm 1$ with probability $1/2$ and x_t confined to $-L, -L+1, .., L-1, L$. The figure shows that the vanishing transition probability $g(x)$ near $x = 0$ makes the RRW non-ergodic leaving most of the phase space empty. It is straight forward to derive a Master Equation for the distribution $P(x,t)$

$$P_X(x,t+1) = P_X(x,t) + \frac{1}{2}g(x-1)P_X(x-1,t) + \frac{1}{2}g(x+1)P_X(x+1,t) - g(x)P_X(x,t) \tag{4}$$

subject to the appropriate boundary conditions at $|X| = L$. The insert in Fig. 1 exhibits a solution of the equation. We will discuss $P(x,t)$ in a more extended publication, here we now turn our attention to another distribution.

Let $P(y,x,T)$ denote the probability that $\sum_{t=1}^{T} x_t = y$ and $x_T = x$. The time evolution of this simultaneous probability is controlled by the following ME

$$\begin{aligned} P(y,x;t+1) = P(y,x;t) + \sum_{\Delta \in \{-1,0,1\}} \\ [W(y,x;y-(x-\Delta),x-\Delta)P(y-(x-\Delta),x-\Delta;t)) \\ -W(y+x+\Delta,x+\Delta;y,x)P(y,x;t)]. \end{aligned} \tag{5}$$
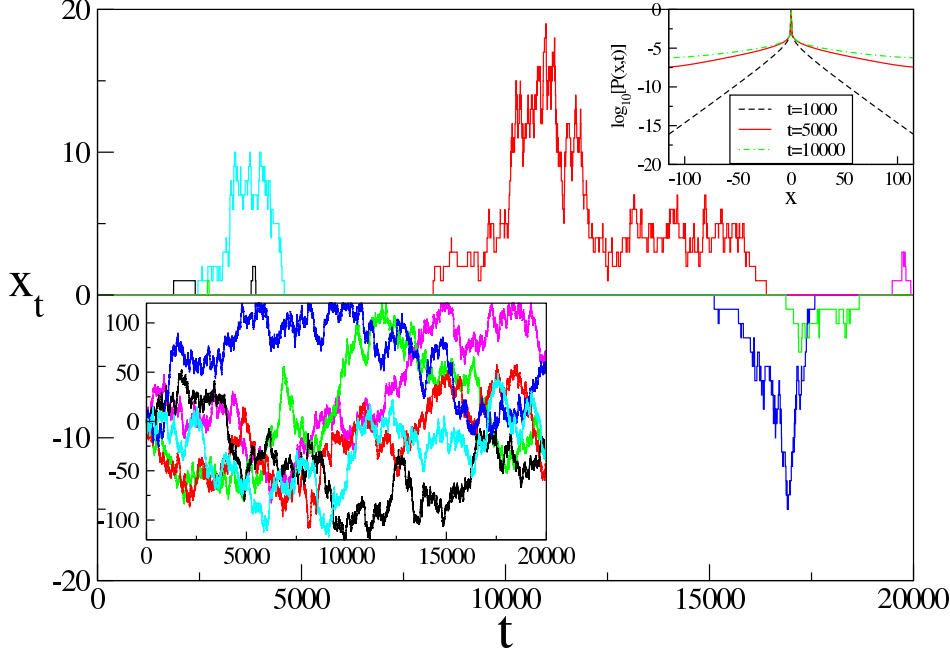
Fig. 1: Six representative trajectories for $L = 120$. Main panel: The restricted RW model ($p = 5 \cdot 10^{-6}$). Lower Inset: The standard RW model ($p = 1$). Non-ergodic behavior of the restricted RW model can easily be seen. Upper Inset: The time evolution of $P(x, t)$ at three different t values ($t = 1000$, 5000 and 10000 from bottom to top).

The transition probabilities W only depend on x and Δ. We have

$$W(y, x; y - (x - \Delta), x - \Delta) = w(x - \Delta, \Delta)$$

and

$$W(y + x + \Delta, x + \Delta; y, x) = w(x, \Delta)$$

where

$$w(z, \Delta) = \begin{cases} g(z)/2 & \text{if } \Delta = \pm 1 \\ 1 - g(z) & \text{if } \Delta = 0 \end{cases} \tag{6}$$

By substituting $g(x)$ we obtain the following simple equation

$$\begin{aligned} P(y, x; T + 1) &= \frac{1}{2}[g(x + 1)P(y - (x + 1), x + 1; T) \\ &+ g(x - 1)P(y - (x - 1), x - 1; T)] \\ &+ (1 - g(x))P(y - x, x; T) \end{aligned} \tag{7}$$

The relevant boundary conditions are straightforward but lengthy to write down.

We now investigate the functional shape of the distribution $P(y, T) = \sum_x P(y, x, T)$ for different values of a, L and T, for the typical value $p = 5 \cdot 10^{-6}$. In Fig. 2 we plot a typical case from where the perfect agreement between the exact and simulation results is evident.
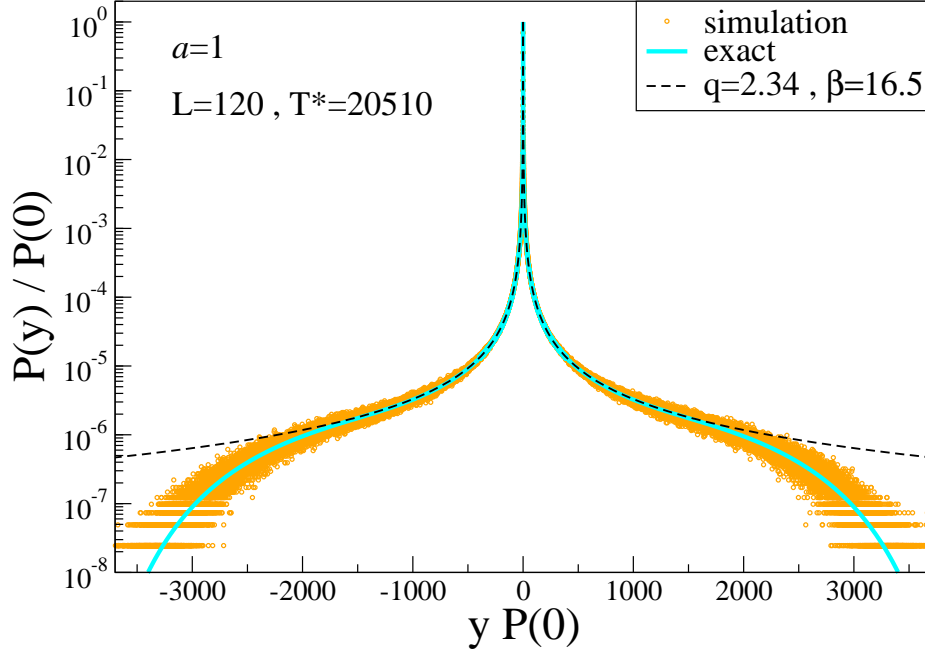
Fig. 2: Exact and simulation results of the case $a = 1$, $L = 120$ and $p = 5 \cdot 10^{-6}$. It is clearly seen that the probability function $P(y, T^*)$ obtained from simulations is completely in accordance with the exact results. The number of experiments used in our simulations are $2x10^8$.

Fig. 3 is concerned with the case $a = 1$ for different values of L. For fixed L we determine the value T^* for which an optimal fit to a q-Gaussian is possible. For $T > T^*$ values, each curve will start to exceed the q-Gaussian tails before the fast drop region, due to the finite-size, is being achieved. It is also interesting to analyze the relationship between T^* and L. We find that $T^* \sim L^2$, a behavior *identical* to the ordinary scaling that relates time and distance for diffusive processes.

Since we have an exact numerical solution we can investigate with great accuracy the nature of the convergence to the q-Gaussian as we increase the domain L of the random walker. In Fig. 4 we demonstrate that as L is increased a trajectory in the T^*-q-β parameter space exists along which $P(y, T^*, L)$ becomes increasingly well described by a q-Gaussian. The Figure contains the scaling combination $Y \equiv \ln_q[P(y, T)/P(0, T)]/[\beta s^2]$, where $s = yP(0)$, $\ln_q(x) = (x^{1-q} - 1)/(1 - q)$ is the q-logarithm and the scaling parameter s_{max} is defined as the s value of each L for which Y significantly starts to deviate from the -1 line, namely when $|Y + 1| > 0.004$. If the dependence on y is exactly q-Gaussian, we would have $Y = -1$ for all y. An appropriate scaling of x-axis yields a clear data collapse. For all values of L we observe the deviation from -1 to be no more than a few parts in a 1000 and as L increases the curves indeed approaches the line $Y = -1$ for large values of the argument y. The oscillations about the $Y = -1$ curve exhibit a subtle dependence on L. Careful inspection of the top panel in Fig. 4 reveals that for increasing values of L the curves actually approach the $Y = -1$ line for both small and large values of the argument s/s_{max}. We therefore believe that asymptotically the distribution $P(y, T^*)$ indeed becomes very well described by the q-Gaussian functional form. It is unfortunately numerically impossible for us to reach very large L-values.

This suggests the distribution $P(y, T, L)$ asymptotically is described by the q-Gaussian form if one let the pair (L, T) vary appropriately. We localize these very precise q and β values so that the curves are as symmetric as possible along the -1 line. Using the q values given in Fig. 4, we obtain an exponential dependence on L from where one can predict the
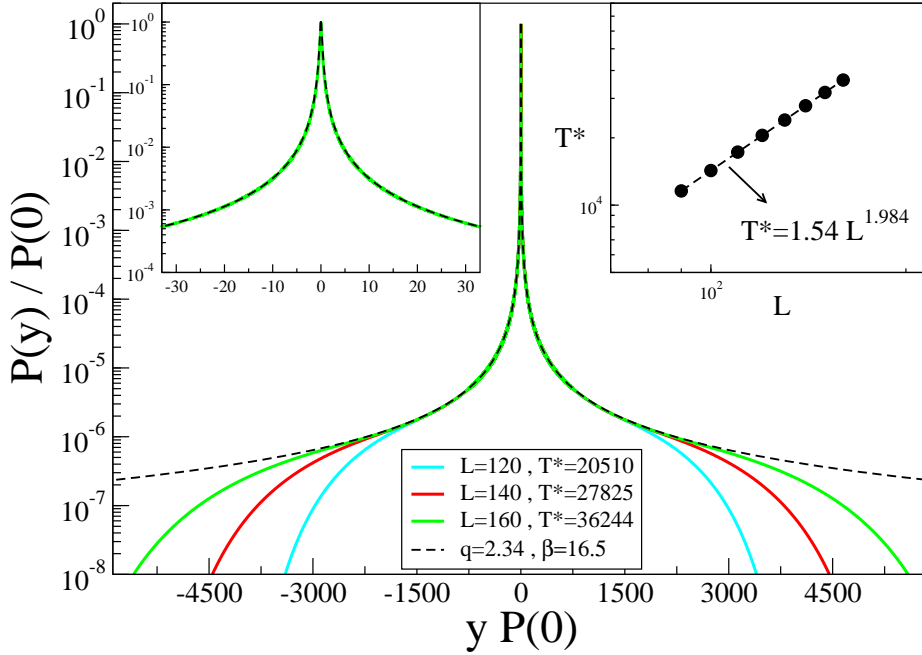
Fig. 3: The case $a = 1$ for $L = 120$, 140 and 160 with $p = 5 \cdot 10^{-6}$. The main panel shows the probability function $P(y, T^*)$. The center of the function is shown in detail in the left inset. The time T^* is chosen to optimize the fit to the q-Gaussian. The scaling of T^* is given in the right inset.

asymptotic value $\lim_{L \to \infty} q(L) = q_\infty \simeq 2.351$, which is evident from Fig. 5. It is interesting to note that for values $5/3 < q < 3$ the variance diverges. So in this respect the distribution behaves similarly to e.g. the Cauchy-Lorentz distribution, which corresponds to $q = 2$. Diverging variance is of course a common feature in complex systems of distributions with power law tails.

Next we consider the effect of changing a to values different from one. Although it is still possible to tune T^* so that the distribution $P(y, T^*)$ is very close to a q-Gaussian, the high resolution Y plot given in Fig. 6 now shows that the order of deviation from straight horizontal line through $Y = -1$ grows significantly whenever $a \neq 1$.

The rôle of correlations. – One might perhaps wonder to what extent the observed deviation from ordinary Gaussian behavior is caused by the peculiar shape of the probability distribution of the individual terms x_t in the sum $y = \sum_{t=1}^{T} x_t$. To check this we solved the Master Equation for the probability distribution for y in the uncorrelated case where all the individual terms in the sum are drawn independently with probability

$$p_{uc}(x) = \frac{\mathcal{N}}{|\frac{x}{L}| + p},$$

(8)

for $x \in \{-L, -L+1, .., 0, .., L\}$ and $\mathcal{N}$ the normalization factor. The motivation for this is simply that for the RRW considered above a term will appear in the sum y a number of times roughly given by $1/g(x)$. In Fig. 7 we show that when the terms are uncorrelated the sum converges towards an ordinary Gaussian. We note that the uncorrelated distribution $P(y, T)$ for small values of T does resemble a q-Gaussian in the region of small y values. However, as T is increased the functional form rapidly changes towards the ordinary Gaussian in stark contrast to the correlated case (left panel in Fig. 7) where the $P(y, T)$ grows towards the q-Gaussian as T is increased up to very large values of T. For the uncorrelated sum no trajectory (L, T^*) which for $L \to \infty$ takes one to the q-Gaussian exists.
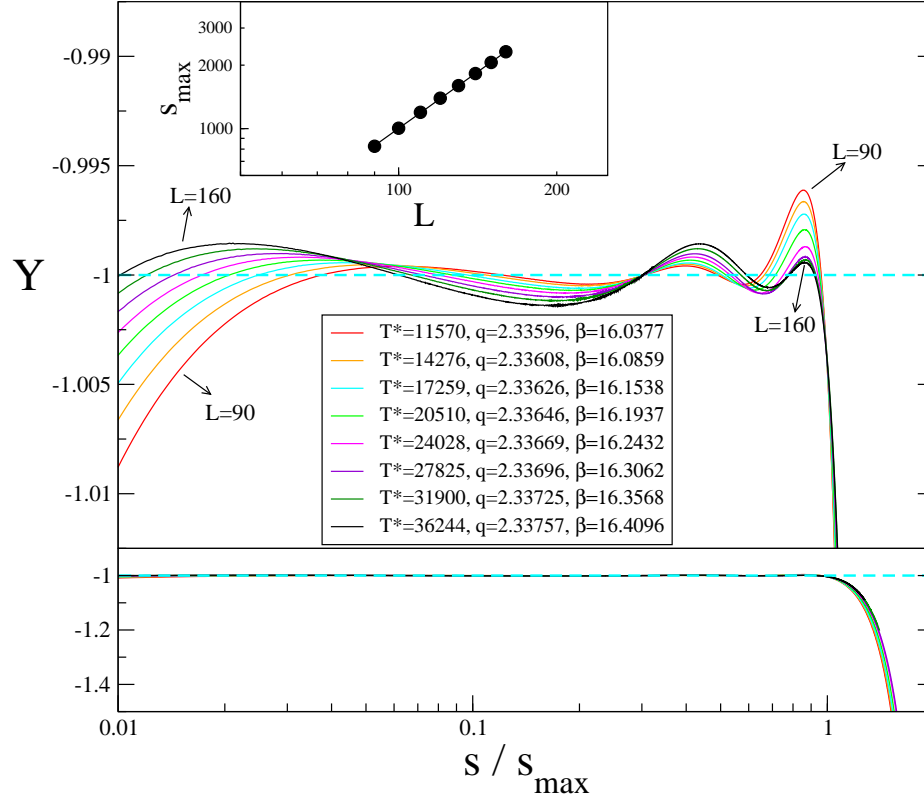
Fig. 4: The case $a = 1$ and $p = 5 \cdot 10^{-6}$ with L values between $L = 90$ to 160. The lower panel shows the data collapse when the x-axis is appropriately scaled. The upper panel shows a zoomed region around $Y = -1$ line. In the inset the scaling of s_{max} with L is given. The straight line is $s_{max} = AL^C$ with $A = 0.2726$ and $C = 1.7828$.

Conclusions. –   We have presented the hitherto most simple setting in which q-Gaussians control asymptotic behavior. We conclude that the q-Gaussian behavior is brought about by the strong correlations and the high reluctance for the walker to move away from the central region of its domain.

The numerical exact solution of Master Equations allows us to present high precision data for the probability function of sums of correlated random variables derived from a restricted random walk (RRW) with position dependent transition probabilities. When the range of the walker L and the number of terms in the sum T is scaled according to $T = 1.54L^2$, q-Gaussians are observed over an increasingly broad interval. For non-linear transition probabilities we are able to identify a subtle oscillatory behavior away from the pure q-Gaussian form. Given the relative simplicity of the RRW it appears likely that the relation between transition probability and the value of q and the existence of oscillatory corrections to the q-Gaussian asymptote can be unraveled analytically. The RRW model presented in the present letter promises this way to significantly increase our understanding of the mechanisms responsible for the often encountered q-Gaussians.

The very weak dependence of q on $T^*(L)$ and the subtle oscillations in the $a = 1$ case and the more essential oscillations present for $a \neq 1$ indicates that the true exact mathematical asymptote might not strictly be q-Gaussian but rather some functional form resembling a q-Gaussian to a high degree of accuracy. It is only because we have numerically iterated the Master Equations exactly that we are able to identify this very slight difference. In studies relying on simulations or observational data the accuracy may not be sufficient to resolve
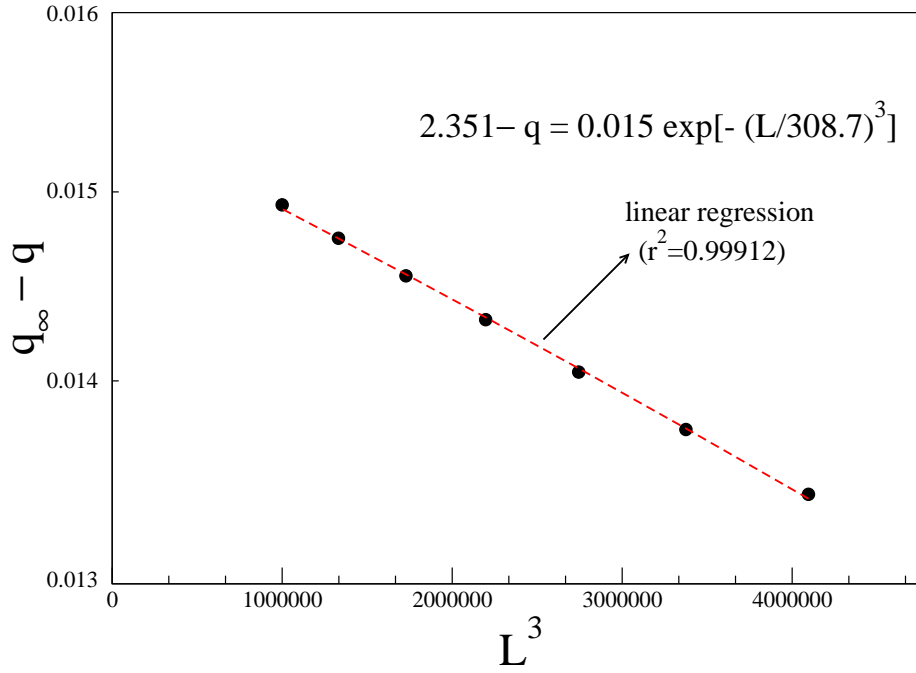
Fig. 5: Linear-log representation for the L dependence of q values. This exponential dependence suggests an asymptotic value around $q_\infty \simeq 2.351$.
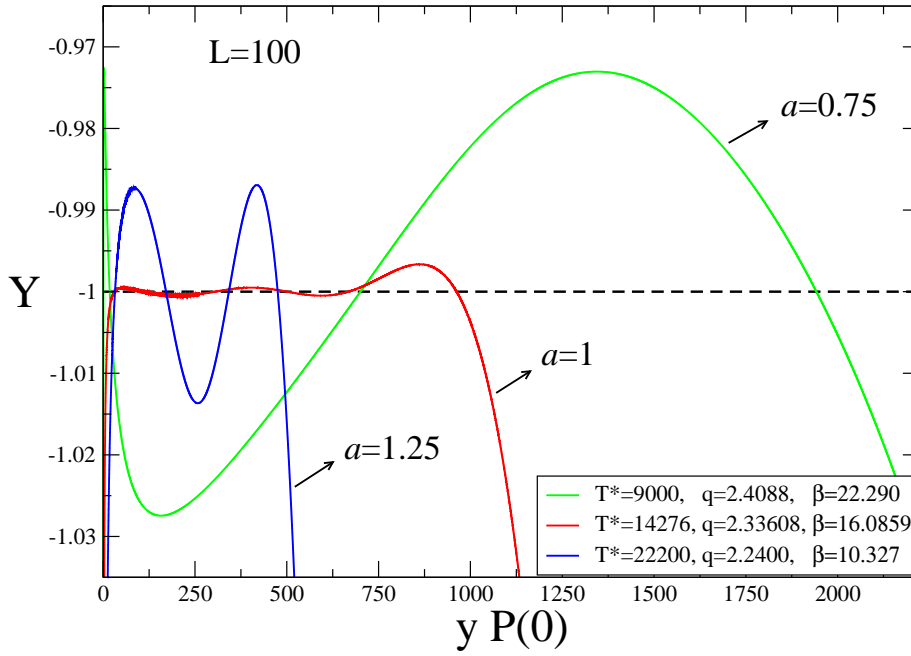

Fig. 6: Y plot of cases $a = 0.75$, $a = 1$ and $a = 1.25$ for a representative L value. Whenever $a \neq 1$, increasing order of deviation from -1 line is evident.
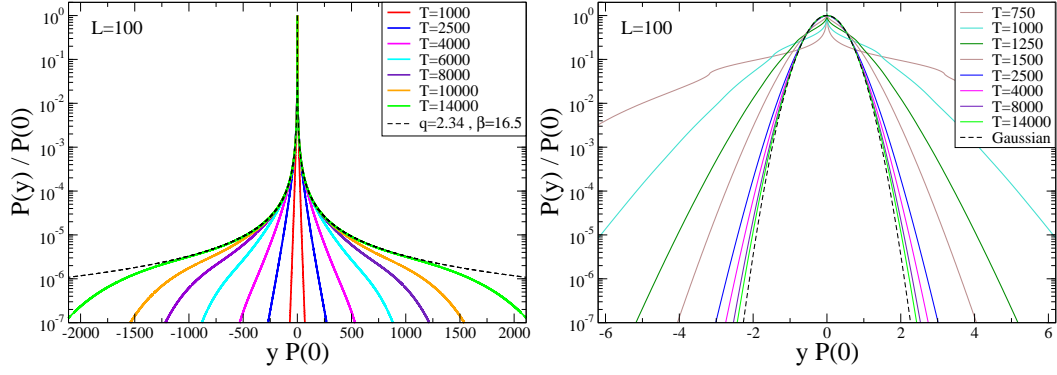
Fig. 7: Comparison of the T dependence for fixed values of L with $p = 5 \cdot 10^{-6}$ for the correlated (left panel) and the uncorrelated (right panel) $P(y,T)$ distribution (see text for the details).

these details and one would conclude that a q-Gaussian is an excellent approximation of the observed behavior.

Let us recall that q-Gaussians have been found previously for more complex processes than the random walk to be able to provide very high quality approximations to relevant distributions. The case $q < 1$ was considered in [8,9] where the authors found for two (scale-invariant) probabilistic models that the large-size limiting distributions are amazingly close to q-Gaussians, but are not exactly q-Gaussians [10]. The work in Ref. [6] provides an analytic example of large-size limiting distributions that are q-Gaussians. We stress that even if q-Gaussians are not always the *exact* analytic form of the probability distributions in question, it is highly intriguing why they provide such exceptionally high accuracy approximations in a large number of cases where correlations are sufficiently strong to make the central limit theorem inapplicable.

$$* * *$$

REFERENCES

[1] VAN KAMPEN N.G., *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam) 1981.
[2] KHINCHIN A.YA., *Mathematical Foundations of Statistical Mechanics* (Dover, New York) 1949.
[3] FELLER W., *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, New York) 1970.
[4] REIF F, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York) 1965.
[5] UMAROV S., TSALLIS C. and STEINBERG S., *Milan J. Math.*, **76** (2008) 307; UMAROV S., TSALLIS C. GELL-MANN M. and STEINBERG S., *J. Math. Phys.*, **51** (2010) 033502; UMAROV S. and TSALLIS C., *Phys. Lett. A*, **372** (2008) 4874; VIGNAT C. and PLASTINO A., *J. Phys. A*, **40** (2007) F969; HAHN M.G., JIANG X.X. and UMAROV S., *J. Phys. A*, **43** (2010) 165208.
[6] RODRIGUEZ A., SCHWAMMLE V. and TSALLIS C., *JSTAT*, **2008** (P09006) ; HANEL R., THURNER S. and TSALLIS C., *Eur. Phys. J. B*, **72** (2009) 263.

[7] TSALLIS C., *J. Stat. Phys.*, **52** (1988) 479; TSALLIS C., *Introduction to nonextensive statistical mechanics: approaching a complex world* (Springer, New York) 2009.

[8] MOYANO L.G., TSALLIS C. and GELL-MANN M., *Europhys. Lett.*, **73** (2006) 813.

[9] THISTLETON W.J., MARSH J.A., NELSON K.P. and TSALLIS C., *Cent. Eur. J. Phys.*, **7** (2009) 387.

[10] HILHORST H.J. and SCHEHR G., *J. Stat. Mech*, (2007) P06003.