JOURNAL OF MATHEMATICAL PHYSICS **56**, 053301 (2015)

Emergence of *q*-statistical functions in a generalized binomial distribution with strong correlations

G. Ruiz[1] and C. Tsallis[2,3]
[1]*Dpto. de Matemática Aplicada y Estadística, Universidad Politécnica de Madrid, Pza. Cardenal Cisneros n. 3, 28040 Madrid, Spain*
[2]*Centro Brasileiro de Pesquisas Físicas and National Institute of Science and Technology for Complex Systems, Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro, Brazil*
[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

(Received 24 November 2014; accepted 22 April 2015; published online 5 May 2015)

We study a symmetric generalization $\mathfrak{p}_k^{(N)}(\eta,\alpha)$ of the binomial distribution recently introduced by Bergeron *et al.*, where $\eta \in [0,1]$ denotes the win probability and α is a positive parameter. This generalization is based on q-exponential generating functions ($e^z_{q^{gen}} \equiv [1+(1-q^{gen})z]^{1/(1-q^{gen})}$; $e^z_1 = e^z$) where $q^{gen} = 1 + 1/\alpha$. The numerical calculation of the probability distribution function of the number of wins k, related to the number of realizations N, strongly approaches a discrete q^{disc}-Gaussian distribution, for win-loss equiprobability (i.e., $\eta = 1/2$) and all values of α. Asymptotic $N \to \infty$ distribution is in fact a q^{att}-Gaussian $e_{q^{att}}^{-\beta z^2}$, where $q^{att} = 1 - 2/(\alpha-2)$ and $\beta = (2\alpha-4)$. The behavior of the scaled quantity k/N^{γ} is discussed as well. For $\gamma < 1$, a large-deviation-like property showing a q^{ldl}-exponential decay is found, where $q^{ldl} = 1 + 1/(\eta\alpha)$. For $\eta = 1/2$, q^{ldl} and q^{att} are related through $1/(q^{ldl}-1) + 1/(q^{att}-1) = 1$, $\forall\alpha$. For $\gamma = 1$, the law of large numbers is violated, and we consistently study the large-deviations with respect to the probability of the $N \to \infty$ limit distribution, yielding a power law, although not exactly a q^{LD}-exponential decay. All q-statistical parameters which emerge are univocally defined by (η,α). Finally, we discuss the analytical connection with the Pólya urn problem. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4919678]

I. INTRODUCTION

Probability distributions that take correlations into account, can be in some cases constructed by deformation of mathematical independent laws.[1–4] Along this line of approach, a variety of generalizations of the binomial distribution have been recently proposed.[5–7] The generalization consists in replacing the sequence of natural numbers that correspond to the variable of the original binomial distribution by an arbitrary sequence of non-negative numbers. Consequently, their factorial and combinatorial numbers are redefined, in such a way that the simple powers of "win" ("loose") probability must be replaced by characteristic polynomials whose degree is the number of wins (loose). These polynomials are obtained through generating functions that force the generalized distributions to still satisfy the conditions of normalization and non-negativeness. Resulting probabilities $\mathfrak{p}_k^{(N)}$ can be symmetrical or asymmetrical, and represent the probability of having k wins and $(N-k)$ losses, in a sequence of N *correlated* trials.

We shall be concerned with a particular set of *symmetrically* generalized binomial distributions, so as to preserve the win-loss symmetry, which is an essential prerequisite for the distribution to be used in the present analysis of strongly correlated systems and their entropic behavior. The generating functions of this particular set are q^{gen}-exponentials ($e^z_{q^{gen}} \equiv [1+(1-q^{gen})z]^{1/(1-q^{gen})}$; $e^z_1 = e^z$), where *gen* stands for *generating* and $q^{gen} = 1 + 1/\alpha > 1$ ($\alpha > 0$), α being a parameter to be soon defined. These generating functions are the only ones, besides the ordinary binomial case (which corresponds $q^{gen} \to 1$), that yield probabilities which obey the Leibnitz triangle rule (see, for instance, Refs. 7 and 8). The family of the generated probabilities depends on two parameters

$(\eta,\ \alpha)$, where η is the "win" probability, $(1-\eta)$ is the "loss" probability, and α characterizes the generating function. A variety of q-statistical functions[9,10] related to the probabilities $\mathfrak{p}_k^{(N)}(\eta,\alpha)$ emerges, all of them univocally defined by (η,α), and whose respective $q(\eta,\alpha)$ indices appear to obey an algebra that reminds the underlying algebra in Refs. 8 and 11. The analogy of the algebras involved points out to a common underlying mathematical structure that remains to be understood better.

In fact, such probabilities provide a probability distribution function of the scaled quantity k/N $(k/N = 0, 1/N, \ldots, 1)$ that is likely to achieve, for all fixed values of (α, N), any of the complete set[12] of compact-support q^{disc}-Gaussian distributions functions, $e_{q^{disc}}^{-\beta^{disc} z^2} = [1-\beta^{disc}(1-q^{disc})z^2]^{1/(1-q^{disc})}$, where *disc* stands for *discrete*, $q^{disc} < 1$, and β^{disc} is a generalized inverse temperature $(\beta^{disc} \in \mathbb{R})$. The q-Gaussian form corresponds to strongly correlated random variables and arises from the extremization of the nonadditive entropy $S_q = k_B(1-\sum_i p_i^q)/(q-1)$ $(q \in \mathbb{R}, S_1 = S_{BG} \equiv -k\sum p_i \ln p_i$, where BG stands for Boltzmann-Gibbs)[9] under appropriate constraints.[14,13]

Since the proposal in Ref. 10, several statistical models which provide, in the $N \to \infty$ limit, q-Gaussian attractors have been constructed.[15–17] Some of them exhibit extensivity of the Boltzmann-Gibbs entropy S_{BG}, and at least one of them exhibits extensivity of the S_q entropy for $q \neq 1$.[17] In particular, the q^{gen}-exponentially generated probabilities $\mathfrak{p}_k^{(N)}(\eta,\alpha)$ exhibit an extensive S_{BG},[18] and they appear to provide q^{att}-Gaussian attractors (where *att* stands for *attractor*). An outstanding fact is that these probabilities have been rigorously deduced *a priori*, imposing a particular structure of the generating function (see also Refs. 19 and 20) under non-negativeness and normalization probability conditions.

In addition to these properties, these probabilities can be shown to correspond to the Pólya urn model.[21] Such a model considers the following urn scheme: from a set of b black balls and r red balls contained in an urn, one extracts one ball and returns it to the urn, together with c balls of the same color. In that case, $\mathfrak{p}_k^{(N)}(\eta,\alpha)$ represents the probability to have k black balls in the urn after the N-th trial, and it can be written as a function of (b,r,c), as $\eta = b/(b+r)$ and $\alpha = (b+r)/c$. All this means that all the properties of this particular set of symmetrically generalized binomial distributions will be translated to a particular urn model.

The paper is organized as follows. Section II presents the symmetric generalized binomial distribution introduced by Bergeron *et al.*,[7] based on the q^{gen}-exponential generating functions. Section III is devoted to the characterization of the involved q-Gaussian distributions (where q refers to q^{disc} for finite N, and to q^{att} in the $N \to \infty$ limit) in the generalized probability distribution of the ratio (number of wins)/(number of realizations). Section IV describes large-deviation-like properties of the distribution of the scaled quantity k/N^γ $(0 < \gamma < 1)$. Section V deals with the behavior of the large-deviation probabilities with regard to the $N \to \infty$ limit distribution, which violates the law of large numbers. We conclude in Sec. VI.

II. THE SYMMETRIC GENERALIZED BINOMIAL DISTRIBUTION

In a sequence of $N \in \mathbb{N}$ independent trials with two possible outcomes, "win" and "loss," the probability of obtaining k wins is given by the binomial distribution,

$$p_k^{(N)}(\eta) = \binom{N}{k}\eta^k(1-\eta)^{N-k} = \frac{N!}{(N-k)!\,k!}\eta^k(1-\eta)^{N-k}, \tag{1}$$

where the parameter η $(0 \le \eta \le 1)$ is the probability of having the outcome "win," $(1-\eta)$ corresponding to the outcome "loss." Therefore, the Bernoulli binomial distribution above preserves the symmetry win-loss.

Let us now consider a strictly increasing infinite sequence of *non-negative real* numbers $\chi = \{x_N\}_{N\in\mathbb{N}}$. With each sequence χ defined above, a Bernoulli-like distribution is constructed,

$$\mathfrak{p}_k^{(N)}(\eta) = \frac{x_N!}{x_{N-k}!\,x_k!}q_k(\eta)q_{N-k}(1-\eta), \tag{2}$$

where the factorials are defined as $x_N! \equiv x_1 x_2 \ldots x_N$, $x_0! \equiv 1$, η is a running parameter on the interval $[0,1]$ and $q_k(\eta)$ are polynomials of degree k. Observe that the symmetry win-loss of binomial-like distribution (2) is preserved, as invariance under $[k,\eta] \to [(N-k),(1-\eta)]$ is verified. That means that no bias can exist favoring either win or loss when $\eta = 1/2$.

The polynomials $q_k(\eta)$ are to be defined in such a way that quantities $\mathfrak{p}_k^{(N)}(\eta)$ represent the probabilities of having k wins and $(N-k)$ losses in a sequence of *correlated N trials*. Consequently, $\mathfrak{p}_k^{(N)}(\eta)$ must be constrained by the normalization equation,

$$\sum_{k=0}^{N} \mathfrak{p}_k^{(N)}(\eta) = 1, \qquad \forall N \in \mathbb{N}, \forall \eta \in [0,1], \tag{3}$$

and the non-negativeness condition,

$$\mathfrak{p}_k^{(N)}(\eta) \geq 0, \qquad \forall N, k \in \mathbb{N}, \forall \eta \in [0,1]. \tag{4}$$

Different sets of polynomials can be associated with (3) and (4). With this aim, some generating functions of polynomials can be considered. Let us make use of a q^{gen}-exponential generating function $e^{z}_{q^{gen}} = [1 + (1-q^{gen})z]^{1/(1-q^{gen})}$. Such a generating function can be written as $\mathcal{N}(z) = (1 - z/\alpha)^{-\alpha}$ $(\alpha > 0)$, where the q-exponential parameter that characterizes the generating function is $q^{gen} = 1 + 1/\alpha > 1$. The following probability distributions are consequently obtained:[7]

$$\mathfrak{p}_k^{(N)}(\eta,\alpha) = \binom{N}{k} \frac{(\eta\alpha)_k \, ((1-\eta)\alpha)_{N-k}}{(\alpha)_N}, \tag{5}$$

where $(a)_b \equiv a(a+1)(a+2)\ldots(a+b-1)$ is the Pochhammer symbol and $\alpha = 1/(1-q^{gen})$. Observe that the $\alpha \to \infty$ limit recovers the ordinary binomial case ($q^{gen} \to 1$). The expectation value and the variance of (5) are $\langle k \rangle_N(\eta) = \eta N$ and $(\sigma_k)^2_N(\eta,\alpha) = N^2 \eta(1-\eta)\dfrac{1+\alpha/N}{1+\alpha}$, respectively.[7]

In the particular cases $\eta = 1/2$ and $4 \leq \alpha = \dot{2}$ (i.e., $\alpha = 4, 6, 8, \ldots$), we can also use the following equivalent expression:

$$\mathfrak{p}_k^{(N)}(\eta = 1/2, \alpha) = \begin{cases} \dfrac{(\alpha/2)_N}{(\alpha)_N}, & (k = 0, N), \\ \left[\displaystyle\prod_{m=1}^{\alpha-1} \frac{\alpha - m}{N + \alpha - m}\right] \times \left[\displaystyle\prod_{j=1}^{\frac{\alpha}{2}-1} \left(\frac{k+j}{j}\right)\left(\frac{N-k+j}{j}\right)\right], & (k \neq 0, N). \end{cases} \tag{6}$$

This expression is computationally very convenient.

III. EMERGENCE OF *q*-GAUSSIANS

The histograms $p(k/N) \equiv N \mathfrak{p}_k^{(N)}(\eta,\alpha)$ $(0 \leq k/N \leq 1)$ are numerically obtained for fixed values of N, η, and α. Fig. 1 shows that, for $\eta = 1/2$ and $\alpha \to \infty$, $p(k/N)$ approaches the unbiased binomial distribution for all values of N.

From now on, unless specified otherwise, we consider the case where no bias exists, i.e., $\eta = 1/2$.

Fig. 2 illustrates, for $\alpha = 15$ and $N = 100$, the distribution $p(k/N)$ normalized to its maximum p_{max}. Other values of α and N have been studied as well and, in all cases, numerical results strongly suggest $q^{disc}(\alpha,N)$-Gaussian distributions (with $q^{disc} < 1$), i.e.,

$$p(k/N)/p_{max} \simeq e_{q^{disc}}^{-\beta^{disc}(\frac{k}{N}-\frac{1}{2})^2} = \left[1 - \beta^{disc}(1-q^{disc})\left(\frac{k}{N} - \frac{1}{2}\right)^2\right]_+^{\frac{1}{1-q^{disc}}}, \tag{7}$$

where $[x]_+ = x$ if $x > 0$, and $[x]_+ = 0$ otherwise, and where $q^{disc}(\alpha,N)$ dependence is omitted for simplicity. Table I shows the values of q^{disc} for typical values of α and N. For a fixed value of N,

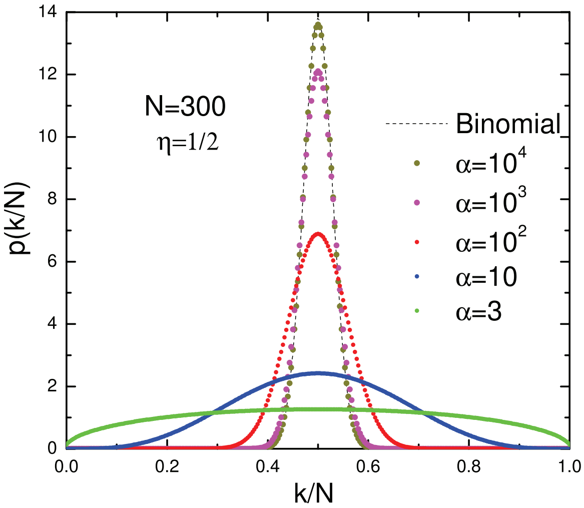

FIG. 1. Probability of having a ratio number of wins k/N $(0 \le k/N \le 1)$ in a sequence of $N = 300$ correlated trials that follow the symmetric generalized binomial distribution (5). The parameters of the model are $\eta = 1/2$, $\alpha = 3, 10, 10^2, 10^3, 10^4$. Observe that the limit distribution tends to a binomial distribution as $\alpha \to \infty$. If, in addition to this limit, we consider $N \to \infty$, the distribution shrinks onto a Dirac delta one.

q^{disc} is a monotonous function of α. Similarly, for a fixed value of α, q^{disc} is a monotonous function of N.

Increasing the value of N, a sequence of $p(k/N)$ distributions is obtained (see Fig. 3). Their corresponding $q^{disc}(N)$-logarithmic representation shows that $q^{disc}(N)$-Gaussian distributions fit very well the data.

In the limit $N \to \infty$, we define

$$q^{att}(\alpha) \equiv \lim_{N\to\infty} q^{disc}(\alpha, N) \tag{8}$$

and the corresponding $(N \to \infty)$ limit distribution is a q^{att}-Gaussian. This result that can in fact be rigorously proved.[18] All this means that, in all cases, finite N distributions are numerically well represented by q^{disc}-Gaussians that approach, in the $N \to \infty$ limit, to a q^{att}-Gaussian limit distribution.

The value of q^{att} depends only on the parameter α, and some values of the (q^{att}, α) pair are shown in Table II. We heuristically conclude that

$$q^{att}(\alpha) = 1 - \frac{2}{\alpha - 2}, \tag{9}$$

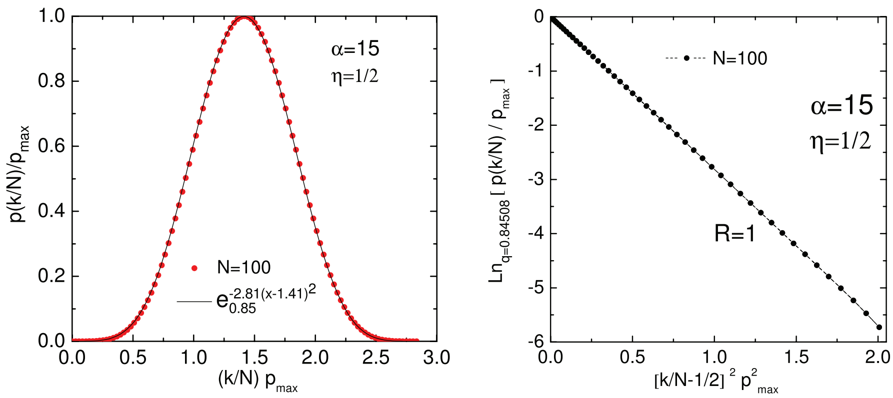

FIG. 2. Normalized symmetric generalized binomial distribution for $\eta = 1/2$, $N = 100$, and $\alpha = 15$. Left panel: The corresponding q^{disc}-Gaussian fitting function is superimposed. Right panel: The q^{disc}-logarithmic representation exhibits that indeed the discrete (i.e., $N < \infty$) distribution is extremely close to a q-Gaussian, the linear regression coefficient being $R = 1$. The slope of linear regression in right panel is proportional to the inverse temperature β^{disc}, as shown in Eq. (7).

TABLE I. Numerical $q^{disc}(\alpha, N)$ parameter of the $q^{disc}(\alpha, N)$-Gaussian distribution strongly suggested by $p(k/N) = N\mathfrak{p}_k^{(N)}$. The last column corresponds to the respective values of $\lim_{N\to\infty} q^{disc}$ (see text).

	$N = 50$	$N = 60$	$N = 70$	$N = 80$	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N \to \infty$
$\alpha = 3$	−0.98 990	−0.99 170	−0.99 306	−0.99 410	−0.99 524	−0.99 800	−0.99 930	−0.99 970	−1
$\alpha = 5$	0.33 117	0.33 185	0.33 214	0.33 233	0.33 264	0.33 311	0.33 324	0.33 328	1/3
$\alpha = 15$	0.84 235	0.84 350	0.84 412	0.84 456	0.84 508	0.84 587	0.84 611	0.84 614	11/13
$\alpha = 25$	0.90 803	0.90 939	0.91 027	0.91 084	0.91 154	0.91 263	0.91 297	0.91 303	21/23
$\alpha = 50$	0.95 127	0.95 298	0.95 413	0.95 494	0.95 598	0.95 763	0.95 821	0.95 830	23/24
$\alpha = 100$	0.97 043	0.97 244	0.97 382	0.97 483	0.97 616	0.97 845	0.97 936	0.97 953	48/49
$\alpha = 500$	0.98 375	0.98 597	0.98 756	0.98 874	0.99 039	0.99 359	0.99 531	0.99 576	248/249
$\alpha = 1000$	0.98 521	0.98 745	0.98 904	0.99 025	0.99 190	0.99 521	0.99 710	0.99 766	498/499
$\alpha \to \infty$	1	1	1	1	1	1	1	1	1

i.e., $1/(1 - q^{att}) = (\alpha - 2)/2$, which we have numerically verified at various scales.

Also, we conclude that the indexes $q^{att}(\alpha)$ and $q^{gen}(\alpha)$ are related by

$$\frac{1}{q^{gen}(\alpha) - 1} - \frac{2}{q^{att}(\alpha) - 1} = 2 \tag{10}$$

that reminds the equations of the algebra indicated in Ref. 11.

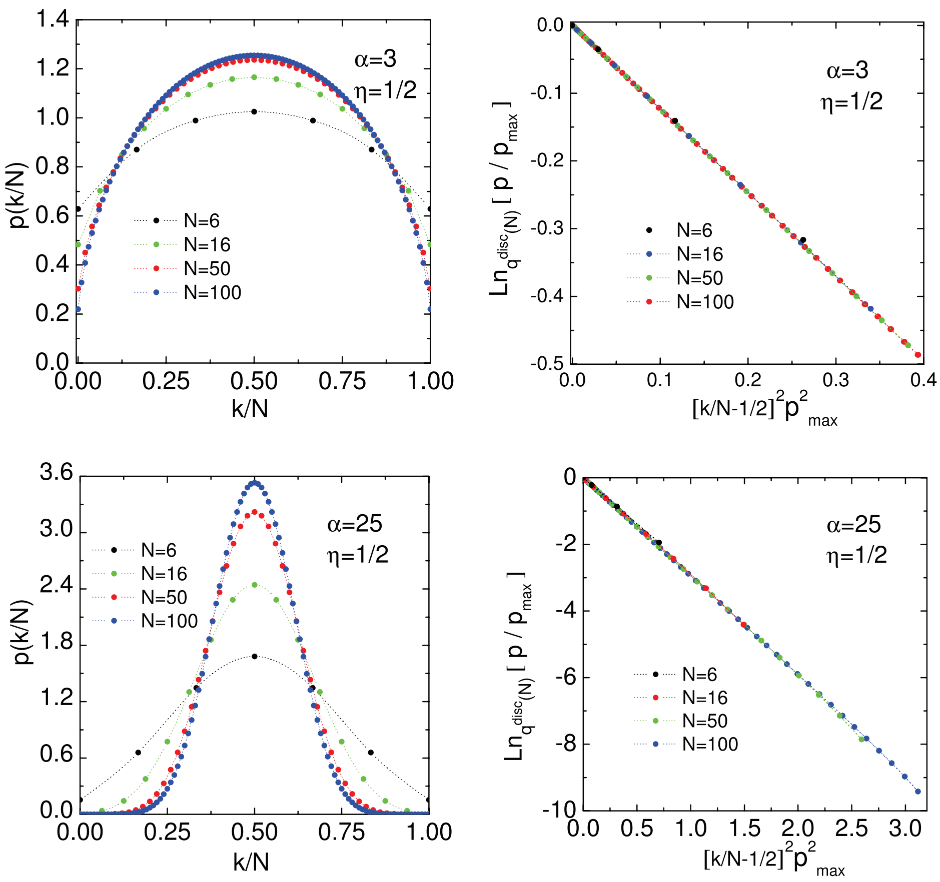

FIG. 3. Left panels: Probability distribution $p(k/N) = N\mathfrak{p}_k^{(N)}$ $(0 \le k/N \le 1)$ for typical values of α and N. Right panels: The corresponding $q^{disc}(N)$-logarithmic representations show that the discrete distributions are very close to q-Gaussians, except for the last point of the tail. Similar results are obtained for other parameter values. We notice that for $\alpha > 4$ $(\alpha < 4)$, which corresponds to $q^{att} > 0$ $(q^{att} < 0)$, the terminal derivative in the linear-linear representation vanishes (diverges). This derivative is finite for $\alpha = 4$, which corresponds to $q^{att} = 0$.

TABLE II. Parameter $q^{att}(\alpha) = \lim_{N\to\infty} q^{disc}(\alpha, N)$ that characterizes the $N \to \infty$ limit q^{att}-Gaussian distribution.

α	2	2.1	2.5	3	4	5	15	25	50	100	500	1000
q^{att}	$-\infty$	−19	−3	−1	0	1/3	11/13	21/23	23/24	48/49	248/249	498/499

Let us now analyze the generalized inverse temperature, β, of non-normalized $q^{att}(\alpha)$-Gaussian distributions. Fig. 4 plots the numerically obtained values, for typical values of $q^{att}(\alpha)$. From our numerics, we heuristically conclude that the following equation is satisfied:

$$\frac{1}{\beta(\alpha)} = \frac{1 - q^{att}(\alpha)}{4}. \tag{11}$$

This corresponds to the cutoffs of the distributions.

We can infer, from (9) and (11), the following simple relation:

$$\beta(\alpha) = 2(\alpha - 2) \quad (\alpha > 0)\,, \tag{12}$$

the $\alpha > 2$ values corresponding to positive values of β and the $0 < \alpha < 2$ values corresponding to negative values of β.

Summarizing, Fig. 5 shows that $\alpha > 2$ provides bell-shaped and compact support q^{att}-Gaussian distributions ($q^{att} < 1$, $\beta > 0$), and $\alpha \in (0,2)$ provides convex and bounded but noncompact support q^{att}-Gaussian distributions ($q^{att} > 2$, $\beta < 0$). In the $\alpha \to 2$ limit, an uniform distribution is obtained. This distribution is the limit of a q^{att}-Gaussian distribution whose $\lim_{\alpha\to 2^+} q^{att}(\alpha) = -\infty$ with $\beta > 0$, as well as $\lim_{\alpha\to 2^-} q^{att}(\alpha) = +\infty$ with $\beta < 0$. In the $\alpha \to 0$ limit, a double peaked delta distribution emerges, i.e., the distribution is the limit of a $q^{att}(\alpha)$-Gaussian distribution whose $\lim_{\alpha\to 0} q^{att}(\alpha) = 2$ with $\beta < 0$. This description illustrates the diagram presented in Ref. 12, where negative generalized temperatures β^{-1} of q-Gaussian distributions are also considered.

Making use of the expression of a normalized q-Gaussian,[14] it can be consequently stated that the generalized distribution $p(k/N)$ defined in Eq. (5) corresponds, for win-loss equiprobability (i.e., $\eta = 1/2$), to the following q^{att}-Gaussian limit distribution:

$$p_\infty(x;\alpha) \equiv \lim_{N\to\infty} p\left(\frac{k}{N};\alpha\right) = \frac{2\,\Gamma\left(\frac{5-3q^{att}}{2(1-q^{att})}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{2-q^{att}}{1-q^{att}}\right)} \left[1 - 4\left(x - \frac{1}{2}\right)^2\right]_+^{\frac{1}{1-q^{att}}}, \tag{13}$$

where $\alpha > 0$, $q^{att} = 1 - 2/(\alpha - 2)$, and $0 \le x \le 1$.

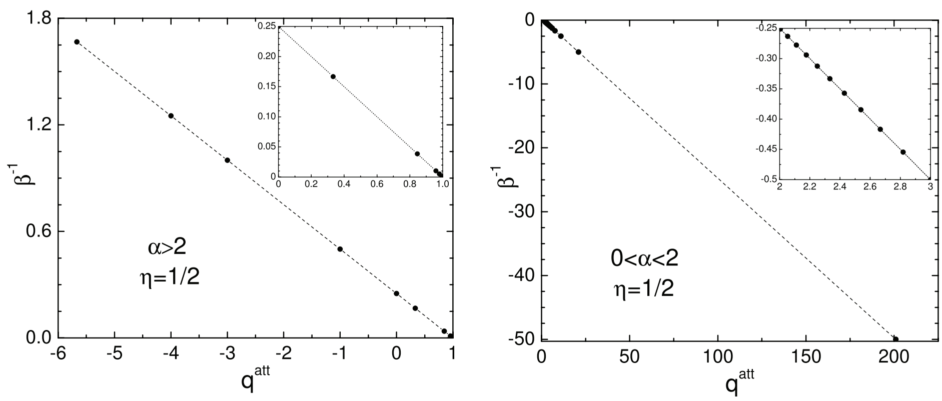

FIG. 4. Non-normalized histogram demonstrates simple $\beta(q^{att})$ and $\beta(\alpha)$ dependence for the q^{att}-Gaussian limit distribution. Left panel: $\alpha > 2$ provides positive generalized temperature of the q^{att}-Gaussian. Right panel: $0 < \alpha < 2$ provides negative generalized temperature of the q^{att}-Gaussian.

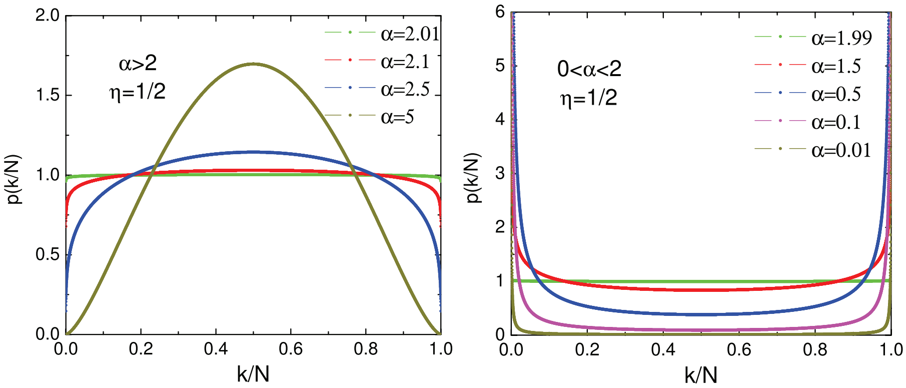

FIG. 5. *Left panel*: The $\alpha > 2$ ($N \to \infty$) limit distributions are concave with bounded and compact support. *Right panel*: The $0 < \alpha < 2$ ($N \to \infty$) limit distributions are convex with bounded but noncompact support. The limiting case $\alpha = 2$ corresponds to the uniform distribution.

IV. LARGE-DEVIATION-LIKE PROPERTIES

Let us now consider that each of the N single variables takes values 1 or 0 (i.e.,"win" or "loose"). Consequently, the value of k corresponds to the sum of all N binary random variables, and, after centering and re-scaling k, the attractors that emerge correspond to the abscissa currently associated with central limit theorems. In the case of the unbiased symmetric generalized probability distribution (i.e., $\eta = 1/2$) defined in Eq. (5), the abscissa measured from its central value scales as N^γ with $\gamma = 1$, and emerging attractors of $p(k/N)$ are the q^{att}-Gaussians defined in Eq. (13). This fact precludes the vanishing limit of the probability P of a deviation of k/N from its central value $k/N = 1/2$, i.e.,

$$\lim_{N\to\infty} P(N; |k/N - 1/2| \geq \epsilon) \neq 0 \ \ (\forall \epsilon > 0), \tag{14}$$

where ϵ is the minimum deviation of k/N with respect to central value k/N and P can be evaluated adding up the weight of the possible values of k which do not fall inside $(N/2 - \epsilon, N/2 + \epsilon)$. Taking into account the symmetry of the distribution and denoting $x \equiv 1/2 - \epsilon$, we can write

$$P(N; |k/N - 1/2| \geq \epsilon) = 2P(N; k/N \leq x) = 2 \sum_{k=0}^{\lfloor Nx \rfloor} \mathfrak{p}_k^{(N)} \tag{15}$$

and conclude that, $\forall \alpha > 0$,

$$\lim_{N\to\infty} P(k/N \leq x; \alpha) = \lim_{N\to\infty} \sum_{k=0}^{\lfloor Nx \rfloor} \mathfrak{p}_k^{(N)}(\alpha) = \int_{0 \leq x \leq \frac{k}{N}} p_\infty(x;\alpha) dx \neq 0, \tag{16}$$

where $\lfloor Nx \rfloor$ is the largest integer number that $\lfloor Nx \rfloor \leq Nx$ and $p_\infty(x;\alpha)$ is defined in Eq. (13). Equation (16) states that these correlated models do *not* satisfy the classical version of the weak law of large numbers.

Due to this fact, let us consider instead the scaled quantity k/N^γ, $\gamma \neq 1$, in analogy with the anomalous diffusion coefficient introduced in nonlinear Fokker-Plank equations.[22] More precisely, the case $\gamma \neq 1$ is similar to anomalous diffusion, where the square space is nonlinear with time.

In the case $\gamma < 1$, we have numerically found a q^{ldl}-exponential (where ldl stands for *large-deviation-like*) decaying behavior of the probability $P(N; k/N^\gamma \leq x)$, for typical values of α, when N increases. In fact (see Fig. 6), it can be straightforwardly verified that

$$\lim_{N\to\infty} P(N; \frac{k}{N^\gamma} \leq x) = \lim_{N\to\infty} \sum_{k=0}^{\lfloor N^\gamma x \rfloor} \mathfrak{p}_k^{(N)} = 0 \tag{17}$$

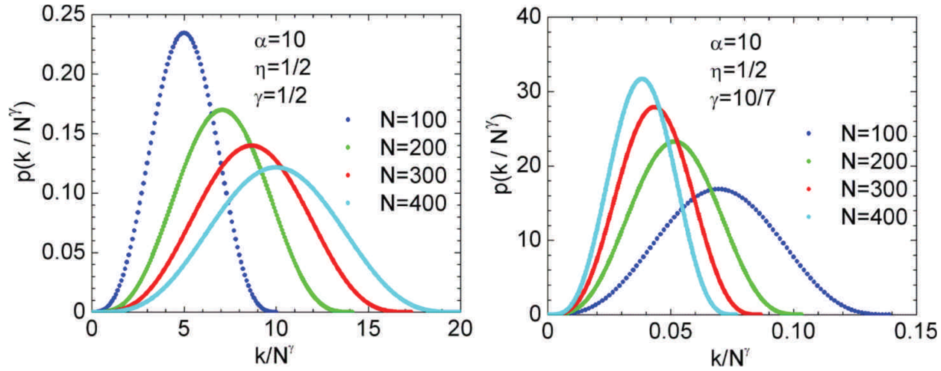

FIG. 6. Non-centered histograms of k/N^γ, $p(k/N^\gamma) = N^\gamma p_k^{(N)}$, for $\eta = 1/2$, $\alpha = 10$, and typical sequences of N. Left panel: Scaling factor $\gamma < 1$ makes the support $N^{1-\gamma}$ to increase with N. Right panel: Scaling factor $\gamma > 1$ makes the support $N^{1-\gamma}$ to decrease with N.

and the zero limit is attained as

$$P(N; \frac{k}{N^\gamma} < x) \sim e_{q^{ldl}}^{-N r_{q^{ldl}}} = \left[1 - (1 - q^{ldl}) N r_{q^{ldl}}\right]^{1/(1-q^{ldl})}, \tag{18}$$

with $q^{ldl} > 1$, where $r_{q^{ldl}}$ represents the decaying rate (*rate function*) of a q^{ldl}-exponential. This result can be illustrated in Fig. 7, where $q^{disc}(N)$-logarithmic representation of $P(N; \frac{k}{N^\gamma} < x)$ as a function of N is shown, for $\alpha = 10$, $\gamma = 1/2$, and $\eta = 1/2$. Straight lines are obtained for all values of α ($\alpha > 0$) and γ ($\gamma < 1$), and even for values of $\eta \neq 1/2$ ($0 \leq \eta \leq 1$), that have been considered.

From our numerics, we have heuristically obtained (see Fig. 8) that for all the values of x for which we have checked, and even for values of $\eta \neq 1/2$, the q^{ldl} index satisfies

$$q^{ldl}(\eta, \alpha) = 1 + \frac{1}{\eta\alpha} . \tag{19}$$

Our numerics show that q^{ldl} does not depend on (γ, x). Consequently, from (9) and (19), we infer that, for $\eta = 1/2$ and for all values of $\alpha > 0$, the $q^{ldl}(\alpha)$-exponential decay index and the $q^{att}(\alpha)$-Gaussian attractor index are related as follows:

$$\left.\frac{1}{q^{att}(\alpha) - 1} + \frac{1}{q^{ldl}(\alpha) - 1}\right|_{\eta=1/2} = 1. \tag{20}$$

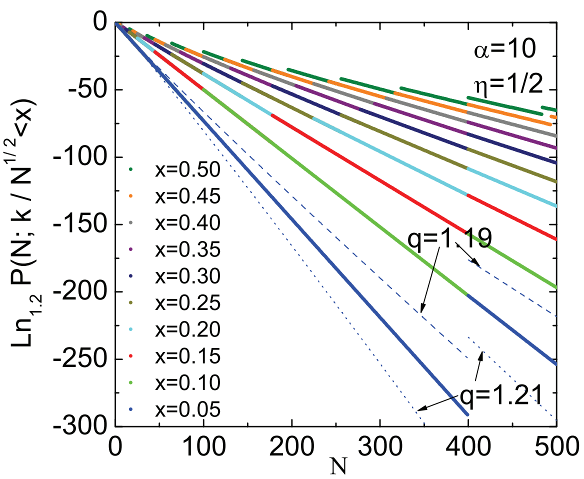

FIG. 7. $P(N; k/N^\gamma < x)$ distribution, for $\alpha = 10$, $\gamma = 1/2$, and $\eta = 1/2$, in semi-q^{ldl}-logarithmic representation. The index of the q^{ldl}-exponential decay is $q^{ldl} = 1.2$, for $\alpha = 10$, no matter the value of γ. Observe the significative deviations from a straight line, for 1% deviations of q^{ldl}.

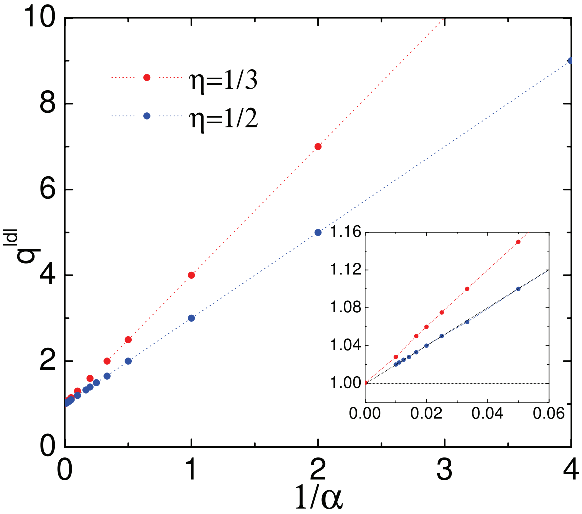

FIG. 8. The α-dependence of the q^{ldl}-exponential index of the decaying probability $P(N;k/N^{1/2} < x)$, for $\eta = 1/2$ and $\eta = 1/3$ models.

Let us now focus on the values of the slopes of the semi-q^{ldl}-logarithmic representation of $P(N;k/N \le x)$. Fig. 7 shows that the q^{ldl}-exponential decaying rate does not only depend on the value of x, i.e., $r_{q^{ldl}} = r_{q^{ldl}}(x,N,\gamma;\eta,\alpha)$. The mechanism that precludes a simple dependence on x can be understood by fixing a particular value of x and γ, as shown in Fig. 9. For a fixed value of x, the sequence of slope values is associated to the maximum value of k involved in $P(N;k/N^\gamma < x)$, i.e., $k_{max}(x,N,\gamma) = \lfloor N^\gamma x \rfloor$.

Summarizing, the deviation re-scaled probability behavior presents a q^{ldl}-exponential decay when $\gamma < 1$ that can be written as

$$P(N;k/N^\gamma < x;\eta,\alpha) = \sum_{k=0}^{\lfloor N^\gamma x \rfloor} \mathfrak{p}_k^{(N)}(\eta,\alpha) \simeq e_{q^{ldl}=1+\frac{1}{\eta\alpha}}^{-Nr(\lfloor N^\gamma x \rfloor;\eta,\alpha)}, \tag{21}$$

which, interesting enough, exhibits a non-trivial dependence of the rate function $r_{q^{ldl}}(\lfloor N^\gamma x \rfloor;\, \eta,\alpha)$.

V. LARGE DEVIATION PROBABILITY WITH RESPECT TO THE $N \to \infty$ LIMIT DISTRIBUTION

Let us now analyze the $N \to \infty$ evolution behavior of the probability of k/N^γ for $\gamma = 1$, i.e., how $p(k/N)$ approaches its attractor $p_\infty(x)$. The probability left deviations of k/N from x $(0 \le x \le 1/2)$, $P(N;k/N < x)$, with respect to the attractor, can be written as

$$\mathbb{P}_k^{(N)}(x;\eta,\alpha) \equiv P\,(k/N \le x;\eta,\alpha) - P^{att}\,(x;\eta,\alpha)\,, \tag{22}$$

where $P^{att}\,(x;\eta,\alpha) \equiv \lim_{N\to\infty} P\,(k/N \le x;\eta,\alpha) = \int_0^x p_\infty(z;\eta,\alpha)dz$.

From Eq. (6), we obtain that the probability left deviations of k/N from x, for $\eta = 1/2$ and even values of $\alpha \ge 4$ (i.e. $4 \le \alpha = \dot{2}$), can be written as

$$P\,(N;k/N \le x;\alpha) = \sum_{k=0}^{\lfloor Nx \rfloor} \mathfrak{p}_k^{(N)}(1/2,\alpha) = \mathfrak{p}_0^{(N)}(1/2,\alpha) + \frac{N!}{(\alpha)_N(\alpha/2-1)!^2} \sum_{k=1}^{\lfloor Nx \rfloor} \prod_{j=1}^{\alpha/2-1} (k+j)(N-k+j)\,. \tag{23}$$

Let us analytically study the simplest model, i.e., $\eta = 1/2$ and $\alpha = 4$. We verify

$$\begin{aligned} P\,(N;k/N \le x) &= \frac{6}{(N+2)(N+3)} + \frac{N!}{(4)_N} \sum_{k=1}^{\lfloor Nx \rfloor} (k+1)(N-k+1) \\ &= \frac{6(N+1) + \lfloor Nx \rfloor(5+9N+3\lfloor Nx \rfloor(N-1) - 2(\lfloor Nx \rfloor)^2}{(3+N)(2+N)(1+N)}. \end{aligned} \tag{24}$$

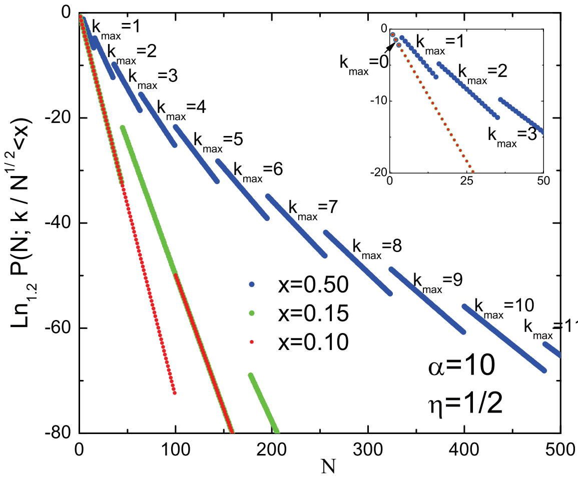

FIG. 9. q^{ldl}-exponential decaying rate of $P(N;k/N^{1/2} < x)$ for $\eta = 1/2$ and $\alpha = 10$ shows that the sequence of slopes that correspond to a value of x are associated to $k_{max} \equiv \max\{\lfloor N^{\gamma}x \rfloor\}$ involved in $P(N;k/N^{\gamma} < x)$.

The corresponding $N \to \infty$ limit distribution is a q^{att}-Gaussian, with $q^{att} = 0$. Consequently, the corresponding asymptotic left deviation of k/N from a fixed value x is given by

$$P^{att}(x) = \int_0^x p_\infty(z)dz = \frac{2\,\Gamma\left(\frac{5}{2}\right)}{\sqrt{\pi}\,\Gamma(2)} \int_0^x \left[1 - 4\left(X - \frac{1}{2}\right)^2\right] dX = 3x^2 - 2x^3. \tag{25}$$

From (24) and (25), we have that, for a certain value of N, the probability left deviations of k/N from x, with respect to the probability left deviation of the $N \to \infty$ limit distribution, is analytically obtained as

$$\mathbb{P}_k^{(N)}(x) = \frac{6(N+1) + \lfloor Nx \rfloor(5 + 9N + 3\lfloor Nx \rfloor(N-1) - 2(\lfloor Nx \rfloor)^2}{(3+N)(2+N)(1+N)} - 3x^2 + 2x^3. \tag{26}$$

The upper bound $\Delta_N^{upp}(x)$ of (26) can be obtained in the case that $\lfloor Nx \rfloor = Nx$. A lower bound $\Delta_N^{low}(x)$ can be also considered as $Nx - 1 < \lfloor Nx \rfloor$, but it is never attained and $\Delta_N^{low}(x) < \mathbb{P}_k^{(N)}(x;\eta,\alpha)$. We can consider, instead of $\Delta_N^{low}(x)$, the maximum of all lower bounds $\Delta_N^{min}(x)$, for each value of x. All these bounds verify the relation

$$\Delta_N^{low}(x) < \Delta_N^{min}(x) \le \mathbb{P}_k^{(N)}(x;\eta = \frac{1}{2}, \alpha = 4) \le \Delta_N^{upp}(x). \tag{27}$$

Fig. 10 exhibits the $(q = 2)$-logarithmic representation of the upper and the minimum bound deviations, $\Delta_N^{upp}(x)$ and $\Delta_N^{min}(x)$. $\Delta_N^{upp}(x)$ appears to q^{LD}-exponentially decay with N (where LD stands for *Large Deviation*), as conjectured in Ref. 23. The hypothesis of $\Delta_N^{upp}(x) q^{LD}$-exponentially decaying behavior can be verified by using the asymptotic expansion of the analytical expressions of the bounding values $\Delta_N^{upp}(x)$, $\forall x$,

$$\Delta_N^{upp}(x) = \frac{3x(x-1)(4x-3)}{N}\left[1 - \frac{50x^2 - 43x + 6}{3x(4x-3)N} + \frac{5(3x-2)(4x-1)}{x(4x-3)N^2} + \cdots\right]. \tag{28}$$

We can compare the respective terms with the asymptotic expansion of a q-exponential function,[23] namely,

$$a(x)e_q^{-r_q(x)N} = \frac{a(x)}{[(q-1)r_q(x)N]^{\frac{1}{q-1}}}$$
$$\times \left[1 - \frac{1}{(q-1)^2 r_q(x)N} + \sum_{m=2}^{\infty}(-1)^m \frac{q(2q-1)\dots[(m-1)q-(m-2)]}{m!(q-1)^{2m}(r_q(x)N)^m}\right]. \tag{29}$$

Equations (28) and (29) would provide, by neglecting higher-order terms, an index $q^{LD} = 2$. In such a situation, and identifying the two first terms of expansions, the best q^{LD}-generalized rate

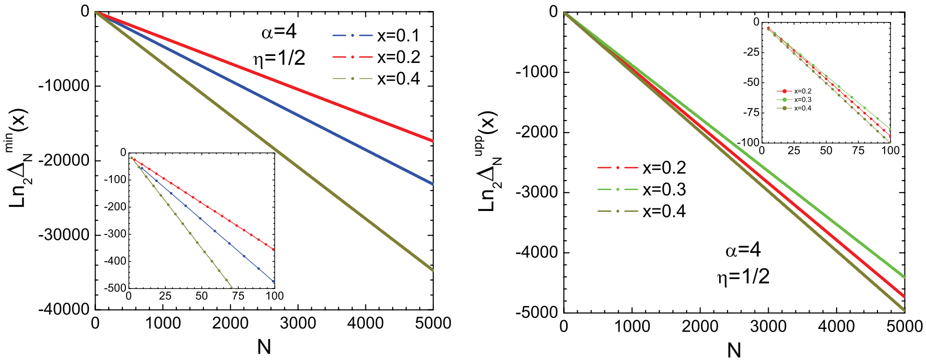

FIG. 10. Left panel: Semi-($q^{LD}=2$)-logarithmic representation of $\Delta_N^{min}(x)$. Observe that, for $x=0.1$, a bias from the linear behavior exists. Right panel: The linear behavior of the ($q=2$)-logarithmic representation of the upper bound $\Delta_N^{upp}(x)$ could reflect its q^{LD}-exponential decay with N.

function and the corresponding q^{LD}-exponential factor $a(x)$ would be

$$r_q^{upp} = \frac{3x(4x-3)}{50x^2-43x+6}, \qquad a^{upp}(x) = \frac{9x^2(x-1)(4x-3)^2}{50x^2-43x+6}. \tag{30}$$

But, in fact, the third term of Eq. (28) is not negligible for some values of x and, in such cases, Eqs. (29) and (30) are not compatible. The q^{LD}-exponential decay of the upper bound of the large deviation to the attractor is precluded, as obtained within a different context in Ref. 24.

Other values of α have been tested and, in all cases, the large deviation probability with respect the attractor presents a power-law decay. We conclude that Eq. (29) roughly describes the large-deviation behavior but, in some region of values of x, higher order terms of Eq. (29) are not negligible, which is not consistent with a q-exponential decay of the large-deviation probability to the attractor.

VI. CONCLUSIONS

A generalized binomial distribution based on q-exponential generating functions is characterized by Eqs. (5) and (6). We know that the corresponding probability function, $\mathfrak{p}_k^{(N)}(\eta,\alpha)$, can be written as a function of the parameters involved in the Pólya urn model, i.e., a sequential extraction of one ball from an urn that contains a set of b black balls and r red balls, followed by the successive reposition of c balls of the same color extracted.

If no bias exists, i.e., for $\eta = 1/2$, the probability to find a relative number of black balls k/N after the N-th trial, closely approaches a $q^{disc}(N)$-Gaussian distribution. The $N\to\infty$ limit probability distribution is in fact a q^{att}-Gaussian[18] whose q^{att} index and β^{-1} generalized temperature can be obtained from (b,r,c) as $q^{att} = \frac{b+r-4c}{b+r-2c}$ and $\beta^{-1} = \frac{c}{2(b+r-2c)}$. In other words, the numerical discrete distributions appear to be very close to a set of $q^{disc}(N)$-Gaussian distributions that, increasing N, evolve towards to a q^{att}-Gaussian attractor. This urn scheme provides a procedure to attain a q^{att}-Gaussian $N\to\infty$ limit distribution that verifies, in all cases ($\forall a,b,c$), the relation $\beta(1-q^{att}) = 4$. When the number of reposition balls verifies $c < (b+r)/2$, such attractors are concave q^{att}-Gaussian distributions with a bounded and compact support (where index $q^{att} < 1$ and generalized temperature $\beta^{-1} > 0$); on the contrary, when $c > (b+r)/2$, such attractors are convex q^{att}-Gaussian distributions with bounded but noncompact support (where index $q^{att} > 2$ and generalized temperature $\beta^{-1} < 0$).

These generalized binomial distributions violate the law of large numbers, but nevertheless present a large-deviation-like property. Indeed, by using, instead of the variable k/N, the rescaled variable k/N^γ ($\gamma < 1$), the left deviation probability behaves as $P(N;k/N^\gamma < x;\eta,\alpha) \simeq e_{q^{ldl}}^{-Nr(\lfloor N^\gamma x\rfloor;\eta,\alpha)}$. Moreover, an interesting result is that when no bias exists (i.e., $\eta = 1/2$), the q^{att}-Gaussian index, the q^{ldl} index, and the q^{gen} index (characterizing the generating function) are univocally defined by the $(b+r)/c$ ratio, and simple mathematical relations exist that remind the algebra indicated in Ref. 11.

ACKNOWLEDGMENTS

We acknowledge useful conversations with J. P. Gazeau and E. M. F. Curado, and partial financial support from Universidad Politécnica de Madrid, Centro Brasileiro de Pesquisas Físicas, CNPq, Faperj, and Capes (Brazilian agencies). One of us (G.R.) acknowledges warm hospitality at CBPF. The other one (C.T.) acknowledges partial financial support from the John Templeton Foundation.

[1] V. V. Dodonov, J. Opt. B: Quantum Semiclassical Opt. **4**(1), R1-R33 (2002), and references therein.
[2] J.-P. Gazeau, *Coherent States in Quantum Physics* (Wiley-VCH, 2009).
[3] R. L. de Matos Filho and W. Vogel, Phys Rev. A **54**, 4560 (1996).
[4] Z. Kis, W. Vogel, and L. Davidovich, Phys Rev. A **64**(3), 033401 (2001).
[5] E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrigues, Phys. Scr. **82**, 038108 (2010).
[6] H. Bergeron, E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrigues, J. Math. Phys. **53**, 103304 (2012).
[7] H. Bergeron, E. M. F. Curado, J. P. Gazeau, and Ligia M. C. S. Rodrigues, J. Math. Phys. **54**, 123301 (2013).
[8] C. Tsallis, M. Gell-Mann, and Y. Sato, Proc. Natl. Acad. Sci. U. S. A. **102**, 15377 (2005).
[9] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[10] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics - Approaching a Complex World* (Springer, New York, 2009).
[11] S. Umarov, C. Tsallis, M. Gell-Mann, and S. Steinberg, J. Math. Phys. **51**, 033502 (2010).
[12] A. Rodriguez and C. Tsallis, J. Stat. Mech.: Theory Exp. **2014**, P12027.
[13] S. Umarov and C. Tsallis, AIP Conf. Proc. **965**, 34-42 (2007).
[14] D. Prato and C. Tsallis, Phys. Rev. E **60**, 2398 (1999).
[15] R. Hanel, S. Thurner, and C. Tsallis, Eur. Phys. J. B **72**, 263 (2009).
[16] A. Rodriguez and C. Tsallis, J. Math. Phys. **53**, 023302 (2012).
[17] J. Ruseckas, Phys. Lett. A **379**, 654 (2015).
[18] H. Bergeron, E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrigues, preprint arXiv:1412.0581 [cond-mat.stat-mech] (2014).
[19] A. Carati, Physica A **348**, 110 (2005).
[20] A. Carati, Physica A **387**, 1491 (2008).
[21] See http://www.encyclopediaofmath.org/ for Pólya distribution, Encyclopedia of Mathematics.
[22] T. D. Frank, *Nonlinear Fokker-Planck Equations: Fundamentals and Applications* (Springer, Berlin - New York, 2005).
[23] G. Ruiz and C. Tsallis, Phys. Lett. A **377**, 491 (2013).
[24] M. Jauregui and C. Tsallis, J. Math. Phys. **56**, 023303 (2015).