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

Guiomar Ruiz$^{1,2}$ and Constantino Tsallis$^{1,3}$

$^1$Centro Brasileiro de Pesquisas Físicas and National Institute of Science and Technology for Complex Systems, Rua Xavier Sigaud 150, 2290-180 Rio de Janeiro-RJ, Brazil
$^2$Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Madrid, Pza. Cardenal Osorros s/n, 28040 Madrid, Spain
$^3$Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

We study a symmetric generalization $p_k^{(N)}(\eta, \alpha)$ of the binomial distribution recently introduced by Bergeron et al, where $\eta \in [0,1]$ denotes the win probability, and $\alpha$ is a positive parameter. This generalization is based on $q$-exponential generating functions ($e_{q^{gen}}^{\eta \alpha} \equiv [1+(1-q^{gen})z]^{1/(1-q^{gen})}; e^\eta = e^z$) 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 $\alpha$. Asymptotic $N \to \infty$ distribution is in fact a $q^{att}$-Gaussian $e^{-\beta z^2}$, where $q^{att} = 1 - 2/(\alpha - 2)$ and $\beta = (2\alpha - 4)$. The behavior of the scaled quantity $k/N^q$ 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 $q^{ldl}$-exponential decay. All $q$-statistical parameters which emerge are univocally defined by $(\eta, \alpha)$. Finally we discuss the analytical connection with the Pólya urn problem.

PACS numbers: 02.50.-r, 05.20.-y, 65.40.gd

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 $p_k^{(N)}$ 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 ($e_{q^{gen}}^{\eta \alpha} \equiv [1+(1-q^{gen})z]^{1/(1-q^{gen})}; e^\eta = e^z$), where gen stands for generating, and $q^{gen} = 1 + 1/\alpha > 1 (\alpha > 0)$, $\alpha$ 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 [7],[8]). The family of the generated probabilities depends on two parameters $(\eta, \alpha)$, where $\eta$ is the “win” probability, $(1-\eta)$ is the “loss” probability, and $\alpha$ characterizes the generating function. A variety of $q$-statistical functions [9],[10] related to the probabilities $p_k^{(N)}(\eta, \alpha)$ emerges, all of them univocally defined by $(\eta, \alpha)$, and whose respective $q(\eta, \alpha)$ indices appear to obey an algebra that reminds the underlying algebra in [8],[11].

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 any of the complete set of $12$ of bounded support $q^{disc}$-Gaussian $e^{-\beta z^2}$ (where disc stands for discrete, and $q^{disc} < 1$), where $\beta$ is a generalized inverse temperature ($\beta \in \mathbb{R}$). The $q$-Gaussian form corresponds to strongly correlated random variables, and arises from the extremization of the nonadditive entropy $S_q = k_B (1 - \sum_i p_i^q)/(q-1)$ ($q \in \mathbb{R}$, $S_1 = S_{BG} = -k \sum_i p_i \ln p_i$, where $BG$ stands for Boltzmann-Gibbs) [9] under
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:

\[
p_k^{(N)}(\eta) = \binom{N}{k} \eta^k (1-\eta)^{N-k} = \frac{N!}{(N-k)! k!} \eta^k (1-\eta)^{N-k}
\]

where the parameter \( \eta \) (\( 0 \leq \eta \leq 1 \)) 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 an strictly increasing infinite sequence of nonnegative real numbers \( \chi = \{x_N\}_{N \in \mathbb{N}} \). With each sequence \( \chi \) defined above, a Bernoulli-like distribution is constructed:

\[
p_k^{(N)}(\eta) = \frac{x_N!}{x_{N-k}! x_k!} q_k(\eta) q_{N-k}(1-\eta)
\]

where the factorials are defined as \( x_N! = x_1 x_2 \ldots x_N \), \( x_0! = 1 \), \( \eta \) 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 \( \binom{N}{k} \) is preserved, as invariance under \([k,\eta] \rightarrow [(N-k),(1-\eta)]\) 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, is such a way that quantities \( p_k^{(N)}(\eta) \) represent the probabilities of having \( k \) wins and \((N-k)\) losses in a sequence of correlated \( N \) trials. Consequently, \( p_k^{(N)}(\eta) \) must be constrained by the normalization equation:

\[
\sum_{k=0}^{N} p_k^{(N)}(\eta) = 1, \quad \forall N \in \mathbb{N}, \forall \eta \in [0,1],
\]

and the non-negativeness condition:

\[
p_k^{(N)}(\eta) \geq 0, \quad \forall N, k \in \mathbb{N}, \forall \eta \in [0,1].
\]

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 \( q^{\text{gen}} \)-exponential generating function \( e_{q^{\text{gen}}}^{\alpha} = [1 + (1-q^{\text{gen}})z]^{1/(1-q^{\text{gen}})} \).

Such a generating function can be written as \( N(z) = (1-z/\alpha)^{-\alpha} \) \((\alpha > 0)\), where the \( q \)-exponential parameter that characterizes the generating function is \( q^{\text{gen}} = 1 + 1/\alpha > 1 \). The following probability distributions are consequently obtained \([7]\):

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

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

\[
p_k^{(N)}(\eta = 1/2, \alpha) = \begin{cases} 
\frac{1}{N!} (\alpha/2)_N & (k = 0, N) \\
\frac{1}{N!} \prod_{m=1}^{\alpha-1} \frac{\alpha-m}{N+\alpha-m} \times \left[ \prod_{j=1}^{q-1} \frac{k+j}{j} \left( \frac{N-k+j}{j} \right) \right] & (k \neq 0, N). 
\end{cases}
\] (6)

This expression is computationally very convenient.

### III. EMERGENCE OF \(q\)-GAUSSIANS

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

Fig. 2 illustrates, for \(\eta = 1/2\) and a typical choice of \(\alpha\) and \(N\), the distribution \(p(k/N)\) normalized to its maximum \(P_{\text{max}}\). Our numerical results strongly suggest that \(p(k/N)\) can be fitted by a \(q^{\text{disc}}\)-Gaussian distribution with \(q^{\text{disc}} = 0.84508\). Other values of \(\alpha\) and \(N\) have been studied as well and, in all cases, numerical results strongly suggest \(q^{\text{disc}}\)-Gaussian distributions (with \(q^{\text{disc}} < 1\), i.e.,

\[
p(k/N)/P_{\text{max}} \simeq e^{-\frac{\beta q^{\text{disc}}}{2}(\frac{k}{N} - 1)^2} \left[ 1 - \beta(1-q^{\text{disc}})\left( \frac{k}{N} - 1 \right)^2 \right]^{1-q^{\text{disc}}} (7)
\]

where \([x]_+ = x \text{ if } x > 0\), and \([x]_+ = 0\) otherwise. Table 1 shows the values of \(q^{\text{disc}}\) for typical values of \(\alpha\) and \(N\) (\(\eta = 1/2\) in all cases). For a fixed value of \(N\), \(q^{\text{disc}}\) is a monotonous function of \(\alpha\). Similarly, for a fixed value of \(\alpha\), \(q^{\text{disc}}\) 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 \(q^{\text{disc}}(N)\)-logarithmic representation show that \(q^{\text{disc}}\)-Gaussian distributions fit very well the data.
FIG. 2: Normalized symmetric generalized binomial distribution for $\eta = 1/2$, $N = 100$ and $\alpha = 15$. **Left panel:** The corresponding $q^{\text{disc}}$-Gaussian fitting function is superimposed. **Right panel:** The $q^{\text{disc}}$-logarithmic representation exhibits that indeed the discrete (i.e., $N < \infty$) distribution is extremely close to a $q$-Gaussian, the linear regression coefficient being $R = 1$.

| $\alpha$ | $N = 50$ | $N = 60$ | $N = 70$ | $N = 80$ | $N = 100$ | $N = 200$ | $N = 500$ | $N = 1000$ | $N \to \infty$ |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 3       | -0.98990 | -0.99170 | -0.99306 | -0.99410 | -0.99524 | -0.99800 | -0.99930 | -0.99970 | -1       |
| 5       | 0.33117  | 0.33185  | 0.33214  | 0.33233  | 0.33264  | 0.33311  | 0.33324  | 0.33328  | 1/3      |
| 15      | 0.84235  | 0.84350  | 0.84412  | 0.84456  | 0.84508  | 0.84587  | 0.84611  | 0.84614  | 11/13    |
| 25      | 0.90803  | 0.90939  | 0.91027  | 0.91084  | 0.91154  | 0.91263  | 0.91297  | 0.91303  | 21/23    |
| 50      | 0.95127  | 0.95298  | 0.95413  | 0.95494  | 0.95598  | 0.95763  | 0.95821  | 0.95830  | 23/24    |
| 100     | 0.97043  | 0.97244  | 0.97382  | 0.97483  | 0.97616  | 0.97845  | 0.97936  | 0.97953  | 48/49    |
| 500     | 0.98375  | 0.98576  | 0.98874  | 0.99039  | 0.99359  | 0.99531  | 0.99576  | 0.99699  | 248/249  |
| $\alpha \to \infty$ | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         | 1         |

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

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

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

and the corresponding $q^{\text{att}}$-Gaussian ($N \to \infty$) limit distributions are characterized by $(q^{\text{att}}, \beta)$. This result that can in fact be rigorously proved [18]. Fig. 5 shows the $N \to \infty$ convergence of the index $q^{\text{disc}}$ to the index $q^{\text{att}}$.

The value of $q^{\text{att}}$ depends only on the parameter $\alpha$, and some values of the $(q^{\text{att}}, \alpha)$ pair are shown in Table II. We heuristically conclude (see Fig. 5) that

$$q^{\text{att}}(\alpha) = 1 - \frac{2}{\alpha - 2}.$$  

Fig. 6 shows the generalized temperature $\beta^{-1}$ of non-normalized $q^{\text{att}}$-Gaussian distribution, as they are numerically obtained from $\alpha$. The $\alpha > 2$ values correspond to positive values of $\beta$, and the $0 < \alpha < 2$ values correspond to

$$q^{\text{att}}(\alpha) = \lim_{N \to \infty} q^{\text{disc}}(\alpha, N)$$

that characterizes the $N \to \infty$ limit $q^{\text{att}}$-Gaussian distribution.

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

$$\begin{array}{ccccccccccc}
\alpha & 2 & 2.1 & 2.5 & 3 & 4 & 5 & 15 & 25 & 50 & 100 & 500 & 1000 \\
q^{\text{att}} & -\infty & -19 & -8 & -3 & -1 & 0 & 1/3 & 11/13 & 21/23 & 25/24 & 48/49 & 248/249 & 498/499 \\
\end{array}$$
FIG. 3: Left panels: Probability distribution \( p(k/N) = Np_k^{(N)} \) \((0 \leq k/N \leq 1)\) for typical values of \(\alpha\) and \(N\). Right panels: The corresponding \(q_{\text{disc}}^{(N)}\)-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 \(q_{\text{att}} > 0\) \((q_{\text{att}} < 0)\), the terminal derivative in the linear-linear representation vanishes (diverges). This derivative is finite for \(\alpha = 4\), which corresponds to \(q_{\text{att}} = 0\).

FIG. 4: Convergence of the \(q_{\text{disc}}\) index to the \(q_{\text{att}}\) index.
negative values of $\beta$. In both cases, the following equation is satisfied:

$$\beta^{-1} = \frac{1 - q^{att}}{4}. \quad (10)$$

This corresponds to the cut-offs of the distributions. We can infer, from (9) and (10), the following simple relation:

$$\beta(\alpha) = 2(\alpha - 2) \quad (\alpha > 0), \quad (11)$$

that can also be checked in Fig. 6 (right panel).

Summarizing, Fig. 6 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 non-compact 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 [12], where negative generalized temperatures $\beta^{-1}$ of $q$-Gaussian distributions are also considered.

The indexes $q^{att}$ and $q^{gen}$ are related by

$$\frac{1}{q^{gen} - 1} - \frac{2}{q^{att} - 1} = 2, \quad (12)$$

equation that reminds the equations of the algebra indicated in [11].

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
IV. LARGE-DEVIAITION-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 \( N^\gamma \) 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.,

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

where \( \epsilon \) 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

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

and conclude that, \( \forall \alpha > 0 \),

$$
\lim_{N \to \infty} P(k/N \leq x; \alpha) = \lim_{N \to \infty} \sum_{k=0}^{\lfloor Nx \rfloor} p_k^{(N)}(\alpha) = \int_{0 \leq x \leq 1/2} p_\infty(x; \alpha)dx \neq 0.
$$

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). Eq. (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 \( k/N^\gamma \), \( \gamma \neq 1 \), 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.
FIG. 8: Non-centered histograms of $k/N^\gamma$, $p(k/N^\gamma) = N^\gamma p_k^{(N)}$, for $\eta = 1/2$, $\alpha = 10$, and typical sequences of $N$. **Left panel:** Scaling factor $\gamma < 1$ makes $\lim_{N \to \infty} p(k/N^\gamma) = 0$, $\forall (k/N^\gamma) \in \mathbb{R}^+$. **Right panel:** Scaling factor $\gamma > 1$ makes $\lim_{N \to \infty} p(k/N^\gamma) = 0$.

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 $\eta$ and $\alpha$, when $N$ increases. In fact (see Fig. 8) it can be straightforwardly verified that

$$\lim_{N \to \infty} P(N; k/N^\gamma \leq x) = \lim_{N \to \infty} \sum_{k=0}^{\left\lfloor N^\gamma x \right\rfloor} p_k^{(N)} = 0,$$

and the zero limit is attained as

$$P(N; k/N^\gamma < x) \sim e^{-N r_{q_{ldl}}} \left[ 1 - (1 - q_{ldl}) N r_{q_{ldl}} \right]^{1/(1-q_{ldl})},$$

with $q_{ldl} > 1$. This result can be checked in Fig. 9 where $q_{disc}(N)$-logarithmic representation of $P(N; k/N^\gamma < x)$ as a function of $N$ is shown, for typical values of $\alpha$ and $\gamma$. Straight lines are obtained for all values of $\eta$ ($0 \leq \eta \leq 1$), $\alpha$ ($\alpha > 0$) and $\gamma$ ($\gamma < 1$) that have been considered.

We have heuristically obtained (see Fig. 10) that, for all the values of $x$ for which we have checked, the $q_{ldl}$ index satisfies

$$q_{ldl}(\eta, \alpha) = 1 + \frac{1}{\eta \alpha}.$$  \hspace{1cm} (19)

Notice that it does not depend on $\gamma$ or $x$ (see Fig. 11).

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:

$$\frac{1}{q_{att}(\alpha)} - 1 + \frac{1}{q_{ldl}(\alpha)} - 1 \bigg|_{\eta=1/2} = 1.$$  \hspace{1cm} (20)

Let us now focus on the values of the slopes of the semi-$q_{ldl}$-logarithmic representation of $P(N; k/N \leq x)$. Fig. 9 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 $\gamma$, as shown in Fig. 12. For a fixed value of $x$, the sequence of slope values are associated to the maximum value of $k$ involved in $P(N; k/N^\gamma < x)$, i.e., $k_{\text{max}}(x, N, \gamma) = \left\lfloor N^\gamma x \right\rfloor$. Summarizing, the deviation re-scaled probability behavior presents a $q_{ldl}$-exponential decay when $\gamma < 1$, that can be written as

$$P(N; k/N^\gamma < x; \eta, \alpha) = \sum_{k=0}^{\left\lfloor N^\gamma x \right\rfloor} p_k^{(N)}(\eta, \alpha) \approx e^{-N r_{q_{ldl}}(\left\lfloor N^\gamma x \right\rfloor; \eta, \alpha)},$$

thus exhibiting a non-trivial dependence of the rate function $r_{q_{ldl}}(\left\lfloor N^\gamma x \right\rfloor; \eta, \alpha)$.
FIG. 9: $P(N; k/N^\gamma < x)$ distribution, for typical values of $\alpha$ ($\eta = 1/2$), in semi-$q^{dl}$-logarithmic representation. The values of $q^{dl}(\alpha)$, related to the $q^{dl}$-exponential decay, are $q^{dl}(100) = 1.02$, $q^{dl}(10) = 1.2$ and $q^{dl}(5) = 1.4$, no matter the value of $\gamma$ ($\gamma = 0.4$, $\gamma = 0.5$ and $\gamma = 0.7$, in figure). The central panel shows the significative deviations from a straight line, for 1% deviations of $q^{dl}$.

FIG. 10: The $\alpha$-dependence of the $q^{dl}$-exponential index of the decaying probability $P(N; k/N^{1/2} < x)$, for $\eta = 1/2$ and $\eta = 1/3$ models.
FIG. 11: The $q^{\text{dim}}$-exponential indices appear to be independent of $\gamma$.

FIG. 12: $q^{\text{dim}}$-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 $P(N; k/N^{\gamma} < x)$.

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 $k/N^{\gamma}$ for $\gamma = 1$, i.e., how $p(k/N)$ approaches its attractor $p_{\infty}(x)$. The probability left deviations of $k/N$ from $x$ $(0 \leq x \leq 1/2)$, $P(N; k/N < x)$, with respect to the attractor, can be written as

$$P_k^N(x; \eta, \alpha) \equiv P(k/N \leq x; \eta, \alpha) - P_{\text{att}}(x; \eta, \alpha),$$

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

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

$$P(N; k/N \leq x; \alpha) = \sum_{k=0}^{\lfloor Nx \rfloor} p_k^N(1/2, \alpha) = p_0^N(1/2, \alpha) + \frac{N!}{(\alpha_N(\alpha/2 - 1))!} \sum_{k=1}^{\lfloor Nx \rfloor} \prod_{j=1}^{\alpha/2-1} (k+j)(N-k+j).$$

(23)
FIG. 13: Left Panel: Semi-($q^{LD} = 2$)-logarithmic representation of $\Delta_N^{\text{up}}(x)$. 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 $\Delta_N^{\text{up}}(x)$ could reflect its $q^{LD}$-exponential decay with $N$.

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

$$P(N; k/N \leq x) = \frac{6}{(N+2)(N+3)} + \frac{N!}{(4)^N} \sum_{k=1}^{[Nx]} (k+1)(N-k+1)$$

$$= \frac{6(N+1) + [Nx](5 + 9N + 3|N_x|(N-1) - 2(|N_x|)^2}{(3 + N)(2 + N)(1 + N)}.$$  \hspace{1cm} (24)

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

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

$$p_k^{(N)}(x) = \frac{6(N+1) + [Nx](5 + 9N + 3|N_x|(N-1) - 2(|N_x|)^2}{(3 + N)(2 + N)(1 + N)}.$$  \hspace{1cm} (26)

The upper bound $\Delta_N^{\text{up}}(x)$ of (26) can be obtained in the case that $|N_x| = N_x$, as shown in Fig. 14. A lower bound $\Delta_N^{\text{low}}(x)$ can be also considered as $N_x - 1 < |N_x|$, but it is never attained and $\Delta_N^{\text{low}}(x) < p_k^{(N)}(x; \eta, \alpha)$. We can consider, instead of $\Delta_N^{\text{low}}(x)$, the maximum of all lower bounds $\Delta_N^{\text{min}}(x)$, for each value of $x$. All these bounds verify the relation

$$\Delta_N^{\text{low}}(x) < \Delta_N^{\text{min}}(x) \leq p_k^{(N)}(x; \eta = \frac{1}{2}, \alpha = 4) \leq \Delta_N^{\text{up}}(x)$$  \hspace{1cm} (27)

FIG. 13 exhibits the ($q = 2$)-logarithmic representation of the upper and the minimum bound deviations, $\Delta_N^{\text{up}}(x)$ and $\Delta_N^{\text{min}}(x)$. $\Delta_N^{\text{up}}(x)$ appears to $q^{LD}$-exponentially decay with $N$ (were $LD$ stands for Large Deviation), as conjectured in [23]. It is not the case of $\Delta_N^{\text{min}}(x)$ for $x = 0.3$, as it is shown in Fig. 14.

The hypothesis of $\Delta_N^{\text{up}}(x)$ $q^{LD}$-exponentially decaying behavior can be verified by using the asymptotic expansion of the analytical expressions of the bounding values $\Delta_N^{\text{up}}(x)$, $\forall x$

$$\Delta_N^{\text{up}}(x) = \frac{3x(x - 1)(4x - 3)}{N} \left[ 1 - \frac{50x^2 - 43x + 6}{3x(4x - 3)N} + \frac{5(3x - 2)(4x - 1)}{x(4x - 3)N^2} + \ldots \right].$$  \hspace{1cm} (28)
We can compare the respective terms with the asymptotic expansion of a $q$-exponential function \[23\], namely
\[
a(x) = \frac{a(x)}{[(q - 1)r_q(x)N]^{1-q}} \times \left[1 - \frac{1}{(q - 1)^2r_q(x)N} + \sum_{m=2}^{\infty} \left(-1\right)^m q(2q - 1)\ldots (m - 1)q - (m - 2)\right]. \tag{29}
\]
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 function and the corresponding $q^{LD}$-exponential factor $a(x)$ would be
\[
\begin{align*}
p^{upper}_q &= \frac{3x(4x - 3)}{50x^2 - 43x + 6}, & a^{upper}(x) &= \frac{9x^2(x - 1)(4x - 3)^2}{50x^2 - 43x + 6}. \tag{30}
\end{align*}
\]
But, in fact, the third term of Eq. \[30\] is not negligible for some values of $x$ and, in such cases, Eq. \[29\] and Eq. \[30\] 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 \[24\].

Other values of $\alpha$ have been tested and, in all cases, the large deviation probability with respect the attractor presents a power-law decay. The large-deviation probability does not in fact $q$-exponentially decay to the attractor, even though, in some cases, Eq. \[29\] roughly describes the large-deviation behavior.

**VI. CONCLUSIONS**

A generalized binomial distribution based on $q$-exponential generating functions is characterized by Eqs. \[5\] and \[6\]. Its probability function $p_k^{(N)}(\eta, \alpha)$ depends on two parameters $(\eta, \alpha)$, and can be considered as 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, $p_k^{(N)}(\eta, \alpha)$ 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 $\alpha = (b + r)/c$.

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 $q$-Gaussian distribution. The $N \rightarrow \infty$ limit probability distribution is in fact a $q$-Gaussian whose $q^{\text{att}}$ index and $\beta^{-1}$ generalized temperature, can be obtained from $(b, r, c)$ as $q^{\text{att}} = \frac{b+r-4c}{b+r}$. In other words, the numerical discrete distributions appear to be very close to a set of $q\text{disc}(N)$-Gaussian distributions that, increasing $N$, evolve towards to a $q^{\text{att}}$-Gaussian attractor. This urn scheme provides a procedure to attain a $q^{\text{att}}$-Gaussian $N \rightarrow \infty$ limit distribution that verifies, in all cases ($\forall a, b, c$), the relation $\beta(1 - q^{\text{att}}) = 4$. When the number of reposition balls verifies $c < (b+r)/2$, such attractors are concave $q^{\text{att}}$-Gaussian distributions with a bounded support; when $c > (b+r)/2$ such attractors are convex $q^{\text{att}}$-Gaussian distributions with bounded support.

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 $k/N^\gamma$ ($\gamma < 1$), the left deviation probability behaves as $P(N; k/N^\gamma < x; \eta, \alpha) \simeq e^{-N\gamma x/(N^\gamma x; \eta, \alpha)}$. Moreover, an interesting result is that, when no bias exists (i.e., $\eta = 1/2$), the $q^{\text{att}}$-Gaussian index, the $q^{\text{att}}$ index, and the $q^{\text{gen}}$ 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 [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 (GR) acknowledges warm hospitality at CBPF. The other one (CT) acknowledges partial financial support from the John Templeton Foundation.

[1] V. V. Dodonov, J. Opt. B: Quantum Semiclass. Opt. 4, 1 (2002), and references therein.
[2] J.-P. Gazeau, Coherent states in quantum physics, Wiley-VCH, 2009.
[3] R. de Matos Filho and W. Vogel Phys Rev. A 54, 4560 (1996).
[4] Z. Kis, W. Vogel, and L. Davidovich, Phys Rev. A 64, 003401 (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, L. M. C. S. Rodrigues, J. Math. Phys. 146, 103304-1-22 (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. Sc. USA 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, private communication.
[13] S. Umarov and C. Tsallis, American Institute of Physics Conference Proceedings 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. Rodrigues and C. Tsallis, J. Math. Phys. 53, 023302 (2012).
[17] J. Ruseckas, [cond-mat.stat-mech] 1408.0088v1.
[18] H. Bergeron, E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrigues, private communication.
[19] A. Carati, Physica A 348, 110 (2005).
[20] A. Carati, Physica A 387, 1491 (2008).
[21] Pólya distribution, Encyclopedia of Mathematics: [http://www.encyclopediaofmath.org/](http://www.encyclopediaofmath.org/)
[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] J. Max and C. Tsallis, private communication.