Approaching a large deviation theory for complex systems

Ugur Tirnakli · Constantino Tsallis · Nihat Ay

Abstract The standard large deviation theory (LDT) is mathematically illustrated by the Boltzmann–Gibbs factor which describes the thermal equilibrium of short-range-interacting many-body Hamiltonian systems, the velocity distribution of which is Maxwellian. It is generically applicable to systems satisfying the central limit theorem (CLT). When we focus instead on stationary states of typical complex systems (e.g., classical long-range-interacting many-body Hamiltonian systems, such as self-gravitating ones), the CLT, and possibly also the LDT, need to be generalized. Specifically, when the \( N \to \infty \) attractor (\( N \) being the number of degrees of freedom) in the space of distributions is a \( Q \)-Gaussian (a nonadditive \( q \)-entropy-based generalization of the standard Gaussian case, which is recovered for \( Q = 1 \)) related to a \( Q \)-generalized CLT, we expect the LDT probability distribution to asymptotically approach a power law. Consistently with available strong numerical indications for probabilistic models, this behavior possibly is that associated with a \( q \)-exponential (defined as \( e_q^x \equiv [1 + (1 - q)x]^{1/(1-q)} \), which is the generalization of the standard exponential form, straightforwardly recovered for \( q = 1 \); \( q \) and \( Q \) are expected to be simply connected, including the particular case \( q = Q = 1 \). The argument of such \( q \)-exponential would be expected to be proportional to \( N \), analogously to the thermodynamic entropy of many-body Hamiltonian systems. We provide here numerical evidence supporting the asymptotic power law by analyzing the standard map, the coherent noise model for biological extinctions and earthquakes, the Ehrenfest dog-flea model, and the random walk avalanches. For the particular case of the strongly chaotic standard map, we numerically verify (below 5% error bar) the validity of the asymptotic exponential behavior predicted by the usual LDT once the initial transient elapses typically beyond \( N \simeq 3 \times 10^6 \). Analogously, for the standard map with vanishing Lyapunov exponent, we provide numerical evidence (below the same error bar) for the asymptotic validity of the \( q \)-exponential behavior once the initial transient elapses typically beyond \( N \simeq 2 \times 10^5 \).
Keywords Nonlinear dynamical systems · Central limit theorem · Large deviation theory · Nonadditive entropies · Model systems

1 Introduction

Boltzmann–Gibbs (BG) statistical mechanics yields various important relations. Still, it is fair to consider as its most important fingerprints the Maxwellian distribution of velocities and the exponential distribution of energies (BG weight or BG factor) [1, 2]. These behaviors correspond mathematically to the central limit theorem (CLT) [3, 4] which leads, when the number $N$ of involved random variables increases indefinitely, to convergence toward Gaussian distributions, and to the Large Deviation Theory (LDT) [5–7] which characterizes the speed at which Gaussians are approached while $N$ increases. To be more precise, the BG distribution $p_{BG}$ associated with a many-body Hamiltonian $\mathcal{H}_N$ at thermal equilibrium is given by $p_{BG} \propto e^{-\beta \mathcal{H}_N}$ whenever $\mathcal{H}_N$ includes short-range interactions or no interactions at all. We may then write that $p_{BG} \propto e^{-[\beta \mathcal{H}_N / N]}$, where, consistently with thermodynamics, $[\beta \mathcal{H}_N / N]$ is an intensive quantity. The corresponding LDT statement concerns the probability $P_N(Y_N / N > z) \approx 0, 1$ of the random variable $Y_N / N$ taking values larger than a fixed value $z \in \mathbb{R}$ for increasingly large values of $N$. Under the hypothesis of probabilistic independence, or similar settings, we expect $P_N(Y_N / N > z) \approx e^{-r_1(z)N}$, where the rate function $r_1$ equals a relative entropy per particle. Therefore, $r_1(z)N$ plays the role of the thermodynamic total entropy which, consistently with the Legendre structure of classical thermodynamics, is extensive, i.e., $r_1(z)N \propto N (N \gg 1)$.

Here, we focus on systems with nonlocal space-time correlations by generalizing the BG theory [8, 9]. The basis of this generalization consists in optimizing entropies which differ from $S_{BG} = -k \sum_i p_i \ln p_i$, such as $S_q = k \frac{1}{q-1} \sum_i p_i^q$ with $q \in \mathbb{R}$, and $S_1 = S_{BG}$ (see [10] for an interesting discussion of the admissibility of such generalization on statistical inference grounds). If $A$ and $B$ are two probabilistically independent systems, we straightforwardly verify that $S_q(A + B) / k = S_q(A) / k + S_q(B) / k + (1 - q)[S_q(A) / k][S_q(B) / k]$. In other words, $S_q$ is nonadditive for $q \neq 1$ whereas $S_{BG}$ is additive. The optimization of $S_q$ with simple constraints yields a probability distribution $p_q \propto e^{-\beta_q \mathcal{H}_N}$, where $\beta_q$ plays the role of an inverse temperature and $e^{\mathcal{E}} \equiv [1 + (1 - q)z]^q_{+}$ with $e^1 = e^z$ and $[\ldots]_+ = [\ldots]_0$ if $[\ldots]_0 > 0$, and zero otherwise (whose inverse function is the $q$-logarithm defined as $\ln_q z \equiv (z^{1-q} - 1) / (1-q)$). The form $e^{-\alpha z^a} (a > 0)$ is usually referred to as $q$-Gaussian. This $q$-generalized statistical mechanics (nonextensive statistical mechanics or $q$-statistics for short) typically tackle with long-range-interacting Hamiltonian systems, e.g., self-gravitating systems [11], violently relaxing systems [12], ionic crystals [13, 14], among other nontrivial systems such as cold atoms in dissipative lattices [15, 16], granular matter [17], high-energy collisions of elementary particles [18], overdamped systems like type-II superconductors [19], matter–antimatter astrophysical observations [20], stellar physics [21], theory of finances [22], complex networks [23]. The associated distributions of velocities appear to typically be $Q$-Gaussians with $Q > 1$ (see, for instance, [24, 25] for the $Q$-XY ferromagnet, [26] for the $Q$-Heisenberg ferromagnet, and [27, 28] for the $Q$-Fermi–Pastà–Ulam model), with $Q$ approaching unity when the range of the interactions approaches the short-range regime. This corresponds mathematically to a $Q$-generalized central limit theorem ($Q$-CLT) which leads, when the number $N$ of strongly correlated random variables increases to infinity, to a convergence on a $Q$-Gaussian distribution. Sufficient conditions for the $Q$-CLT to hold are already available [29] (see also [30, 31]) but the necessary conditions for a $Q$-CLT remain as a challenge. A $Q$-generalized approach of long-range-interacting Hamiltonian systems can be seen in [32]. However, the fact that various thermostatistical quantities are computable within the BG theory by no means guarantees that the theory correctly handles the many-body dynamics of those systems at any experimentally accessible time for any experimentally accessible size of the system. Analytical computability is necessary but not sufficient. Indeed, in the various anomalous systems listed above, the $Q$-CLT poorly fits reality [13–28].

Within $q$-statistics we have, for the total energy of the system at its stationary, or quasi-stationary, state, $p_q \propto e^{-\beta_q \mathcal{H}_N}$, with $\mathcal{H}_N$ being super-extensive, i.e., not proportional to $N$, consistently with long-range interactions. For say two-body (attractive) interactions decaying like $1 / (distance)^{\alpha}$ ($\alpha \in [0, \infty)$) within a $d$-dimensional system, we may rewrite $p_q \propto e^{-\beta_q \mathcal{H}_N}$.
where $\bar{N} \equiv \frac{N^{1 - \alpha/d} - 1}{1 - \alpha/d}$ is, for $N$ increasingly large, constant for $\alpha/d > 1$ (short range), increases like $N^{1 - \alpha/d}$ for $0 \leq \alpha/d < 1$ (long range), and increases like $\ln N$ for $\alpha/d = 1$. Notice that both $(\beta_q \bar{N})$ and $(\mathcal{H}_N/N \bar{N})$ are intensive quantities (see details in [33] and references therein). The desirable mathematical correspondence would of course be to have a $q$-Large Deviation Theory ($q$-LDT) with a probability $P(N, Y/N > \lambda) \approx e^{-r_q(z)N}$, where the rate function $r_q$ would once again equal some sort of relative nonadditive entropy $S_q$ per particle. Therefore, $r_q(z)N$ is expected to play the role of the total system thermodynamic entropy which, as before, should be extensive, i.e., $\propto N$ ($N \gg 1$). Naturally, in order to unify all the above situations, we expect $q = f(Q)$, $f(Q)$ being a smooth function which satisfies $f(1) = 1$, thus recovering the usual LDT.

The above $q$-LDT scenario has already been numerically verified for a purely probabilistic model with strong correlations [34–36]. In the present paper, we follow along those lines and focus on the possible emergence of the same type of probability for four well-known dynamical models, namely the standard map, the coherent noise model for biological extinctions and earthquakes, the Ehrenfest dog-flea model, and the random walk avalanches. The Ehrenfest model is a genuine $N$-body problem, whereas what plays the role of $N$ in the other three models is the number of successive iterations. In these four models, we numerically verify that, in the space of probability distributions, convergence toward $Q$-Gaussians indeed occurs when summing increasingly large number $N$ of random variables. Here, we focus on the speed at which these $Q$-Gaussians are attained when $N$ increases. In other words, we provide dynamical examples approaching the mathematical development of a $q$-LDT.

Concerning the standard LDT, several nontrivial calculations, either analytical or numerical, are available in the literature, where various classes of systems are focused on [37–44]. Such systems may be either stochastic or Hamiltonian ones. For example, rogue waves, Fermi–Pasta–Ulam–Tsingou chains, population dynamics including birth–death processes, deterministic systems such as the Lorentz gas, Markov dynamics involving both symmetric and asymmetric exclusion processes, glass models exhibiting various dynamical phenomena such as super-Arrhenius temperature dependence of characteristic times, nonexponential relaxation, spatially heterogeneous dynamics, transport decoupling, aging and memory effects.

2 Results

We address now a low-dimensional conservative system, namely the standard map [45]:

$$U_{i+1} = U_i - K \sin X_i$$

$$X_{i+1} = X_i + U_{i+1} \quad (K \geq 0),$$

(1)

($U$ and $X$ are taken as modulo $2\pi$). This is a highly paradigmatic system in the study of Hamiltonian low-dimensional maps [46,47] and has been analyzed deeply in the context of physical applications [48–50] as well as mathematical aspects in the theory of dynamical systems [51,52]. Its applications include particle confinement in magnetic traps, particle dynamics in accelerators, comet dynamics, ionization of Rydberg atoms and electron magneto-transport.

This map is integrable for $K = 0$ and nonintegrable otherwise. If $K \ll 1$ the phase space of the system is dominated by the stability islands. As $K$ increases, chaotic behavior starts to set in for a tiny portion of the phase space. There is a critical $K$ value ($K_c \approx 0.97$) below which the chaotic regions in the phase space do no communicate, and there is a unique chaotic sea for $K > K_c$. If $K \gg 1$ the chaotic sea dominates a large portion of the phase space and for some $K$ values the stability islands become virtually invisible [53,54]. On the other hand, it also appears that, as $K \to \infty$, there will be islands for a residual set of $K$ values [55]. For the phase space portions with stability islands, the system exhibits weak chaotic dynamics, characterized by zero local Lyapunov exponent. Instead, for the portions of chaotic sea, it exhibits strong chaotic dynamics with a positive Lyapunov value. Now let us concentrate on the $X$ variable of the map and define a random variable as the sum of iterates of the map:

$$Y_N := \sum_{i=1}^{N} (X_i - \langle X_i \rangle),$$

(2)

where $\langle \cdots \rangle$ represents the expectation value and we approximate this by sampling over a large number of $M$ initial conditions taken randomly from uniform distribution in $[0, 2\pi]$, that is,

$$\langle X_i \rangle \approx \frac{1}{M} \sum_{j=1}^{M} X_{i(j)}.$$

(3)
For large $M$, this will converge to a number $\mu$ independent of $i$. Therefore, Eq. (2) translates to

$$Y_N = \sum_{i=1}^{N} X_i - N\mu,$$

(4)

which enables us to create a sequence of data points centered around zero from a deterministic dynamical system.

Let us focus on the $K = 10$ case. The CLT basically states that the sum of $N$ independent identically distributed (i.i.d.) random variables (appropriately centered and rescaled), converges onto a Gaussian distribution in the limit $N \to \infty$ (property that is also currently referred to by saying that a Gaussian is the attractor in the space of distributions, not to be confused with the dynamical attractor in phase space, a completely different concept). Although it is evident that the iterates of a deterministic dynamical system would never be completely independent, one can still prove CLTs if the i.i.d. assumption is replaced by the property that the system is strongly mixing [3, 56], which is guaranteed by strong chaos. Therefore, in the context of the standard map, one would expect the usual CLT to hold for $K = 10$. More precisely, $\sqrt{N}(Y_N/N)$ should (weakly) converge, for $N \to \infty$, to a variable with a Gaussian distribution [57]:

$$\lim_{N \to \infty} P_N \left( a \leq \sqrt{N}Y_N/N \leq b \right) = \int_{a}^{b} p(y; \sigma)dy$$

(5)

or equivalently

$$\lim_{N \to \infty} P_N \left( \sqrt{N}Y_N/N \geq z \right) = \int_{z}^{\infty} p(y; \sigma)dy,$$

where the probability density is given by

$$p(y; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{y}{\sigma} \right)^2 \right].$$

(7)

We remind that the prefactor $\sqrt{N}$ before $Y_N$ emerges in order to have gradual data collapse for increasingly large $N$.

For $K = 10$, the dynamics displays strong chaos within the full (or nearly full) phase space. This implies the convergence of $Y_N/N$ to the Gaussian distribution. In Fig. 1a, the distribution of $Y_N/N$ is shown for two representative values of $N$. The Gaussian shape is evident and a representative $z > 0$ value is also indicated in the figure.

The numerical procedure for the CLT results can be summarized as follows: we firstly need to generate a long time series for the random variable of the system under consideration. This random variable is given by Eq. (4) for the standard map and by Eq. (15) for all the other models. The size of the time series is $5 \times 10^7$ for the standard map and even larger than that for the others. Finally, the summation in Eq. (4) must be performed with several different values for $N$. After having enough data, one can generate the histograms for each case (with different $N$ values) using an appropriate box size allowing also a fair sampling of the central part. Two typical results obtained by this procedure are shown in Fig. 1a, c for the standard map.

At this point, we analyze the large deviation behavior of the system, i.e., the speed of convergence of $Y_N/N$ for $N \to \infty$, by numerically exhibiting

$$P_N(Y_N/N > z) \approx C(z)e^{-r_1(z)N}$$

(8)

with the rate function $r_1(z)$ defined by

$$\lim_{N \to \infty} \frac{1}{N} \ln[P_N(Y_N/N > z)/C(z)] = -r_1(z) \quad (r_1(0) = 0),$$

(9)

where $C(z)$ is a factor which, by definition, yields $C(z = 0) = 1/2$. Eq. (9) can be checked for $K = 10$: see Fig. 1c. The exponential decay is evident in the figure, from where the rate function is calculated. Let us focus on the $K = 0$ case. For this case, the random variables that we have defined previously using Eq. (4) are not anymore close to i.i.d., but strongly correlated instead. Therefore, the usual CLT cannot apply here and this constitutes a typical example to test the $Q$-CLT. If this kind of $Q$-generalized CLT holds, the sum of the correlated random variables is expected to converge to a $Q$-Gaussian, which is defined as:

$$p_Q(y; \beta_Q) = A_Q\sqrt{\beta_Q} \exp \left[ -\beta_Q y^2 \right] \quad (\beta_Q > 0),$$

(10)

since this distribution optimizes, under simple constraints, the continuous entropic form $S_q = k \int [1 - f(x)]^{q} dx / q - 1$ with $S_1 = S_{BG} = -k \int f(x) \ln f(x)$; the $Q$-Gaussian distribution is normalized for $Q < 3$, and its second moment is finite for $Q < 5/3$.

Here, $1/\sqrt{\beta_Q}$ characterizes the distribution width of $p_Q(y; \beta_Q)$, $A_Q$ being the normalization factor.
Fig. 1 (Color online) The probability density function of the standard map for two representative values of $N$ for (a) $K = 10$ and (b) $K = 0$. If we multiply both the ordinate and the abscissa by $\sqrt{N}$ in (a) and by $N^\gamma$ with $\gamma \approx 0.65$ in (b), the present data collapse onto a single Gaussian ($Q$-Gaussian with $Q \approx 1.935$) for $K = 10$ ($K = 0$). A typical value $z > 0$ is indicated as well. In (c), (d), we see, respectively, that the large deviation probability $P_N(Y_N/N > z)$ asymptotically decays with $N$ exponentially for $K = 10$ and as a power law (possibly $q$-exponentially with $q = 2.0$) for $K = 0$. The slopes provide the rate function $r_q(z)$ as shown in the Insets. Finally in (e), (f), we represent the ratio between the data and fitted function for $K = 10$ and $K = 0$, respectively. The discrepancy bars at $1 \pm 0.05$ are indicated in dotted lines.
Fig. 2 (Color online) a $K = 10$: $L$-dependence of the variance; b $K = 0$: $L$-dependence of the variance; c $K = 0$: $L$-dependence of the $Q$-variance. The $Q$-variance for continuous distributions is defined as

$$\langle y^2 \rangle_Q = \frac{\int_{\gamma_{i-1}}^{\gamma_i} dy' y^2 |p_Q(y; \beta_Q)|^2}{\int_{\gamma_{i-1}}^{\gamma_i} dy' |p_Q(y; \beta_Q)|^2}$$

(see details in [9,58,59]). For $Q$-Gaussians, it is straightforwardly obtained $\beta_Q(y^2)_Q = \frac{1}{\pi y^2}$ (see, for instance, [60]). The green dashed lines indicate the analytical results, namely $1/2$ for $Q = 1$ ($K = 10$) and $0.939$ for $Q = 1.935$ ($K = 0$).

such a generalized central limit theorem appears to hold

$$\lim_{N \to \infty} P_N \left( N^{\gamma} Y_N / N > z \right) = \int_{z}^{\infty} p_Q(y; \beta_Q) dy.$$  (12)

For this to hold, we have to choose $Q \simeq 1.935$ [57] and $\gamma \simeq 0.65$ (adjusted from the best fit of the data); the exponent $\gamma$ plays, for $K = 0$, the role of $1/2$ for $K = 10$. A sufficient, but not necessary, condition for such a generalized central limit theorem is given by $Q$-independence [29]. Note that, within the context of our example for $K = 0$, $Q$-independence is not satisfied. Indeed, our $\gamma$ differs from $1/(4 - 2Q)$, the exponent which corresponds to the case of $Q$-independence. However, the probability density function is very well approximated by a $Q$-Gaussian with $Q \simeq 1.935$ as can

Note that, as $Q \to 1$, we recover the Gaussian distribution with the density given in Eq. (7). (For completeness, let us mention that it is claimed in [61,62] that the rather standard mathematical connection between the discrete and the continuous forms of $S_q$ carries some odd peculiarities. These claims have, however, been severely counter argued in [63,64]). Recent works [57] provide strong numerical evidence that, in this case,

$$A_Q = \begin{cases} 
\frac{\Gamma \left[ \frac{5 - 3Q}{2(1-Q)} \right]}{\Gamma \left[ \frac{2-Q}{1-Q} \right]} \sqrt{\frac{1-Q}{\pi}} & \text{if } Q < 1, \\
\frac{1}{\sqrt{\pi}} & \text{if } Q = 1, \\
\frac{\Gamma \left[ \frac{Q-1}{2} \right]}{\Gamma \left[ \frac{3 - Q}{2(1-Q)} \right]} \sqrt{\frac{Q-1}{\pi}} & \text{if } 1 < Q < 3.
\end{cases}$$

(11)
be seen in Fig. 1b. Let us emphasize at this point that \( Q \simeq 1.935 \) is the present numerical approximation for the recently established analytical results \( Q = 2 \) [65]. For \( K = 0 \), the problem of slow convergence of numerics to the analytical \( Q = 2 \) result is due to strong correlations among the various initial conditions and one would need extremely large values of \( N \) as well as of the number of initial conditions to tackle this. Obviously, this is not so for the uncorrelated case \( (K = 10) \) where numerical results approach the analytical ones rapidly.

Now we are in the position of studying the speed of convergence of \( Y_N/N \) for \( N \to \infty \), as we have already done for \( K = 10 \) case, whereby providing a large deviation analysis consistent with nonextensive statistical mechanics. Similarly, in this case, we provide numerical evidence for a power law emerging asymptotically for large values of \( N \). Moreover, given the numerous works mentioned above that provide support to the field of \( q \)-statistics, and very especially in the LDT case in [36], we believe that the following form is particularly distinguished:

\[
P_N(Y_N/N > z) \approx C_q(z)e^{-r_q(z)N}, \quad (13)
\]

in the sense that an unique value \( q \) might exist such that

\[
\lim_{N \to \infty} \frac{1}{N} \ln_q[P_N(Y_N/N > z)/C_q(z)] = -r_q(z) \quad (14)
\]

where the \( q \)-rate function \( r_q(z) \geq 0 \), the equality holding for \( z = 0 \). We naturally expect \( C_q(z = 0) = 1/2 \).

In the \( q \to 1 \) limit, we recover \( r_1(z) \) for \( K = 10 \), as expected.

Typical results are given in Fig. 1d. The behavior is analogous to that of the usual case, the exponential being replaced by a \( q \)-exponential with \( q = 2.0 \); compare Fig. 1b of the usual LDT with Fig. 1d for \( q \)-LDT. For \( r_q(z) \), a numerical error up to \((q - 1)/\ln_q(2P_{\max}(z = 0))\) occurs, which vanishes for \( q = 1 \), \( \forall P_{\max}(z = 0) \), and for \( 2P_{\max}(z = 0) = 1 \), \( \forall q \).

In practice, for the present data it does not appear to overcome 1%. The \( N \)-dependence of the data-fit ratios of \( P_N \) is depicted in Fig. 1e, f. The heights of the transients are of the same order (roughly 1.6 for typical values of \( z \)) for \( K = 10 \) and \( K = 0 \), but their durations are sensibly different. Indeed, we verify that the exponential behavior is attained, below a 5% discrepancy, beyond \( N \simeq 3 \times 10^6 \) for \( K = 10 \), whereas the possible \( q \)-exponential behavior is attained beyond \( N \simeq 2 \times 10^5 \) for \( K = 0 \), i.e., 15 times earlier.

As a consistency check we have also calculated the variances corresponding to the \( K = 10 \) and \( K = 0 \) cases. For \( K = 10 \) we calculated, from the computational discrete data, the second moment \((y^2)\) for increasing values of the cutoff \( L \in [0, 1/2] \) such that the data that are taken into account are those corresponding to \( y/N \leq L \), and compared it with the analytical expression obtained from the continuous Gaussian: see Fig. 2a. For \( K = 0 \), we calculated, from the computational discrete data, the second moment \((y^2)\) as well as the \( Q \)-second moment \((y^2)_Q \) with \( Q = 1.935 \), and compared the latter with the analytical expression obtained from the continuous \( Q \)-Gaussian in Eq. (10): see Fig. 2b, c. Naturally, for \( K = 0 \) the \( L \)-dependences of \((y^2)\) and \((y^2)_Q \) differ dramatically. Indeed, since \( Q = 1.935 > 5/3 \), the standard variance is mathematically ill-defined (see, for instance, [60]).

We address next the return distributions of some paradigmatic models, namely, the coherent noise model (CNM), the Ehrenfest dog-flea model (EDFM), and the random walk avalanches model (RWAM). The CNM has been introduced firstly for analyzing biological extinctions [66], but then it has been also used as a simple mean-field model for earthquakes [67]. The model consists of \( N \) agents, each of which having a threshold \( x_i \) against an external stress \( \eta \). In the model, these parameter values are chosen randomly from probability distributions \( p_{\text{thresh}}(x) \) and \( p_{\text{stress}}(\eta) \), respectively. Generically, exponential distribution \( p_{\text{stress}}(\eta) = (1/\sigma) \exp(-\eta/\sigma) \) is used for the external stress, whereas, for \( p_{\text{thresh}}(x) \), the uniform distribution \((0 \leq x \leq 1) \) is chosen. The dynamics of the model can be given in three steps: (i) a random stress \( \eta \) is generated from \( p_{\text{stress}}(\eta) \) and all agents with \( x_i \leq \eta \) are replaced by new agents with new threshold drawn from \( p_{\text{thresh}}(x) \), (ii) finally, a small fraction \( f \) of \( N \) agents is chosen and new thresholds drawn from \( p_{\text{thresh}}(x) \) are assigned, then (iii) these steps are used for the next time step. The number of agents replaced in the first step determines the event size \( s \) [66,67]. The return distribution is noted \( p(\Delta s) \), where

\[
\Delta s = s(t + 1) - s(t) \quad (15)
\]

is the difference between two consecutive event sizes. This quantity here plays the role of the random variable \( Y_N \) that we have used for the standard map. The cor-
responding distributions were studied in [68–70]. The return distributions appear to follow the $Q$-Gaussian form. For example, for $\sigma = 0.05$ it is $Q \simeq 2.10$ [68]. We verify here that the large deviation probability decays as a power law which is consistent with a $q$-exponential with $q \simeq 2.20$ (see Fig. 3a).

The Ehrenfest dog-flea model (EDFM) has been introduced in 1907 by Ehrenfest and Ehrenfest [71]. It is a simple and paradigmatic model of generation–recombination Markov chain describing the process of approaching an equilibrium state in a large set of uncoupled two state systems together with fluctuations avalanches around this state [72]. The simple dynamics of the model is the following: It has $N$ dynamical sites represented by the total number of fleas shared by two dogs, namely dog $A$ and dog $B$. Suppose that there are $N_A$ fleas on $A$ and $N_B$ fleas on $B$ which leads to the population $N = N_A + N_B$. At every time step, a randomly chosen flea jumps from one dog to the other which results in $N_A \to N_A \pm 1$ and $N_B \to N_B \mp 1$. This procedure is repeated for an arbitrary number of times. In the long run, the mean number of fleas on both $A$ and $B$ converges to the equilibrium value, $\langle N_A \rangle = \langle N_B \rangle = N/2$ with some fluctuations. A single fluctuation is described as a process that starts once the number of fleas on one of the dogs becomes larger (or smaller) than the equilibrium value $N/2$ and stops when it gets back to it for the first time. Therefore, termination of a fluctuation specifies the start of the other one. The length ($\lambda$) of a fluctuation is determined by the number of time steps elapsed until the fluctuation ends. The return distribution is defined as in the CNM through Eq. (15). It was analyzed in [73] and it obeys a $Q$-Gaussian form with $Q \simeq 2.32$. We found that the large deviation probability decays as a power law which is consistent with a $q$-exponential with $q = 3$, as seen in Fig. 3b.

It was shown in the RWAM [74] that the avalanche size of a one-dimensional directed sandpile model can be mapped to the area under a Brownian curve with an absorbing boundary at the origin. This is equivalent to a random walker on $[0, \infty)$ with an absorbing boundary at the origin. If we denote the trajectory of the random walker by $x(i)$ with $i = 0, 1, \ldots, N$, the avalanche size can be described as $s = \sum_{i=1}^{N} x(i)$ ($x(0) = 1$). The return distribution of this model is defined as in the CNM, i.e., using Eq. (15), and it was numerically found a $Q$-Gaussian with $Q \simeq 2.32$ [75]. We have determined here that the large deviation probability decays as a power law consistent with a $q$-exponential with $q \simeq 3.0$, as seen in Fig. 3c.

Finally, let us summarize the numerical procedure for the CLT and LDT results. We firstly need to generate a long time series for the random variable of the system under consideration. This random variable is given by Eq. (4) for the standard map and by Eq. (15) for all the other models. The size of the time series is $5 \times 10^7$ for the standard map and even larger than that for the others. Finally, the summation in Eq. (4) must be performed with several different values for $N$. After having enough data, one can generate the histograms for each case (with different $N$ values) using an appropriate box size allowing also a fair sampling of the central part. Two typical results obtained by this procedure are shown in Fig. 1a, c for the standard map. To achieve the large deviation analysis, firstly one needs to localize an $x$-axis value in the histogram, denoted by $z$, as seen in Fig. 1a, b. Then, we calculate the total probability value larger than this $z$. This allows us to construct a plot that represents this value with respect to various $N$ for any particular $z$ value chosen. This can be seen in Fig. 1b, d for the standard map and in Fig. 3 for the other models.

3 Discussion

Let us conclude by reminding that our aim is to approach the fingerprints of Boltzmann–Gibbs statistical mechanics, namely the Maxwellian distribution of velocities and the BG exponential weight for the energies, within a more general context. Indeed, in the realm of nonextensive statistical mechanics based on nonadditive entropies, a $Q$-Gaussian distribution emerges for the velocities and a $q$-exponential weight emerges for the energies, with $(Q, q) \neq (1, 1)$, the equality $Q = q = 1$ holding precisely for the BG theory. These generalized results appear to, respectively, follow along the lines of the central limit theorem and the large deviation theory. This scenario has already been successfully verified for a purely probabilistic model (see [36] and references therein). We numerically checked here this conjectural path with iterative dynamical models, namely the paradigmatic two-dimensional conservative standard map, the coherent noise model, and the random walk avalanches, as well as a genuine $N$-body system, the Ehrenfest dog-flea model. In all cases, the conjecture appears to lurk. Notice however that, for the
standard map, we define sums of $N$ terms precisely as in the central limit theorem. For the other three illustrations, we simply use the return distributions, which in all cases are of the $Q$-Gaussian form. In the present four examples, the Large Deviation probability asymptotically decays with $N$ as a power law which appears to correspond to a $q$-exponential form with the argument $r_q N$ being proportional to $N$, which would in turn be consistent with the Legendre structure of thermodynamics. Indeed, the intensive quantity $r_q$ possibly is related to a nonadditive relative entropy per particle. Along the lines of the present promising results, analytical approaches would naturally be very welcome, either for the specific models studied here, or in the ambitious form of a $q$-generalized theorem for large deviations based on a $Q$-generalized central limit theorem for strongly correlated random variables. To be more precise, we generically expect that many complex systems would exhibit an asymptotic power law at their large deviation behavior. On the other hand, whenever the central limit behavior is concerned, many systems present $Q$-Gaussians as attractors and we then consistently expect $q$-exponentials for the large deviation probabilities with a value for $q$ which univocally depends on $Q$, and which satisfies $q = Q = 1$. It would certainly be interesting that future analytical and/or numerical efforts would focus along these lines for typical complex natural, artificial and social systems. It is important to have in mind that, whereas the asymptotic power-law behavior appears naturally for large values of $N$ (as shown here), the establishment of the distinctive $q$-exponential form demands the exploration along all scales for $N$, particularly those involving relatively small values of $N$. Let us nevertheless emphasize that the ubiquity of these power laws (here and in Ref. [36]) calls for a sensible generalization of the standard LDT exponential behavior. An admissible such generalization should in principle satisfy two conditions: (i) to asymptotically exhibit, of course, the power law, and (ii) its argument should be extensive, i.e., proportional to $N$. The $q$-exponential possibility with $q > 1$ satisfies both.

Acknowledgements We are thankful to two anonymous Referees for useful remarks. Also, we acknowledge partial financial support by the Santa Fe Institute, New Mexico, and the Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, for support of the SFI Micro Working Group “Large Deviations in Complex Systems” meeting. CT acknowledges as well CNPq and Faperj (Brazilian agencies), U.T. is a member of the Science Academy, Bilim Akademisi, Turkey and acknowledges partial support from TUBITAK (Turkish Agency) under the Research Project number 121F269.

Data Availability Statement The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Reif, F.: Fundamentals of Statistical and Thermal Physics. Waveland Press, Long Grove (2008)
2. Pathria, R.K., Beale, P.D.: Statistical Mechanics. Academic Press, New York (2011)
3. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
4. van Kampen, N.G.: Stochastic Processes in Physics and Chemistry. North-Holland, Amsterdam (1981)
