Asymmetric scaling in large deviations
for rare values bigger or smaller than the typical value

Cécile Monthus

Institut de Physique Théorique, Université Paris Saclay, CNRS, CEA, 91191 Gif-sur-Yvette, France

In various disordered systems or non-equilibrium dynamical models, the large deviations of some observables have been found to display different scalings for rare values bigger or smaller than the typical value. In the present paper, we revisit the simpler observables based on independent random variables, namely the empirical maximum, the empirical average, the empirical non-integer moments or other additive empirical observables, in order to describe the cases where asymmetric large deviations already occur. The unifying starting point to analyze the large deviations of these various empirical observables is given by the Sanov theorem for the large deviations of the empirical histogram : the rate function corresponds to the relative entropy with respect to the true probability distribution and it can be optimized in the presence of the appropriate constraints. Finally, the physical meaning of large deviations rate functions is discussed from the renormalization perspective.

I. INTRODUCTION

In macroscopic systems with a large number N of degrees of freedom, it is interesting to analyze how much physical observables can fluctuate as a function of the size N. When u is some intensive variable, one usually distinguishes the three following levels of descriptions for its probability distribution $P_N(u)$:

(i) for large N, the probability distribution $P_N(u)$ becomes concentrated on the typical value u^{typ} that does not depend on N

$$P_N(u) \underset{N\to+\infty}{\simeq} \delta(u - u^{typ}) \tag{1}$$

This statement is the analog of the law of large numbers for the empirical average of independent random variables. Another well-known example is the typical Lyapunov exponent for product of random matrices [1, 2].

(ii) zooming in Eq. 1 reveals the order $\frac{1}{T_N}$ of the small typical fluctuations around the typical value u^{typ}, where T_N is some appropriate scale that grows with N (for instance like a power-law N^χ with some exponent $0 < \chi < 1$ or like a power of $\ln(N)$)

$$u \underset{N\to+\infty}{\simeq} u_{typ} + \frac{v}{T_N} \tag{2}$$

and the rescaled variable $v \equiv T_N(u - u^{typ})$ is distributed with some universal limiting distribution $V(v)$. This statement is the analog of the Central Limit Theorem with the scale $T_N = \sqrt{N}$ and where $V(v)$ is the Gaussian distribution for the universality class of probability distributions whose two first moments are finite (if they are not finite, one obtains the other universality classes involving Lévy stable laws). Another famous example is given by the three universality classes Gumbel-Fréchet-Weibull of Extreme Value Statistics [3, 4], with many applications in various physics domains (see the reviews [5–7] and references therein).

(iii) in the field of large deviations, one is interested instead in evaluating how rare it is for large N to observe some finite value u different from u^{typ}. The standard theory of large deviations is based on the exponential decay [8–10]

$$P_N(u) \underset{N\to+\infty}{\simeq} e^{-NI(u)} \tag{3}$$

where the rate function is positive $I(u) \geq 0$ and vanishes only for the typical value u^{typ} of Eq 1

$$I(u^{typ}) = 0 \tag{4}$$

While the region (ii) of universal typical fluctuations has been traditionally the main focus of studies for various physical observables, the theory of large deviations (iii) is nowadays considered as the unifying language for the statistical physics of equilibrium, non-equilibrium and dynamical systems (see the reviews [8–10] and references therein). In particular, the large deviations with respect to the large time limit of dynamical trajectories has produced an appropriate statistical physics approach for various Markovian processes (see the reviews [11–17] and the PhD Theses [18–21] and the HDR Thesis [22]).

However the recent huge activity on large deviations in the field of random matrices has shown that the maximal eigenvalue [23–28] and many other observables involving the eigenvalues [29–39] display asymmetric scaling in large

deviations, i.e. the probability $P_N(u)$ to observe bigger values than typical $u > u^{typ}$ and smaller values than typical $u < u^{typ}$ are governed by two different scalings $D_N^{\pm}$ (for instance two different power-laws $D_N^{\pm} = N^{\theta_{\pm}}$) and two rate functions $I_{\pm}(u)$

$$P_N(u) \underset{N \to +\infty}{\simeq} e^{-D_N^+ I_+(u)} \text{ for } u \geq u^{typ}$$

$$P_N(u) \underset{N \to +\infty}{\simeq} e^{-D_N^- I_-(u)} \text{ for } u \leq u^{typ} \tag{5}$$

instead of the standard form of Eq. 3. For the maximal eigenvalue [23–28], the physical interpretation of this asymmetry is that to push the maximal eigenvalue inside the Wigner sea, one needs to reorganize all the other eigenvalues, whereas to pull the maximal eigenvalue outside the Wigner sea, one may leave the other eigenvalues unchanged. Via mapping between models belonging to the Kardar-Parisi-Zhang universality class (see the list in the review [27] and references therein), these asymmetric large deviations properties for the biggest eigenvalue of some random matrices ensembles can be rephrased in many other frameworks, in particular :

(a) for the Asymmetric Exclusion process, which is one of the most studied models in the field of the non-equilibrium dynamics of interacting particles (see the reviews [11, 16, 17] and references therein), the interpretation of the asymmetric large deviations is that to slow down the traffic, it is sufficient to slow down a single particle, whereas to speed up the traffic, one needs to speed up all particles [40].

(b) for the Directed Polymer in random medium in dimension $d = 2$, which is one of the simplest disordered model displaying a low temperature glassy frozen phase (see the review [41] and references therein), the interpretation is that an anomalously good ground state energy requires only L anomalously good on-site energies along the polymer, while an 'anomalously bad' ground state energy requires L^d bad on-site energies in the sample.

These examples and their very clear physical meanings show that asymmetric large deviations of Eq. 5 are likely to occur in many other problems in the fields of non-equilibrium dynamics or disordered systems, while they are not considered in the standard theory of large deviations [8–10] based on Eq. 3. As a consequence, it seems useful to revisit simpler observables based on independent random variables where asymmetric large deviations have been found to occur, in particular for the empirical maximum [42, 43], for the empirical average [44–47], and in joint linear statistics [48]. Since these problems have been already studied in details in these references by exact methods, the goal of the present paper is to give a unifying perspective based on the large deviation properties of the empirical histogram in the presence of constraints corresponding to the observables under study. This point of view also allows to make the link with the studies of large deviations in the field of random matrices [23–39] where the Coulomb gas technique is based on the large deviations of the empirical histogram of eigenvalues, in the presence of constraints corresponding to the observables under study. The main difference is that the large deviations of the empirical histogram are governed by the Coulomb interaction energy in the case of random matrices eigenvalues, while it is governed by the relative entropy of the Sanov theorem for the case of independent variables [8–10]. This large deviation framework also makes the link with the Gibbs theory of ensembles in equilibrium statistical physics [8–10] and thus allows to understand why it is natural to expect the possibility of phase transitions in large deviation rate functions (see the recent review [49] and references therein).

The paper is organized as follows. In section II, we recall how the empirical histogram of independent random variables allows to reconstruct interesting observables like the empirical maximum, the empirical average, or other additive empirical observables, with the consequences for typical values. In section III, the large deviations of the empirical histogram is presented as the unifying starting point to analyze the large deviations of empirical observables. In section IV, the asymmetry in the large deviations of the empirical maximum is analyzed on various scales. In section V, the asymmetry in the large deviations of the empirical average is described for the case of stretched exponential decay or power-law decay of the initial distribution. In section VI, the generalization for the large deviations of arbitrary non-integer empirical moments is discussed. In section VII, these large deviations properties are analyzed from the renormalization perspective. Our conclusions are summarized in section VIII.

II. EMPIRICAL OBSERVABLES FOR INDEPENDENT RANDOM VARIABLES

A. Notations

Since our main goal is to analyze the asymmetry in large deviation properties that may occur for simple observables involving N independent random variables x_i drawn with some probability distribution $\pi(x)$, we will focus on the cases of positive variables $0 \leq x < +\infty$, where the decay of the probability distribution $\pi(x)$ for large $x \to +\infty$ is :

(i) either an exponential decay with some exponent $\alpha > 0$

$$\pi_{\alpha,\nu}^{exp}(x) \underset{x \to +\infty}{\simeq} K x^{\nu-1} e^{-x^\alpha} \tag{6}$$

with possibly some power-law prefactor if $\nu \neq 1$, while K is a constant amplitude.

ii) or a power-law decay with some exponent $\mu > 2$ (in order to ensure the existence of the two first moments $\overline{x}$ and $\overline{x^2}$)

$$\pi_\mu^{power}(x) \underset{x \to +\infty}{\simeq} \frac{K}{x^{1+\mu}} \tag{7}$$

But of course these assumptions are not restrictive, and if one is interested into another cases, one can adapt the methods described below by considering the various possible tail behaviors for $x \to -\infty$.

B. Empirical histogram

If one is not interested in the order of appearance of the variables $[x_i]_{1 \leq i \leq N}$ (otherwise see the pedagogical introduction [50] and references therein), all the information is contained in the empirical histogram

$$p_N(x) \equiv \frac{1}{N} \sum_{i=1}^N \delta(x - x_i) \tag{8}$$

Its typical value is of course the 'true' probability distribution $\pi(x)$

$$p_N^{typ}(x) = \pi(x) \tag{9}$$

while the large deviations around this typical value are postponed section III. Let us first recall how the empirical histogram allows to reconstruct the usual empirical observables of interest.

C. Empirical maximum

The information on the empirical maximum

$$x_N^{max} \equiv \max_{1 \leq i \leq N} (x_i) \tag{10}$$

is contained in the empirical histogram of Eq. 8 as follows : the empirical maximum x_N^{max} corresponds to the value x where the empirical number of variables bigger than x

$$\mathcal{N}_N(x) \equiv \sum_{i=1}^N \theta(x_i - x) = N \int_x^{+\infty} dy p_N(y) \tag{11}$$

jumps from 0 to 1

$$0 = \mathcal{N}_L(x_N^{max} + 0) = N \int_{x_N^{max}+0}^{+\infty} dy p_N(y)$$

$$1 = \mathcal{N}_L(x_N^{max} - 0) = N \int_{x_N^{max}-0}^{+\infty} dy p_N(y) \tag{12}$$

The typical value of the empirical histogram of Eq. 9 yields that the typical value M_N of the empirical maximum x_N^{max} of Eq. 10

$$M_N \equiv (x_N^{max})^{typ} \tag{13}$$

is given by the typical position of the jump of Eq. 12

$$1 = N \int_{M_N}^{+\infty} dy p_N^{typ}(y) = N \int_{M_N}^{+\infty} dy \pi(y) \equiv N C(M_N) \tag{14}$$

where we have introduced the complementary cumulative distribution function that measures the integrated tail above the threshold x

$$C(x) \equiv \int_x^{+\infty} dx' \pi(x') \tag{15}$$

The large deviations properties of the ratio

$$r_N \equiv \frac{x_N^{max}}{M_N} = \frac{1}{M_N} \left(\max_{1 \leq i \leq N} (x_i) \right) \tag{16}$$

of typical value unity

$$r_N^{typ} = 1 \tag{17}$$

will be discussed in section IV. To be self-contained, let us now recall the behavior of the typical values M_N as a function of N for the two types of decay under study here.

1. Typical value M_N of the maximum for the exponential decay

For the asymptotic behavior of Eq. 6, the asymptotic behavior of its primitive of Eq. 15 reads

$$C(x) \underset{x \to +\infty}{\simeq} \frac{K}{\alpha} x^{\nu - \alpha} e^{-x^\alpha} \tag{18}$$

Then Eq. 14 determining the typical value M_N of the empirical maximum becomes for large N

$$\frac{1}{N} = C(M_N) \underset{N \to +\infty}{\simeq} \frac{K}{\alpha} M_N^{\nu - \alpha} e^{-M_N^\alpha} \tag{19}$$

and the inversion yields at leading order the well-known logarithmic behavior [3, 4]

$$M_N \underset{N \to +\infty}{\simeq} \left[\ln N + \frac{\nu - \alpha}{\alpha} \ln(\ln N) - \ln \left(\frac{\alpha}{K} \right) \right]^{\frac{1}{\alpha}} \tag{20}$$

2. Typical value M_N of the maximum the power-law decay

For the power-law decay of Eq. 7, the asymptotic behavior of its primitive of Eq. 15

$$C(x) = \int_x^{+\infty} dx' \pi(x') = \frac{K}{\mu x^\mu} \tag{21}$$

yields that the solution of Eq. 14 follows the well-known power-law [3, 4]

$$M_N = \left(\frac{KN}{\mu} \right)^{\frac{1}{\mu}} \tag{22}$$

D. Empirical additive observables

The empirical histogram of Eq. 8 allows to reconstruct any additive observable G_N involving some function $g(x)$

$$G_N \equiv \frac{1}{N} \sum_{i=1}^N g(x_i) = \int_0^{+\infty} dx g(x) p_N(x) \tag{23}$$

Th most studied observable in the whole history of probability is of course the empirical average

$$a_N \equiv \frac{1}{N} \sum_{i=1}^N x_i = \int_0^{+\infty} dx x p_N(x) \tag{24}$$

The empirical moments of arbitrary non-integer order q

$$a_N^{(q)} \equiv \frac{1}{N} \sum_{i=1}^{N} x_i^q = \int_0^{+\infty} dx\, x^q p_N(x) \tag{25}$$

have been also considered [6, 24] in order to interpolate between the case $q = 1$ of the empirical average of Eq. 24 and the empirical maximum of Eq. 10 that should dominate the empirical moment of Eq 25 for large $q \to +\infty$. Another examples include the exponential case $g(x) = e^{tx}$ considered in Ref [51] or the logarithmic case $g(x) = \ln(x)$.

The typical value of the empirical histogram of Eq. 9 yields that the typical values of additive observables of Eq. 23 are simply

$$G_N^{typ} = \int_0^{+\infty} dx\, g(x) p_N^{typ}(x) = \int_0^{+\infty} dx\, g(x) \pi(x) \tag{26}$$

In particular the typical value of the empirical average of Eq. 24 corresponds to the first moment $\overline{x}$ of $\pi(x)$

$$a_N^{typ} = \int_0^{+\infty} dx\, x p_N^{typ}(x) = \int_0^{+\infty} dx\, x \pi(x) = \overline{x} \tag{27}$$

The possibility of asymmetric large deviations properties around this typical value will be discussed in section V.

III.   ANALYSIS BASED ON THE LARGE DEVIATIONS OF THE EMPIRICAL HISTOGRAM

A.   Reminder on the Sanov theorem involving the relative entropy

The large deviations of the empirical histogram $p_N(x)$ of Eq. 8 around its typical value $p_N^{typ}(x) = \pi(x)$ of Eq 9 are described by the Sanov theorem (see the reviews [8–10] and the pedagogical introduction [50])

$$\mathcal{P}_N[p_N(.)] \underset{N \to +\infty}{\simeq} \delta\left(1 - \int dx\, p_N(x)\right) e^{-N S^{rel}(p_N(.)|\pi(.))} \tag{28}$$

that involves of course the normalization constraint of the empirical histogram $\int dx\, p_N(x) = 1$, while the rate function in the exponential is the relative entropy of the empirical histogram $p_N(x)$ with respect to the true probability distribution $\pi(x)$

$$S^{rel}(p_N(.)|\pi(.)) \equiv \int dx\, p_N(x) \ln\left(\frac{p_N(x)}{\pi(x)}\right) \tag{29}$$

B.   Exact generating function of the empirical histogram for finite N

Among the various derivations of Eq. 28, one is based on the exact generating function $\mathcal{Z}_N[\kappa(.)]$ of the empirical histogram for any finite N

$$\mathcal{Z}_N[\kappa(.)] \equiv \int \mathcal{D}p_N(.) \mathcal{P}_N[p_N(.)]\, e^{N \int dx\, \kappa(x) p_N(x)} = \int dx_1 \pi(x_1) ... \int dx_N \pi(x_N)\, e^{\sum_{i=1}^{N} \kappa(x_i)}$$

$$= \prod_{i=1}^{N} \left(\int dx_i \pi(x_i) e^{\kappa(x_i)}\right) = \left(\int dx\, \pi(x) e^{\kappa(x)}\right)^N \equiv e^{N\Phi[\kappa(.)]} \tag{30}$$

where the scaled cumulant generating function

$$\Phi[\kappa(.)] = \ln\left(\int dx\, \pi(x) e^{\kappa(x)}\right) \tag{31}$$

is related to the relative entropy of Eq. 29 via the appropriate Legendre transform (see [50] for more details on the Legendre transforms in the two directions).

C. Constraint to reproduce the cumulative distribution of the maximum x_N^{max}

The probability distribution $\mathcal{X}_N(x_N^{max})$ of the empirical maximum x_N^{max} of Eq. 10 is well known to be [3, 4]

$$\mathcal{X}_N(x_N^{max}) = N\pi(x_N^{max})\left[1 - C(x_N^{max})\right]^{N-1} \tag{32}$$

in terms of the cumulative function $C(x)$ introduced in Eq. 15.

Here it is instructive to mention how it can be reproduced from the large deviations of the empirical histogram of Eq. 28 : the cumulative probability of the maximum x_N^{max} amounts to replace the normalisation constraint $\int dx p_N(x) = 1$ by the two constraints

$$\int_{x_N^{max}}^{+\infty} dx p_N(x) = 0$$
$$\int_0^{x_N^{max}} dx p_N(x) = 1 \tag{33}$$

leading to

$$\int_0^{x_N^{max}} dx \mathcal{X}_N(x) = \int \mathcal{D}p_N(.)\delta\left(1 - \int_0^{x_N^{max}} dx p_N(x)\right)\delta\left(\int_{x_N^{max}}^{+\infty} dx p_N(x)\right)e^{-NS^{rel}(p_N(.)|\pi(.))} \tag{34}$$

Introducing the Lagrange multiplier ω, one needs to optimize the functional

$$\mathcal{L}(p_N(.)) = -S^{rel}(p_N(.)|\pi(.)) + \omega\left(1 - \int_0^{x_N^{max}} dx p_N(x)\right)$$
$$= -\int dx p_N(x)\ln\left(\frac{p_N(x)}{\pi(x)}\right) + \omega\left(1 - \int_0^{x_N^{max}} dx p_N(x)\right) \tag{35}$$

over the empirical histogram $p_N(.)$

$$0 = \frac{\partial\mathcal{L}(p_N(.))}{\partial p_N(x)} = -\ln\left(\frac{p_N(x)}{\pi(x)}\right) - 1 - \omega \tag{36}$$

The optimal solution is thus simply proportional to the true distribution $\pi(x)$ on $[0, x_N^{max}]$ (while it vanishes for $x > x_N^{max}$)

$$p_N^*(x) = \pi(x)e^{-1-\omega}\theta(0 \le x \le x_N^{max}) \tag{37}$$

where the normalization constraint determines the Lagrange multiplier ω

$$1 = \int_0^{x_N^{max}} dx p_N^*(x) = e^{-1-\omega}\int_0^{x_N^{max}} dx\pi(x) = e^{-1-\omega}\left[1 - C(x_N^{max})\right] \tag{38}$$

Plugging the corresponding optimal value of the functional of Eq. 35

$$\mathcal{L}(p_N^*(.)) = -\int dx p_N^*(x)\ln\left(\frac{p_N^*(x)}{\pi(x)}\right) = 1 + \omega = \ln\left[1 - C(x_N^{max})\right] \tag{39}$$

into Eq. 34

$$\int_0^{x_N^{max}} dx \mathcal{X}_N(x) \underset{N\to+\infty}{\simeq} e^{N\mathcal{L}(p_N^*(.))} = e^{N\ln[1-C(x_N^{max})]} \tag{40}$$

thus allows to recover the exact cumulative distribution of Eq. 32. In this derivation, Eq. 34 thus corresponds to the entropic cost for the emptiness of the region $[x_N^{max}, +\infty[$. Section IV will be devoted to the asymmetric large deviations properties of this distribution.

D. Standard large deviations for additive empirical observables

The probability distribution $\mathcal{G}_N(G_N)$ of the additive observable of Eq. 23

$$G_N \equiv \frac{1}{N} \sum_{i=1}^{N} g(x_i) = \int_0^{+\infty} dx g(x) p_N(x) \tag{41}$$

can be directly characterized by its exact generating function for finite N by applying Eq 30 to the case $\kappa(x) = kg(x)$

$$Z_N(k) \equiv \int dG \mathcal{G}_N(G) e^{NkG} = \int \mathcal{D}p_N(.)\mathcal{P}_N[p_N(.)] e^{Nk \int dx g(x) p_N(x)} = \left(\int dx \pi(x) e^{kg(x)} \right)^N \equiv e^{N\phi(k)} \tag{42}$$

with the scaled cumulant generating function $\phi(k)$

$$\phi(k) = \ln \left(\int dx \pi(x) e^{kg(x)} \right) \tag{43}$$

The alternative evaluation of Eq. 42 based on the standard Large deviation form for the probability

$$\mathcal{G}_N(G) \underset{N \to \infty}{\simeq} e^{-NI(G)} \tag{44}$$

yields

$$e^{N\phi(k)} \equiv \int dG \mathcal{G}_N(G) e^{NkG} \underset{N \to +\infty}{\simeq} \int dG e^{N(kG - I(G))} \tag{45}$$

via the saddle-point method for the integral over G that $\phi(k)$ is the Legendre transform of the rate function $I(G)$

$$\begin{aligned} \phi(k) &= kG - I(G) \\ 0 &= k - I'(G) \end{aligned} \tag{46}$$

with the reciprocal Legendre transform

$$\begin{aligned} I(G) &= kG - \phi(k) \\ 0 &= G - \phi'(k) \end{aligned} \tag{47}$$

Another way to understand the physical meaning of these Legendre transforms consists in evaluating the probability $\mathcal{G}_N(G)$ via the addition of the sum constraint in the large deviations of the empirical histogram of Eq. 28 :

$$\mathcal{G}_N(G) = \int \mathcal{D}p_N(.)\delta \left(1 - \int_0^{+\infty} dx p_N(x) \right) \delta \left(G - \int_0^{+\infty} dx g(x) p_N(x) \right) e^{-NS^{rel}(p_N(.)|\pi(.))} \tag{48}$$

Introducing the two Lagrange multiplier ω and k, one needs to optimize the functional

$$\begin{aligned} \mathcal{L}(p_N(.)) &= -S^{rel}(p_N(.)|\pi(.)) + \omega \left(1 - \int_0^{+\infty} dx p_N(x) \right) - k \left(G - \int_0^{+\infty} dx g(x) p_N(x) \right) \\ &= -\int dx p_N(x) \ln \left(\frac{p_N(x)}{\pi(x)} \right) + \omega \left(1 - \int_0^{+\infty} dx p_N(x) \right) - k \left(G - \int_0^{+\infty} dx g(x) p_N(x) \right) \end{aligned} \tag{49}$$

over the empirical histogram $p_N(.)$

$$0 = \frac{\partial \mathcal{L}(p_N(.))}{\partial p_N(x)} = -\ln \left(\frac{p_N(x)}{\pi(x)} \right) - 1 - \omega + kg(x) \tag{50}$$

The optimal solution reads

$$p_N^*(x) = e^{-1-\omega+kg(x)} \pi(x) \tag{51}$$

where the the Lagrange multipliers are fixed by the two constraints

$$
\begin{aligned}
1 &= \int_0^{+\infty} dx p_N^*(x) = e^{-1-\omega} \int_0^{+\infty} dx e^{kg(x)} \pi(x) \\
G &= \int_0^{+\infty} dx g(x) p_N^*(x) = e^{-1-\omega} \int_0^{+\infty} dx g(x) e^{kg(x)} \pi(x)
\end{aligned}
\tag{52}
$$

i.e. in terms of the function $\phi(k)$ introduced in Eq. 43

$$
\begin{aligned}
1 + \omega &= \phi(k) \\
G &= \phi'(k)
\end{aligned}
\tag{53}
$$

The corresponding optimal value of the functional of Eq. 49 using 53 thus involves the Legendre transform $I(G)$ of $\phi(k)$ (Eqs 46 and 47)

$$
\mathcal{L}(p_N^*(.)) = -\int dx p_N^*(x) \ln\left(\frac{p_N^*(x)}{\pi(x)}\right) = 1 + \omega - kG = \phi(k) - kG = -I(G)
\tag{54}
$$

as it should to recover via Eq. 48

$$
\mathcal{G}_N(G) \underset{N \to +\infty}{\simeq} e^{N\mathcal{L}(p_N^*(.))} = e^{-NI(G)}
\tag{55}
$$

Here the analogy with the Gibbs theory of ensembles in equilibrium statistical physics is obvious : the effective distribution of Eq 51 for an individual random variable x is the analog of the Boltzmann distribution in the canonical ensemble, where the Lagrange multiplier k is conjugated to the quantity $g(x)$, whose average G over the N variables is fixed.

Of course these computations make sense only if the integrals of Eq. 52 converge : depending on the function $g(x)$ defining the empirical observable G_N under study, these integrals may diverge in some region of the Lagrange multiplier k. The consequences for the non-standard large deviations properties will be discussed for the case of the empirical average in section V and for the empirical moments in section VI.

E. Large deviations for joint additive empirical observables

If one is interested in the large deviations of the joint probability of two additive empirical observables, one needs to add another constraint in Eq. 48, as described in detail in Ref [48] for the joint probability of the empirical average and of an empirical moment of order q. More generally, one can add as many constraints as needed for the problem one is interested in.

IV. ASYMMETRY IN THE LARGE DEVIATIONS OF THE EMPIRICAL MAXIMUM

A. Probability distribution of the ratio $r_N = \frac{x_N^{max}}{M_N}$

Via the change of variables $r = \frac{x_N^{max}}{M_N}$ of Eq. 16, the probability distribution $\mathcal{X}_N(x_N^{max})$ of Eq. 32 becomes the probability distribution

$$
\mathcal{R}_N(r) = N M_N \pi(r M_N) \left[1 - C(r M_N)\right]^{N-1}
\tag{56}
$$

Since the typical value M_N is large as a consequence of the equation $C(M_N) = \frac{1}{N}$, while the ratio r is finite, the value of $(r M_N)$ is also large, i.e. the value of $C(r M_N)$ is also small . Then Eq. 56 becomes

$$
\mathcal{R}_N(r) \underset{N \to +\infty}{\simeq} N M_N \pi(r M_N) e^{-N C(r M_N)}
\tag{57}
$$

It is thus more convenient to substitute $N = \frac{1}{C(M_N)}$ in order to write everything in terms of the single scale M_N

$$
\mathcal{R}_N(r) \underset{N \to +\infty}{\simeq} \frac{M_N \pi(r M_N)}{C(M_N)} e^{-\frac{C(r M_N)}{C(M_N)}}
\tag{58}
$$

B.   Case of the exponential decay

For the exponential decay of Eq. 6, the asymptotic behavior of the function $C(x)$ given in Eq. 18 yields in Eq. 58

$$\mathcal{R}_N(r) \underset{N\to+\infty}{\simeq} \alpha M_N^\alpha r^{\nu-1} e^{M_N^\alpha(1-r^\alpha)} e^{-r^{\nu-\alpha} e^{M_N^\alpha(1-r^\alpha)}} \tag{59}$$

This means that the rescaled variable

$$y \equiv M_N^\alpha(r^\alpha - 1) + (\alpha - \nu)\ln r \underset{M_N\to+\infty}{\simeq} M_N^\alpha(r^\alpha - 1) \tag{60}$$

is distributed with the Gumbel probability distribution [3, 4]

$$G(y) \equiv e^{-y} e^{-e^{-y}} \tag{61}$$

whose very strong asymmetry for the asymptotic behaviors for $y \to \pm\infty$ is well-known

$$G(y) \underset{y\to+\infty}{\simeq} e^{-y}$$
$$G(y) \underset{y\to-\infty}{\simeq} e^{-e^{-y}} \tag{62}$$

As a consequence when r is finite and different from the typical value $r^{typ} = 1$, the variable y of Eq. 60 will be near $(\pm\infty)$ depending on the sign of $(r-1)$: the asymptotic behaviors of Eqs 62 will thus produce completely different scalings in the region bigger than typical $r > 1$ [42, 43]

$$\mathcal{R}_N(r) \underset{N\to+\infty}{\propto} e^{-M_N^\alpha(r^\alpha - 1)} \qquad \text{for } r > 1 \tag{63}$$

and in the region smaller than typical $0 < r < 1$

$$\mathcal{R}_N(r) \underset{N\to+\infty}{\propto} e^{-r^{\beta+1-\alpha} e^{M_N^\alpha(1-r^\alpha)}} \qquad \text{for } 0 < r < 1 \tag{64}$$

The link with the region of small typical fluctuations around the typical value $r^{typ} = 1$ usually considered in the Extreme Value Statistics [3, 4] corresponds here to the Taylor expansion at first order of the variable y of Eq. 60

$$y \underset{r\to 1}{\simeq} M_N^\alpha \left[\alpha(r-1) + O(r-1)^2) \right] \tag{65}$$

i.e. the appropriate rescaling to have a finite variable y distributed with the Gumbel distribution $G(y)$ is

$$r = 1 + \frac{y}{\alpha M_N^\alpha} \tag{66}$$

i.e. for the unrescaled maximum of Eq. 10

$$x_N^{max} = r M_N = \left(1 + \frac{y}{\alpha M_N^\alpha} \right) M_N = M_N + \frac{y}{\alpha} M_N^{1-\alpha} \tag{67}$$

where the behavior of M_N as a function of N was recalled in Eq. 20.

C.   Case of the power-law decay

For the power-law decay of Eq. 7, the asymptotic behavior of its primitive of Eq. 21 yields in Eq. 58

$$\mathcal{R}_N(r) \underset{N\to+\infty}{\simeq} \frac{\mu}{r^{1+\mu}} e^{-\frac{1}{r^\mu}} \equiv F_\mu(r) \tag{68}$$

where the Fréchet distribution $F_\mu(r)$ of parameter μ appears for any finite value r, while the scale M_N has completely disappeared, in contrast to the exponential decay case described above. So here the probability of the values $r \neq r^{typ} \neq 1$ do not decay with N.

D. Asymmetry beyond the regime of finite ratio r

In the following section, we will need the probability of the maximum of Eq 32 beyond the regime of finite ratio r with respect to the typical value M_N. In the region much bigger than typical $x_N^{max} \gg M_N$ where $C(x_N^{max}) \ll C(M_N) = \frac{1}{N}$, the factor $[1 - C(x_N^{max})]^{N-1}$ can be neglected in Eq 32 and one obtains the leading behavior

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \gg M_N}{\simeq} N\pi(x_N^{max}) \tag{69}$$

The physical meaning is that one just needs to draw the anomalously big value x_N^{max}, while the other $(N-1)$ variables may remain typical and thus have no probabilistic cost.

On the contrary, in the region much smaller than typical $x_N^{max} \ll M_N$ where $C(x_N^{max}) \gg C(M_N) = \frac{1}{N}$, the factor $[1 - C(x_N^{max})]^{N-1}$ is the leading behavior in Eq 32 and produces an extensive cost in N in the exponential

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \ll M_N}{\simeq} [1 - C(x_N^{max})]^{N-1} \underset{x_N^{max} \ll M_N}{\simeq} e^{-NC(x_N^{max})} \tag{70}$$

This asymmetry is thus very strong, and reads for the exponential decay of Eq. 6 with Eq. 18,

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \gg M_N}{\simeq} KN(x_N^{max})^{\nu-1} e^{-(x_N^{max})^\alpha}$$

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \ll M_N}{\simeq} e^{-N\frac{K}{\alpha}(x_N^{max})^{\nu-\alpha} e^{-(x_N^{max})^\alpha}} \tag{71}$$

while for the power-law decay of Eq 7 with Eq 21, it is given by

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \gg M_N}{\simeq} \frac{NK}{(x_N^{max})^{1+\mu}}$$

$$\mathcal{X}_N(x_N^{max}) \underset{x_N^{max} \ll M_N}{\simeq} e^{-\frac{NK}{(x_N^{max})^\mu}} \tag{72}$$

V. POSSIBLE ASYMMETRY IN THE LARGE DEVIATIONS OF THE EMPIRICAL AVERAGE

Whenever the first moment $\overline{x}$ and the variance $\sigma^2 \equiv \overline{x^2} - (\overline{x})^2$ are finite, the Central Limit Theorem means that the empirical average of Eq 24 will display typical fluctuations of order $\frac{1}{\sqrt{N}}$ around the typical value $\overline{x}$

$$a_N \equiv \frac{1}{N} \sum_{i=1}^N x_i \underset{N \to +\infty}{\simeq} \overline{x} + \frac{v}{\sqrt{\sigma^2 N}} \tag{73}$$

where v is a Gaussian random variable of zero mean and variance unity. In this section, we focus on the large deviations properties for the probability distribution $\mathcal{A}_N(a)$ of the empirical average a

$$\mathcal{A}_N(a) \equiv \int_0^{+\infty} dx_1 \pi(x_1)... \int_0^{+\infty} dx_N \pi(x_N) \delta\left(a - \frac{1}{N} \sum_{i=1}^N x_i\right) \tag{74}$$

to discuss how rare it is to observe $a \neq \overline{x}$ for large N and when an asymmetry will occur.

A. Standard Large deviation theory for the exponential decay with exponent $\alpha \geq 1$

The standard large deviation theory recalled in section III D for additive empirical observable applies to the empirical average $G = a$ with $g(x) = x$: the probability to observe a value $a \neq a^{typ} = \overline{x}$ is exponentially small in N

$$\mathcal{A}_N(a) \underset{N \to \infty}{\simeq} e^{-NI(a)} \tag{75}$$

and the rate function $I(a)$ can be either evaluated directly or can be computed as the Legendre transform (Eqs 46 and 47) of the scaled cumulant generating function of Eq. 43

$$\phi(k) = \ln\left(\int_0^{+\infty} dx \pi(x) e^{kx}\right) \tag{76}$$

For the exponential decay of Eq. 6 with an exponent $\alpha > 1$, the scaled cumulant generating function of Eq. 76 is defined for any $k \in]-\infty, +\infty[$, and the above large deviation theory can be applied.

For the exponential decay of Eq. 6 with an exponent $\alpha = 1$, the scaled cumulant generating function of Eq. 76 is defined only for $k \in]-\infty, 1[$, and it is thus useful to describe the example of the gamma distribution of parameter ν

$$\gamma_\nu(x) \equiv \frac{1}{\Gamma(\nu)} x^{\nu-1} e^{-x} \tag{77}$$

Its Laplace transform is simply

$$\int_0^{+\infty} dx e^{-px} \gamma_\nu(x) \quad = \frac{1}{(1+p)^\nu} \tag{78}$$

so that computing its power N simply amounts to change the parameter ν into $(N\nu)$. As a consequence, the sum of N variables x_i is distributed with the gamma distribution $\gamma_{N\nu}(.)$ of parameter $(N\nu)$ that corresponds to the convolution of N distributions $\gamma_\nu(.)$. After the rescaling by N, the probability distribution of the empirical average is thus exactly given by

$$\mathcal{A}_N(a) \quad = N\gamma_{N\nu}(Na) = N\frac{1}{\Gamma(N\nu)}(Na)^{N\nu-1} e^{-Na} \tag{79}$$

For large N, the Stirling approximation for $\Gamma(N\nu)$ yields the large deviation form of Eq. 75 where the rate function

$$I_\nu(a) = a - \nu - \nu \ln\left(\frac{a}{\nu}\right) \tag{80}$$

is well defined for $a \in]0, +\infty[$ and measures how rare it is to observe a value a different from the typical value $a^{typ} = \nu$.

The corresponding scaled generating cumulant function $\phi(k)$ of Eq. 76 is defined only for $k < 1$.

$$\phi_\nu(k) = -\nu \ln(1-k) \tag{81}$$

The correspondence between k and a via the Legendre transform of Eq. 46 is

$$k_a = I'_\nu(a) = 1 - \frac{\nu}{a} \tag{82}$$

or equivalently via the reciprocal Legendre transform of Eq. 47

$$a_k = \phi'_\nu(k) = \frac{\nu}{1-k} \tag{83}$$

So the region $k \in]-\infty, 0[$ allows to parametrize the whole smaller than typical region $a \in]0, \nu[$ while the region $k \in]0, 1[$ allows to parametrize the whole bigger than typical region $a \in]\nu, +\infty[$ without problems.

B. Asymmetry in the large deviations for stretched exponential decay $0 < \alpha < 1$

1. Usual large deviation form in the region smaller than typical $a \leq a^{typ} = \overline{x}$

For the exponential decay of Eq. 6 with an exponent $0 < \alpha < 1$, the scaled cumulant generating function of Eq. 76 is defined only for $k \in]-\infty, 0]$ that corresponds to the region smaller than typical $a \leq a^{typ} = \overline{x}$, where the usual large deviation form will thus be valid

$$\mathcal{A}_N(a) \underset{N\to+\infty}{\propto} e^{-NI_-(a)} \quad \text{for} \quad a \leq a^{typ} = \overline{x} \tag{84}$$

The rate function $I_-(a)$ for $a \leq a^{typ} = \overline{x}$ corresponds to the Legendre transform of the function $\phi(k)$ defined for $k \leq 0$.

2. Unusual large deviation in the region $a > a^{typ}$

For $k > 0$ that corresponds to the region bigger than the typical value $a > a^{typ} = \overline{x}$, the function $\phi(k)$ as defined by Eq. 76 does not exist as a consequence of the divergence of the integral at $(+\infty)$ when $\pi(x)$ decays only as a stretched exponential with $0 < \alpha < 1$

$$\int^{+\infty} dx \pi(x) e^{kx} \propto \int^{+\infty} dx e^{kx - x^\alpha} = +\infty \tag{85}$$

This suggests to consider the strategy based on the maximum alone : one considers that $(N - 1)$ variables have their typical sum $(N - 1)\overline{x}$, which happens with probability one for large N, i.e. with no probabilistic cost, while the remaining variable, that will have to coincide with the maximum x_N^{max} of Eq. 10, should be anomalously big in order to satisfy the sum constraint

$$x_N^{max} = Na - (N - 1)\overline{x} \underset{N \to +\infty}{\propto} N(a - \overline{x}) \tag{86}$$

So the cost of this strategy directly involves the probability $\mathcal{X}_N(x_N^{max})$ of Eqs 69 and 71 of the anomalously extensive value of Eq. 86

$$\mathcal{A}_N^{StrategyMax}(a) \underset{N \to +\infty}{\propto} \mathcal{X}_N(x_N^{max} \simeq N(a - \overline{x})) \underset{x_N^{max} \gg M_N}{\simeq} \simeq K N^\nu (a - \overline{x})^{\nu - 1} e^{-N^\alpha (a - \overline{x})^\alpha} \underset{N \to +\infty}{\propto} e^{-N^\alpha I_+(a)} \tag{87}$$

that decays only as the stretched exponential of exponent $\alpha \in]0, 1[$. The corresponding rate function

$$I_+(a) = (a - \overline{x})^\alpha \quad \text{for} \quad a > \overline{x} \tag{88}$$

has been proven to be valid in the whole region $a > \overline{x}$ in Refs [44, 46].

C. Asymmetry in the large deviations for power-law decay

For the power-law decay of Eq. 7, one has the same scenario as for the stretched exponential case discussed above:
(i) in the region smaller than typical $a \leq a^{typ} = \overline{x}$, the usual large deviation form is valid

$$\mathcal{A}_N(a) \underset{N \to +\infty}{\propto} e^{-N I_-(a)} \quad \text{for} \quad a \leq a^{typ} = \overline{x} \tag{89}$$

where the rate function $I_-(a)$ corresponds to the Legendre transform of the function $\phi(k)$ defined for $k \leq 0$.

(ii) in the region bigger than the typical value $a > a^{typ} = \overline{x}$ where the function $\phi(k)$ of Eq. 76 does not exist as a consequence of the divergence of the integral, the strategy based on the anomalous maximum of Eq. 86 leads to the probability (using Eqs 69 and 72)

$$\mathcal{A}_N^{StrategyMax}(a) \underset{N \to +\infty}{\propto} \mathcal{X}_N(x_N^{max} \simeq N(a - \overline{x})) \underset{x_N^{max} \gg M_N}{\simeq} \simeq \frac{K}{N^\mu (a - \overline{x})^{1+\mu}} \tag{90}$$

that decays only as the power-law $N^{-\mu}$. This phenomenon of condensation in the power-law case has been studied in great detail in the references [45, 47, 48], with motivations coming from the zero-range process (see explanations and references in [45, 47, 48]).

VI. ASYMMETRY IN THE LARGE DEVIATIONS OF THE EMPIRICAL MOMENT OF ORDER $q > 0$

The analysis of the previous section concerning the empirical average can be directly generalized to obtain the large deviations properties of the empirical moment of arbitrary non-integer order $q > 0$ of Eq 25.

A. Standard Large deviation theory for the exponential decay with exponent $\alpha \geq q$

The standard Large deviation theory recalled in section III D for additive empirical observable applies to the empirical moment $G_N = a_N^{(q)}$ of arbitrary non-integer order $q > 0$ of Eq 25 with $g(x) = x^q$: the probability to observe a value $a_q \neq \overline{x^q}$ is exponentially small in N

$$\mathcal{A}_N^{(q)}(a_q) \underset{N \to \infty}{\simeq} e^{-N I_q(a_q)} \tag{91}$$

and the rate function $I(a_q)$ corresponds to the Legendre transform (Eqs 46 and 47) of the scaled cumulant generating function of Eq. 43

$$\phi_q(k) \quad = \ln\left(\int_0^{+\infty} dx\pi(x)e^{kx^q}\right) \tag{92}$$

which is well defined for any $k \in]-\infty, +\infty[$ when $\pi(x)$ displays the exponential decay of Eq. 6 with an exponent $\alpha > q$.

B.   Asymmetry in the large deviations for stretched exponential decay $0 < \alpha < q$

1.   Usual large deviation form in the region smaller than typical $a_q \leq \overline{x^q}$

For the exponential decay of Eq. 6 with an exponent $0 < \alpha < q$, the scaled cumulant generating function of Eq. 92 is defined only for $k \in]-\infty, 0]$ that corresponds to the region smaller than typical $a_q \leq \overline{x^q}$, where the usual large deviation form will thus be valid

$$\mathcal{A}_N(a_q) \underset{N\to+\infty}{\propto} e^{-NI_q^-(a)} \quad \text{for} \quad a \leq \overline{x^q} \tag{93}$$

The rate function $I_q^-(a_q)$ corresponds to the Legendre transform of the function $\phi_q(k)$ defined for $k \leq 0$.

2.   Unusual large deviation in the region $a_q > \overline{x^q}$

For $k > 0$ that corresponds to the region bigger than the typical value $a_q > \overline{x^q}$ the function $\phi_q(k)$ as defined by Eq. 92 does not exist. The strategy based on the maximum alone explained in the previous section can be then considered: $(N-1)$ variables x_i^q have their typical sum $(N-1)\overline{x^q}$, which happens with probability one for large N, i.e. with no probabilistic cost, while the remaining variable will have to coincide with the power q of the maximum x_N^{max} of Eq. 10. This maximum should be anomalously big in order to satisfy the sum constraint

$$x_N^{max} = [Na_q - (N-1)\overline{x^q}]^{\frac{1}{q}} \underset{N\to+\infty}{\propto} N^{\frac{1}{q}}(a_q - \overline{x^q})^{\frac{1}{q}} \tag{94}$$

So the cost of this strategy reads in terms of the probability $\mathcal{X}_N(x_N^{max})$ of Eqs 69 and 71 of the anomalously big value of Eq. 94

$$\mathcal{A}_N^{(q)StrategyMax}(a_q) \underset{N\to+\infty}{\propto} \mathcal{X}_N\left(x_N^{max} \simeq N^{\frac{1}{q}}(a_q - \overline{x^q})^{\frac{1}{q}}\right) \underset{x_N^{max}\gg M_N}{\simeq} KN^{1+\frac{\nu-1}{q}}(a_q - \overline{x_q})^{\frac{\nu-1}{q}}e^{-N^{\frac{\alpha}{q}}(a_q-\overline{x^q})^{\frac{\alpha}{q}}}$$

$$\underset{N\to+\infty}{\propto} e^{-N^{\frac{\alpha}{q}}I_q^+(a_q)} \tag{95}$$

that decays only as the stretched exponential of exponent $\frac{\alpha}{q} \in]0, 1[$, with the corresponding rate function

$$I_q^+(a_q) = (a_q - \overline{x^q})^{\frac{\alpha}{q}} \quad \text{for} \quad a_q > \overline{x^q} \tag{96}$$

3.   Discussion

So for any exponential decay with exponent $\alpha > 0$ in Eq. 6, only the empirical moments of order $q < \alpha$ display a standard form of large deviations, while the empirical moments of order $q > \alpha$ will be characterized by asymmetric large deviations. For instance for the gamma distribution of Eq. 77 corresponding to $\alpha = 1$, where the large deviations of the empirical average a corresponding to $q = 1$ are still standard (with the rate function of Eq. 80), all the empirical moments of order $q > 1$ will have asymmetric large deviations, in particular the empirical second moment corresponding to $q = 2$.

VII.   RENORMALIZATION INTERPRETATION OF LARGE DEVIATIONS RATE FUNCTIONS

The region of typical fluctuations around the typical value (see the Introduction around Eq. 2) has been analyzed in detail from the renormalization point of view, both for the sum of random variables [52–54] and for the maximum of random variables [55–59]. In this section, it is thus interesting to discuss the meaning of large deviations from the renormalization perspective.

A.   Merging two sets of N variables

To see more clearly the renormalization meaning of large deviations, it is interesting to consider the merging of two sets of N random variables :

(1) the first set $[x_i]_{1 \leq i \leq N}$ of variables is drawn with the probability distribution $\pi_1(x)$ and is characterized by the empirical histogram

$$p_N^{(1)}(x) \equiv \frac{1}{N} \sum_{i=1}^{N} \delta(x - x_i) \tag{97}$$

(2) the second set $[x_i]_{N+1 \leq i \leq 2N}$ of variables is drawn with the probability distribution $\pi_2(x)$ and is characterized by its empirical histogram

$$p_N^{(2)}(x) \equiv \frac{1}{N} \sum_{i=N+1}^{2N} \delta(x - x_i) \tag{98}$$

B.   Renormalization for the large deviations of the empirical histogram

Each of these two sets labelled by $b = 1, 2$ is characterized by the large deviation properties of its empirical histogram $p_N^{(b)}(x)$ (Eq. 28 and 29)

$$\mathcal{P}_N^{(b)}[p_N^{(b)}(.)] \underset{N \to +\infty}{\simeq} \delta \left(1 - \int dx p_N^{(b)}(x)\right) e^{-NS^{rel}(p_N^{(b)}(.)|\pi_b(.))} \tag{99}$$

or its exact generating function $\mathcal{Z}_N^{(b)}[\kappa(.)]$ of Eq. 30 for any finite N

$$\mathcal{Z}_N^{(b)}[\kappa(.)] \quad \equiv \int \mathcal{D} p_N^{(b)}(.) \mathcal{P}_N^{(b)}[p_N^{(b)}(.)] \, e^{N \int dx \kappa(x) p_N^{(b)}(x)} = \left(\int dx \pi_b(x) e^{\kappa(x)}\right)^N \equiv e^{N\Phi_b[\kappa(.)]} \tag{100}$$

Via the merging of the data of Eqs 97 and 98, the global histogram for the $(2N)$ variables is simply the average of the two histograms

$$p_{2N}(x) \equiv \frac{1}{2N} \sum_{i=1}^{2N} \delta(x - x_i) = \frac{p_N^{(1)}(x) + p_N^{(2)}(x)}{2} \tag{101}$$

with the typical value

$$p_{2N}^{typ}(x) = \frac{\pi_1(x) + \pi_2(x)}{2} \tag{102}$$

Its generating function is simply the products of the generating functions of Eq. 100

$$\mathcal{Z}_{2N}[\kappa(.)] = \mathcal{Z}_N^{(1)}[\kappa(.)] \mathcal{Z}_N^{(2)}[\kappa(.)] = \left(\int dx \pi_1(x) e^{\kappa(x)}\right)^N \left(\int dx \pi_2(x) e^{\kappa(x)}\right)^N = e^{N(\Phi_1[\kappa(.)] + \Phi_2[\kappa(.)])} \equiv e^{2N\Phi[\kappa(.)]} \tag{103}$$

i.e. the scaled cumulant generating function follows the renormalization rule

$$\Phi[\kappa(.)] = \frac{\Phi_1[\kappa(.)] + \Phi_2[\kappa(.)]}{2} \tag{104}$$

In particular, when the two sets are drawn with the same probability distribution $\pi_1 = \pi_2 = \pi$, the scaled cumulant generating function $\Phi[\kappa(.)]$ is exactly conserved along the RG flow.

In terms of large deviations form of Eq. 99, the probability of the histogram of Eq. 101

$$\mathcal{P}_{2N}[p_{2N}(.)] = \int \mathcal{D}p_N^{(1)}(.)\mathcal{P}_N^{(1)}[p_N^{(1)}(.)] \int \mathcal{D}p_N^{(2)}(.)\mathcal{P}_N^{(2)}[p_N^{(2)}(.)]\delta\left(p_{2N}(.) - \frac{p_N^{(1)}(.) + p_N^{(2)}(.)}{2}\right)$$

$$\underset{N\to+\infty}{\simeq} \int \mathcal{D}p_N^{(1)}(.) \int \mathcal{D}p_N^{(2)}(.)\delta\left(1 - \int dx p_N^{(1)}(x)\right)\delta\left(1 - \int dx p_N^{(2)}(x)\right)\delta\left(p_{2N}(.) - \frac{p_N^{(1)}(.) + p_N^{(2)}(.)}{2}\right)$$

$$e^{N\left[-S^{rel}(p_N^{(1)}(.)|\pi_1(.)) - S^{rel}(p_N^{(2)}(.)|\pi_2(.))\right]}$$

(105)

corresponds to the optimization of the function $\left[-S^{rel}(p_N^{(1)}(.)|\pi_1(.)) - S^{rel}(p_N^{(2)}(.)|\pi_2(.))\right]$ in the exponential in the presence of the constraints that can be taken into account via Lagrange multipliers. One obtains the optimal solution

$$p_N^{(1)}(x) = p_{2N}(x)\frac{2\pi_1(x)}{\pi_1(x) + \pi_2(x)}$$

$$p_N^{(2)}(x) = p_{2N}(x)\frac{2\pi_2(x)}{\pi_1(x) + \pi_2(x)}$$

(106)

and the corresponding sum of the relative entropies in the exponential

$$N\left[-S^{rel}(p_N^{(1)}(.)|\pi_1(.)) - S^{rel}(p_N^{(2)}(.)|\pi_2(.))\right] = N\left[-\int dx p_N^{(1)}(x)\ln\left(\frac{p_N^{(1)}(x)}{\pi_1(x)}\right) - \int dx p_N^{(2)}(x)\ln\left(\frac{p_N^{(2)}(x)}{\pi_2(x)}\right)\right]$$

$$= -2N\int dx p_{2N}(x)\ln\left(\frac{2p_{2N}(x)}{\pi_1(x) + \pi_2(x)}\right) = -2NS^{rel}\left(p_{2N}(x)|p_{2N}^{typ}(x) = \frac{\pi_1(x) + \pi_2(x)}{2}\right)$$

(107)

coincides with the relative entropy of the histogram $p_{2N}(x)$ with respect to its typical value of Eq. 102 as it should for consistency. In particular, when the two sets are drawn with the same probability distribution $\pi_1 = \pi_2 = \pi$, the optimal solution to produce an anomalous empirical histogram $p_{2N}(x)$ consists in choosing the same anomalous empirical histogram for the two subsets (Eq. 106).

C. Renormalization for the large deviations of the empirical maximum

For each set $b = 1, 2$, the cumulative probability distribution of the empirical maximum of Eq. 32 can be interpreted as an exact large deviation form

$$\int_0^x dx' \mathcal{X}_N^{(b)}(x') = [1 - C_b(x)]^N = e^{-NJ_b(x)}$$

(108)

where the rate function reads

$$J_b(x) = -\ln\left[1 - C_b(x)\right]$$

(109)

in terms of the complementary cumulative distribution function (Eq 15) associated to each distribution $\pi_b(x)$

$$C_b(x) \equiv \int_x^{+\infty} dx' \pi_b(x')$$

(110)

The empirical maximum of the $(2N)$ variables is of course the maximum of the two maximal values associated to the two sets of Eqs 97 and 98

$$x_{2N}^{max} \equiv \max\left(x_N^{max(1)}, x_N^{max(2)}\right)$$

(111)

So the corresponding cumulative distribution

$$\int_0^x dx' \mathcal{X}_{2N}(x') = \left[\int_0^x dx' \mathcal{X}_N^{(1)}(x')\right]\left[\int_0^x dx' \mathcal{X}_N^{(2)}(x')\right] = e^{-N[J_1(x)+J_2(x)]} \equiv e^{-2NJ(x)}$$

(112)

is written exactly in a large deviation form with the rate function

$$J(x) = \frac{J_1(x) + J_2(x)}{2} = -\frac{\ln\left[1 - C_1(x)\right] + \ln\left[1 - C_2(x)\right]}{2} \tag{113}$$

When the two sets are drawn with the same probability distribution $\pi_1 = \pi_2 = \pi$, one obtains that the rate function $J(x)$ is exactly conserved along the RG flow.

D. Renormalization for the large deviations of the empirical average

1. Case of standard large deviations

In the case of standard large deviations, each set $b = 1, 2$ is described by the large deviation form of Eq. 75) for its empirical average a_b

$$\mathcal{A}_N^{(b)}(a_b) \underset{N \to \infty}{\simeq} e^{-N I_b(a_b)} \tag{114}$$

where the rate function $I_b(a_b)$ is the Legendre transform of the scaled cumulant generating function of Eq. 76

$$\phi_b(k) = \ln\left(\int_0^{+\infty} dx \pi_b(x) e^{kx}\right) \tag{115}$$

involved in the generating function

$$Z_N^{(b)}(k) \equiv \int da_b \mathcal{A}_N^{(b)}(a_b) e^{Nka_b} \equiv e^{N\phi_b(k)} \tag{116}$$

Via the merging of the data of Eqs 97 and 97, the empirical average of the $(2N)$ variables is simply the average of the two empirical averages of the two sets $b = 1, 2$

$$a \equiv \frac{1}{2N} \sum_{i=1}^{2N} x_i = \frac{a_1 + a_2}{2} \tag{117}$$

Its generating function is simply the products of the generating functions of Eq. 116

$$\begin{aligned}
Z_{2N}(k) &\equiv \int da \mathcal{A}_{2N}(a) e^{2Nka} = \int da_1 \mathcal{A}_N^{(1)}(a_1) \int da_2 \mathcal{A}_N^{(2)}(a_2) \, e^{2Nk(a_1 + a_2)} \\
&= Z_N^{(1)}(k) Z_N^{(2)}(k) = e^{N[\phi_1(k) + \phi_2(k)]} \equiv e^{2N\phi(k)}
\end{aligned} \tag{118}$$

so the renormalization rule for the scaled cumulant generating function is simply

$$\phi(k) = \frac{\phi_1(k) + \phi_2(k)}{2} \tag{119}$$

In particular, when the two sets are drawn with the same probability distribution $\pi_1 = \pi_2 = \pi$, the scaled cumulant generating function $\phi(k)$ is exactly conserved along the RG flow.

In terms of large deviations form of Eq. 114, the probability of the empirical average reads for this case $\pi_1 = \pi_2 = \pi$

$$\mathcal{A}_{2N}(a) = \int da_1 \mathcal{A}_N(a_1) \int da_2 \mathcal{A}_N(a_2) \delta\left(a - \frac{a_1 + a_2}{2}\right) \underset{N \to +\infty}{\simeq} \int da_1 e^{-N[I(a) - I(2a - a_1)]} \tag{120}$$

The saddle-point evaluation of this integral requires to find the maximum of the function in the exponential

$$\mathcal{L}(a_1) = -I(a_1) - I(2a - a_1) \tag{121}$$

The vanishing of the first derivative

$$0 = \frac{\partial \mathcal{L}(a_1)}{\partial a_1} = -I'(a_1) + I'(2a - a_1) \tag{122}$$

gives the symmetric solution $a_1 = a = 2a - a_1 = a_2$ which is indeed a maximum if the second derivative is negative

$$0 = \frac{\partial^2 \mathcal{L}(a_1)}{\partial^2 a_1}\Big|_{a_1=a} = [-I''(a_1) - I''(2a - a_1)]_{a_1=a} = -2I''(a) < 0 \tag{123}$$

For instance for the gamma distribution of parameter ν of Eq. 77 the second derivative of the rate function of Eq. 80

$$I_\nu''(a) = \frac{\nu}{a^2} > 0 \tag{124}$$

satisfies this condition.

2. Case of large deviations with asymmetric scaling

It is now interesting to compare the above discussion with the case of the large deviations for stretched exponential decay $0 < \alpha < 1$ that display the asymmetric scaling (Eqs 84 and 87)

$$\mathcal{A}_N(a) \underset{N\to+\infty}{\propto} e^{-NI_-(a)} \quad \text{for} \quad a \leq \overline{x}$$
$$\mathcal{A}_N(a) \underset{N\to+\infty}{\propto} e^{-N^\alpha I_+(a)} \quad \text{for} \quad a > \overline{x} \tag{125}$$

Then Eq. 120 is replaced by the sum of four possible contributions of various orders with respect to N

$$\begin{aligned}
\mathcal{A}_{2N}(a) &= \int da_1 \left[e^{-NI_-(a_1)}\theta(a_1 \leq \overline{x}) + e^{-N^\alpha I_+(a_1)}\theta(\overline{x} < a_1) \right] \\
&\quad \left[e^{-NI_-(2a-a_1)}\theta(2a - \overline{x} \leq a_1) + e^{-N^\alpha I_+(2a-a_1)}\theta(a_1 < 2a - \overline{x}) \right] \\
&= \int da_1 e^{-N[I_-(a_1)+I_-(2a-a_1)]}\theta(2a - \overline{x} \leq a_1 \leq \overline{x}) \\
&\quad + \int da_1 e^{-NI_-(a_1)-N^\alpha I_+(2a-a_1)}\theta(a_1 \leq \overline{x})\theta(a_1 < 2a - \overline{x}) \\
&\quad + \int da_1 e^{-N^\alpha I_+(a_1)-NI_-(2a-a_1)}\theta(\overline{x} < a_1)\theta(2a - \overline{x} \leq a_1) \\
&\quad + \int da_1 e^{-N^\alpha[I_+(a_1)+I_+(2a-a_1)]}\theta(\overline{x} < a_1 < 2a - \overline{x}) \tag{126}
\end{aligned}$$

The fourth contribution of order e^{-N^α} will be the leading contribution whenever the domain of integration for a_1 is not empty, i.e. in the region $a > \overline{x}$: then the saddle-point evaluation requires the maximization of the function involving the rate function $I_+(a) = (a - \overline{x})^\alpha$ of Eq. 88

$$\mathcal{L}(a_1) = -I_+(a_1) - I_+(2a - a_1) \tag{127}$$

However the symmetric solution $a_1 = a = 2a - a_1 = a_2$ is a minimum here as a consequence of the sign of the second derivative for any $0 < \alpha < 1$

$$\mathcal{L}''(a) = -2I_+''(a) = 2\alpha(1 - \alpha)(a - \overline{x})^{\alpha-2} > 0 \tag{128}$$

The maximization of Eq 121 occurs instead at the boundaries $a_1 = \overline{x}$ and $a_2 = 2a - \overline{x}$ (or vice-versa) and one obtains the leading contribution in the region $a > \overline{x}$

$$\mathcal{A}_{2N}(a) \simeq e^{-N^\alpha I_+(2a-\overline{x})} = e^{-N^\alpha(2a-2\overline{x})^\alpha} = e^{-(2N)^\alpha(a-\overline{x})^\alpha} = e^{-(2N)^\alpha I_+(a)} \quad \text{for} \quad a > \overline{x} \tag{129}$$

as it should for consistency with Eq. 125 in the region $a > \overline{x}$.

VIII. CONCLUSION

In this paper, we have revisited the empirical observables based on independent random variables, namely the empirical maximum, the empirical average, the empirical non-integer moments or other additive empirical observables,

in order to describe the cases where asymmetric large deviations occur. We have stressed the analogy with equilibrium statistical mechanics : the Sanov theorem for the large deviations of the empirical histogram that involves as rate function the relative entropy with respect to the true probability distribution has been taken as the unifying starting point. The various empirical observables have been then analyzed by optimizing this relative entropy in the presence of the appropriate constraints. Finally, we have discussed the physical meaning of large deviations rate functions from the renormalization perspective.

---

[1] J.M. Luck, Aléa Saclay (1992) "Systèmes désordonnés unidimensionnels".
[2] A. Crisanti, G. Paladin, A. Vulpiani, "Products of random matrices in statistical physics", Springer Verlag (1993).
[3] E.J. Gumbel, " Statistics of extreme" (Columbia University Press, NY 1958).
[4] J. Galambos, " The asymptotic theory of extreme order statistics" (Krieger , Malabar, FL 1987).
[5] J. P. Bouchaud and M. Mézard, J. Phys. A: Math. Gen. 30, 7997 (1997)
[6] M. Clusel and E. Bertin, Int. J. Mod. Phys. B 22, 3311 (2008)
[7] J.Y. Fortin and M. Clusel, J. Phys. A: Math. Theor. 48 183001 (2015).
[8] Y. Oono, Progress of Theoretical Physics Supplement 99, 165 (1989).
[9] R.S. Ellis, Physica D 133, 106 (1999).
[10] H. Touchette, Phys. Rep. 478, 1 (2009).
[11] B. Derrida, J. Stat. Mech. P07023 (2007).
[12] R J Harris and G M Schütz, J. Stat. Mech. P07020 (2007).
[13] E.M. Sevick, R. Prabhakar, S. R. Williams, D. J. Searles, Ann. Rev. of Phys. Chem. Vol 59, 603 (2008).
[14] H. Touchette and R.J. Harris, chapter "Large deviation approach to nonequilibrium systems" of the book "Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond", Wiley 2013.
[15] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim Rev. Mod. Phys. 87, 593 (2015).
[16] A. Lazarescu, J. Phys. A: Math. Theor. 48 503001 (2015).
[17] A. Lazarescu, J. Phys. A: Math. Theor. 50 254004 (2017).
[18] A. de La Fortelle, PhD (2000).
[19] V. Lecomte, PhD Thesis (2007).
[20] R. Chétrite, PhD Thesis (2008).
[21] B. Wynants, PhD Thesis (2010), arXiv:1011.4210.
[22] R. Chétrite, HDR Thesis (2018).
[23] D.S. Dean and S.N. Majumdar, Phys. Rev. Lett. 97, 160201 (2006) and Phys. Rev. E 77, 041108 (2008).
[24] G. Biroli, J. P. Bouchaud and M. Potters, JSTAT P07019 (2007).
[25] S. N. Majumdar and M. Vergassola, Phys. Rev. Lett. 102, 060601 (2009).
[26] S. N. Majumdar, chapter in the book " The Oxford Handbook of Random Matrix Theory" ed. by G. Akemann, J. Baik and P. Di Francesco, Oxford University Press (2011).
[27] S. N. Majumdar and G. Schehr, J. Stat. Mech. P01012 (2014)
[28] S. N. Majumdar, G. Schehr, D. Villamaina and P. Vivo, J. Phys. A: Math. Theor. 46, 022001 (2013).
[29] P. Vivo, S. N. Majumdar and O. Bohigas, Phys. Rev. B 81, 104202 (2010)
[30] S. N. Majumdar, C. Nadal, A. Scardicchio and P. Vivo, Phys. Rev. Lett. 103, 220603 (2009) and Phys. Rev. E 83, 041105 (2011)
[31] C. Nadal, S. N. Majumdar and M. Vergassola, Phys. Rev. Lett. 104, 110501 (2010) and J. Stat. Phys. 142, 403 (2011).
[32] S. N. Majumdar and P. Vivo, Phys. Rev. Lett. 108, 200601 (2012)
[33] C. Texier and S. N. Majumdar, Phys. Rev. Lett. 110, 250602 (2013)
[34] R. Marino, S. N. Majumdar, G. Schehr and P. Vivo, Phys. Rev. Lett. 112, 254101 (2014).
[35] R. Marino, S. N. Majumdar, G. Schehr and P. Vivo, J. Phys. A: Math. Theor. 47, 055001 (2014)
[36] A. Grabsch and C. Texier, J. Phys. A : Math. Theor. 49, 465002 (2016)
[37] F. D. Cunden, P. Facchi and P. Vivo, J. Phys. A: Math. Theor. 49, 135202 (2016)
[38] A. Grabsch, S. N. Majumdar and C. Texier, J. Stat. Phys. 167(2), 234 (2017).
[39] B. Lacroix-A-Chez-Toine, J. A. Monroy Garzon, C. S. H. Calva, I. Perez Castillo, A. Kundu, S. N. Majumdar and G. Schehr , arXiv:1904.01813.
[40] B. Derrida and J.L. Lebowitz, Phys. Rev. Lett. 80, 209 (1998)
[41] T. Halpin-Healy and Y.-C. Zhang, Phys. Repts., **254**, 215 (1995).
[42] R. Giuliano and C. Macci, Communications in Statistics - Theory and Methods, 43:6, 1077-1098 (2014)
[43] P. Vivo, Eur. J. Phys. 36 055037 (2015).
[44] S.V. Nagaev, Theory of Probability and its Applications, 14(1):5164. (1969)
[45] M. R. Evans and S. N. Majumdar, J. Stat. Mech. P05004 (2008).
[46] N. Gantert, K. Ramanan and F. Rembart, Electronic Communications in Probability 19 , 41, 1 (2014).
[47] C. Godrèche, arXiv:1812.02513.
[48] J. Szavits-Nossan, M. R. Evans and S. N. Majumdar, Phys. Rev. Lett. 112, 020602 (2014) and J. Phys. A: Math. Theor. 47 (2014) 455004 (2014).

[49] F. Corberi and A. Sarracino, arXiv:1903.09547.
[50] C. Monthus, arXiv:1903.05899
[51] G. Ben Arous, L.V. Bogachev and S.A. Molchanov 132 (4) 579 (2005).
[52] G. JonaLasinio, Nuovo Cimento 26, 98 (1975).
[53] G. JonaLasinio, Phys. Rep. 352, 439 (2001).
[54] I. Calvo, J. C. Cuchi, J. G. Esteve and F. Falceto, J. Stat. Phys. 141, 409 (2010).
[55] G. Gyorgyi, N. R. Moloney, K. Ozogany, and Z. Racz Phys. Rev. Lett. 100, 210601 (2008).
[56] G. Györgyi, N. R. Moloney, K. Ozogány, Z. Rácz, and M. Droz, Phys. Rev. E 81, 041135 (2010)
[57] E. Bertin and G. Györgyi, J. Stat. Mech. P08022 (2010).
[58] I. Calvo, J. C. Cuchi, J. G. Esteve, and F. Falceto, Phys. Rev. E 86 041109 (2012)
[59] F. Angeletti, E. Bertin, and P. Abry, J. Phys. A: Math. Theor. 45 115004 (2012).