Central limit theorem for a class of globally correlated random variables

Adrián A. Budini
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Centro Atómico Bariloche, Avenida E. Bustillo Km 9.5, (8400) Bariloche, Argentina
and Universidad Tecnológica Nacional (UTN-FRBA), Fanny Newbery 111, (8400) Bariloche, Argentina
(Received 22 February 2016; revised manuscript received 26 April 2016; published 9 June 2016)

The standard central limit theorem with a Gaussian attractor for the sum of independent random variables may lose its validity in the presence of strong correlations between the added random contributions. Here, we study this problem for similar interchangeable globally correlated random variables. Under these conditions, a hierarchical set of equations is derived for the conditional transition probabilities. This result allows us to define different classes of memory mechanisms that depend on a symmetric way on all involved variables. Depending on the correlation mechanisms and statistics of the single variables, the corresponding sums are characterized by distinct probability densities. For a class of urn models it is also possible to characterize their domain of attraction, which, as in the standard case, is parametrized by the probability density of each random variable. Symmetric and asymmetric q-Gaussian attractors ($q < 1$) arise in a particular two-state case of these urn models.

DOI: 10.1103/PhysRevE.93.062114

I. INTRODUCTION

The standard central limit theorem (CLT) is a cornerstone of probability theory [1–4]. It establishes that a sum of independent (identical) random variables, under suitable rescaling, converges to a Gaussian distribution. It plays a fundamental role in the formulation of statistical thermodynamics and also provides a rigorous basis for assuming Gaussian statistics for describing fluctuations in equilibrium and nonequilibrium systems.

There exist a few remarkable examples where the standard CLT was generalized. The Gaussian attractor arises when considering independent random variables with a finite second moment. As is well known, when this condition is lifted up the attractor becomes a Levy distribution [5]. In addition, Gumbel distribution arises from the study of extreme value statistics and describes the fluctuations of the largest value in a large set of identically distributed independent random variables [6]. Interestingly, this problem can in general be related with the statistics of sums of correlated random variables [7]. Departure from Gaussian statistics was also analyzed for global correlations where the characteristic function of the total sum is defined by a nonmultiplicative Fourier structure [8].

Recently it was argued that the presence of global correlations in stationary equilibrium and nonequilibrium systems is a situation where nonextensive statistical mechanics may apply [9–12]. Consistently, many theoretical studies were devoted to finding global memory mechanisms that lead to attractors defined by q-Gaussian probability densities [13–17]. These statistical objects also arise from maximizing Tsallis entropy [9], from superstatistical models [18], as well as from specific transformations of Gamma distributed random variables [19].

Global correlations are a mechanism that may lead to departures from Gaussian statistics. Nevertheless, establishing a generalization of the CLT on the basis of only this feature is a formidable task. In fact, to our knowledge, there does not exist general rigorous mathematical criteria for splitting correlations in weak ones (leading to Gaussian statistics) and stronger ones (departure from normal distribution). Therefore, as in the previous literature [7–17], one is naturally forced to study particular cases. Of special interest is to find generalizations that rely on simple correlation mechanisms or symmetries, which in turn also allow defining or studying its basin of attraction. In general, this last issue is hard to solve.

In this paper we analyze the departure from the standard CLT for a specific class of global correlations. Similar interchangeable random variables [20–24] are considered. This property or symmetry, originally introduced by de Finetti [20] in probability theory, is defined by random variables whose joint probability density is invariant under arbitrary permutations of its arguments.

The main goal is twofold. First, we give a general characterization of possible correlation mechanisms consistent with interchangeability. This objective is achieved by characterizing the correlations not through the joint probability densities but through the transition probabilities. These functions tell us how the probability density of a given variable depends on the previous values assumed by the rest of the random variables. We demonstrate that these objects obey a set of hierarchical equations that resemble a Chapman-Kolmogorov equation for Markovian chains [2–4]. From this result we construct different correlation models that allow us to achieve the second main goal, that is, the characterization of the departure from Gaussian statistics as well as to study their basin of attraction. For a class of urn models [25–28], we demonstrate that their basin of attraction is as wide as in the standard case. Asymmetric and symmetric q-Gaussian attractors [19] with $q < 1$ arise as a particular two-state case of these urn models.

The paper is outlined as follows. In Sec. II, based on the interchangeability property of the joint probabilities, we derive a hierarchical set of equations for the transition probabilities. Section III is devoted to finding different solutions to the previous equations, which are based on a generalization of Pólya urn scheme [25–28]. In Sec. IV, departure from Gaussian statistics and their basin of attraction are analyzed. In Sec. V we provide the conclusions. In the Appendices we show some calculus details and study other correlation models (additive memory, de Finetti representation).

 ©2016 American Physical Society

II. HIERARCHY OF TRANSITION PROBABILITIES FOR SIMILAR INTERCHANGEABLE RANDOM VARIABLES

We consider a set of n random variables $X_1, X_2, \dots X_n$. For a given n, their statistics is completely characterized by the n-joint probability density $P_n(x_1, x_2, \dots x_n)$, which defines the probability that each variable falls in an infinitesimal range dx_i around x_i. Similar interchangeable variables are defined by the following two symmetries.

Similarity (or scale invariance [9]) means that for any n the following relation is fulfilled:

$$P_{n-1}(x_1, x_2, \dots x_{n-1}) = \int dx_n P_n(x_1, x_2, \dots x_n). \quad (1)$$

Therefore, the joint probability density of the subset of $(n-1)$ random variables coincides with the marginal density corresponding to n variables. This property tells us that by adding a new stochastic variable, $(n-1) \to n$, the marginal statistics of the previous $(n-1)$ variables remains invariant. In general, this property is not valid. In fact, a problem with n arbitrary random variables may not have any relation with a problem with $(n-1)$ variables (see Ref. [16]). Then, the similarity condition Eq. (1) must not be confused with the usual definition of marginal densities [3], which certainly always applies to any fixed $n, P_n^{(k)}(x_1, x_2, \dots x_k) \equiv \int P_n(x_1, \dots x_k, x_{k+1} \dots x_n) dx_{k+1} \dots dx_n$. In terms of them, condition Eq. (1) reads $P_{n-1}(x_1, x_2, \dots x_{n-1}) = P_n^{(n-1)}(x_1, x_2, \dots x_{n-1})$.

Interchangeability is defined by the invariance of the joint probability density under arbitrary permutations of its arguments,

$$P_n(\dots, x_k, \dots, x_l, \dots) = P_n(\dots, x_l, \dots, x_k, \dots), \quad (2)$$

that is, for any k and l in the interval $(1, 2, \dots n)$, the joint probability density does not change under the (arbitrary) interchange $x_k \leftrightarrow x_l$. These relations are assumed valid for all values of n.

Notice that, in particular, the previous two symmetries imply that all random variables $\{X_i\}_{i=1}^n$ are characterized by the same density $P_1(x)$. In fact, for any n and $k = 1, \dots n$, it follows $P_n^{(1)}(x_k) = P_1(x_k)$ where $P_n^{(1)}(x_k) \equiv \int P_n(x_1, \dots x_k, \dots x_n) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n$.

The joint probability density $P_n(x_1, x_2, \dots x_n)$ completely characterizes the random variables $\{X_i\}_{i=1}^n$. Nevertheless, an extra aspect is enlightened by introducing a conditional probability density defined by the relation

$$P_n(x_1, \dots x_n) = P_{n-1}(x_1, \dots x_{n-1}) T_{n-1}(x_1, \dots x_{n-1} | x_n). \quad (3)$$

Hence, the function $T_{n-1}(x_1, \dots x_{n-1} | x_n)$ gives the probability density of the variable X_n *given* that the previous ones $\{X_i\}_{i=1}^{n-1}$ assumed the values $x_1, \dots x_{n-1}$ [29]. By definition, it satisfies the normalization condition $\int dx_n T_{n-1}(x_1, \dots x_{n-1} | x_n) = 1$.

From Eq. (3), iteratively it follows

$$P_n(x_1, \dots x_n) = P_1(x_1) T_1(x_1 | x_2) T_2(x_1, x_2 | x_3) \dots \cdots \times T_{n-1}(x_1, \dots x_{n-1} | x_n). \quad (4)$$

Therefore, the set of functions $T_k(x_1, \dots x_k | x_{k+1})$, with $k = 1, \dots n-1$ provide the same information than the n-joint probability density. Furthermore, from Eq. (4) one can easily read how the correlations between the random variables are built up. In fact, having an explicit expression for the transition probabilities it is possible to numerically generate the corresponding realizations of the correlated variables $\{X_i\}_{i=1}^n$.

The main problem that we solve in this section is to determine which set of transition probabilities are consistent with the similarity and interchangeability properties. *Given* an arbitrary probability density $P_1(x_1)$, the symmetry does not impose any condition. For $n = 2$, interchangeability implies $P_2(x_1, x_2) = P_2(x_2, x_1)$, or equivalently $P_1(x_1) T_1(x_1 | x_2) = P_1(x_2) T_1(x_2 | x_1)$. After integration, and by using the similarity property, it follows the condition

$$\int dx_1 P_1(x_1) T_1(x_1 | x_2) = P_1(x_2). \quad (5)$$

By using a similar procedure, $T_2(x_1, x_2 | x_3)$ must fulfill

$$T_1(x_1 | x_3) = \int dx_2 T_1(x_1 | x_2) T_2(x_1, x_2 | x_3). \quad (6)$$

For higher conditional probabilities densities (see Appendix A), the following relations,

$$T_{n-1}(x_1 \dots x_{n-1} | x_{n+1}) = \int dx_n T_{n-1}(x_1, \dots x_{n-1} | x_n) \times T_n(x_1, \dots x_n | x_{n+1}), \quad (7)$$

must be fulfilled for all values of n. Furthermore, the function $T_n(x_1 \dots x_n | x_{n+1})$ must be *symmetric* in the conditional arguments $x_1 \dots x_n$, that is, it is invariant under arbitrary permutations of its arguments. The hierarchical set of equations defined by Eq. (7) is the main result presented in this section.

If $T_{n-1}(x_1, \dots x_{n-1} | x_{n+1})$, for all values of n, does not depend on the previous values $x_1, \dots x_{n-1}$, it follows $T_n(x_1, \dots x_n | x_{n+1}) = P_1(x_{n+1})$, that is, we recover the case of *independent identical random variables*. Notice that interchangeability implies that $T_{n-1}(x_1 \dots x_{n-1} | x_n)$ depends symmetrically on the previous arguments $x_1 \dots x_{n-1}$. Therefore, transition probabilities that only depend on one previous value, with a dependence that is independent on the number of previous events, $T_{n-1}(x_1 \dots x_{n-1} | x_n) = T_{n-1}(x_{n-1} | x_n) = T(x_{n-1} | x_n)$, are *not consistent* with interchangeability. This case corresponds to *stationary Markov chains*. In fact, the unique transition probability $T(x|y)$ should satisfy [Eq. (5)]

$$\int dx P_1(x) T(x|y) = P_1(y), \quad (8)$$

while by imposing the previous conditions on Eq. (7), it follows

$$T(x|y) = \int dx' T(x|x') T(x'|y). \quad (9)$$

The stationary property is given by Eq. (8), while the Markov property is defined by the *Chapman-Kolmogorov* [2–4] relation Eq. (9).

The inconsistence between stationary Markov chains and interchangeable random variables can be put in evidence as follows. For the former one, the correlations between "neighbor" random variables, X_k and X_{k+d}, only depend through their "distance," d. Nevertheless, this dependence is inconsistent with interchangeability, which in fact implies that correlations between *any* two random variables is always the

same statistical object. On the other hand, both Eqs. (7) and (9) are defined by integral relations. Therefore, as in the standard Chapman-Kolmogorov case [3,30], we conclude that the set of hierarchical Eqs. (7) define a *necessary* condition for similarity and interchangeability symmetries.

In the following section we search solutions of Eq. (7) where the transition probabilities are based on a generalization of Pólya urn scheme. In the Appendices we studied other solutions that also depend in the same manner on all previous values taken by the random variables, that is, global correlations. For example, an additive memory assumption $T_n(x_1 \dots x_n|x_{n+1}) = \mathcal{T}_n(x_1 + x_2 \cdots + x_n|x_{n+1})$ (Appendix B) leads to consistent solutions for Gaussian and classical spin variables. A generalized de Finetti representation is analyzed in Appendix C.

III. URN SCHEMES

Urn models are examples of random variables defined through their transition probabilities [2,25]. They generate interchangeable random variables [26–28]. Below we review these schemes, which gives us the basis for constructing a generalization consistent with interchangeability.

A. Pólya Urn scheme

In the standard Pólya urn scheme [2,25], initially an urn contains many balls that, for example, are characterized by different colors (the random variables). At each step, one determines the color of one ball taken at random and then put into the urn one extra ball (next random variable) of the same color. A similar process can be defined by starting the urn with only one ball, which is a particular case of the Blackwell-MacQueen urn scheme [26,27], being related to the "Chinese restaurant process" [27,28].

In the present context, the Blackwell-MacQueen urn scheme is defined by an arbitrary density $P_1(x)$, while the transition probabilities are

$$T_n(x_1, \dots x_n|x) = \frac{\lambda P_1(x) + \sum_{i=1}^{n} \delta(x - x_i)}{n + \lambda}. \qquad (10)$$

Here, λ is a dimensionless positive parameter, while $\delta(x)$ is the delta Dirac function. When $\lambda \to \infty$, *identical independent random variables* are recovered, while the limit $\lambda \to 0$ leads to a *fully correlated case*, that is, after the first random value the next ones assume the same value.

After a simple algebra it is possible to proof that the set of functions defined by Eq. (10) satisfy Eq. (5), as well as the hierarchical set of conditions corresponding to interchangeability, Eq. (7). The Pólya urn scheme here is defined as the particular case in which the random variables $\{X_i\}_{i=1}^n$ are discrete. Hence, we write

$$P_1(x) = \sum_{\mu=1}^{M} q_\mu \delta(x - x_\mu), \qquad (11)$$

where $\{x_\mu\}_{\mu=1}^M$ is the set of M possible values and $\{q_\mu\}_{\mu=1}^M$ are the corresponding weights (probabilities), with $\sum_{\mu=1}^M q_\mu = 1$. In this case, the transition probabilities Eq. (10) can be written in terms on the number of times n_μ that each value x_μ was assumed previously,

$$T_n(\{x_i\}|x) = \sum_{\mu=1}^{M} \frac{\lambda q_\mu + n_\mu}{n + \lambda} \delta(x - x_\mu), \qquad (12)$$

where $T_n(\{x_i\}|x) \equiv T_n(x_1, \dots x_n|x)$. Notice that the set of numbers $\{n_\mu\}_{\mu=1}^M$ that the random values $\{X_i\}_{i=1}^n$ assumed the values $\{x_\mu\}_{\mu=1}^M$ satisfy the relation $n = \sum_{\mu=1}^M n_\mu$.

The correlation mechanism associated with Eq. (12) can be read in the following way. With probability $\lambda/(n + \lambda)$ the random variable X_{n+1} is drawn randomly in agreement with the density $P_1(x)$, Eq. (11). Hence, independently of the previous history, it assumes the value x_μ with probability q_μ. Alternatively, with probabilities $n_\mu/(n + \lambda)$, which depends on all previous history, it assumes the value x_μ. The parameter λ measure the weigh of both options.

A central property of the transition probability Eq. (12) is given by its limit in the regime $n \to \infty$, where it becomes similar to that of independent variables. This result, which was characterized previously in the literature [26,27], here can be written as

$$T_\infty(\{x_i\}|x) = \sum_{\mu=1}^{M} F_u \, \delta(x - x_\mu), \qquad (13)$$

where the set of weights $\{F_\mu\}_{\mu=1}^M$ are defined by

$$F_\mu \equiv \lim_{n\to\infty} \frac{\lambda q_\mu + n_\mu}{n + \lambda}, \quad 0 \leqslant F_\mu \leqslant 1, \qquad (14)$$

and consistently satisfy $\sum_{\mu=1}^M F_\mu = 1$. These weights (probabilities) are different for each realization, that is, they are random variables. Their probability density [26,27] is presented in the next section [see Eq. (33)]. Notice that when the previous regime is achieved, *each* realization is in fact equivalent to that of independent random variables.

In order to get a clear understanding of the processes defined by Eq. (12) and its properties, in Fig. 1 we plotted a set of realizations for the random variables $\{X_i\}_{i=1}^n$ (upper panels). They correspond to classical spin variables, that is, we take $x_\mu = \pm 1$ and $M = 2$. For clarity, each value of X_i is continued in the real interval $(i - 1, i)$.

The first value, X_1, is chosen in agreement with $P_1(x)$, Eq. (11). The next values $\{X_i\}_{i=2}^n$ follow from the transition probability $T_n(\{x_i\}|x)$, Eq. (12). We also plotted this object as a function of n and for each value of $x = \pm 1$ (lower panels).

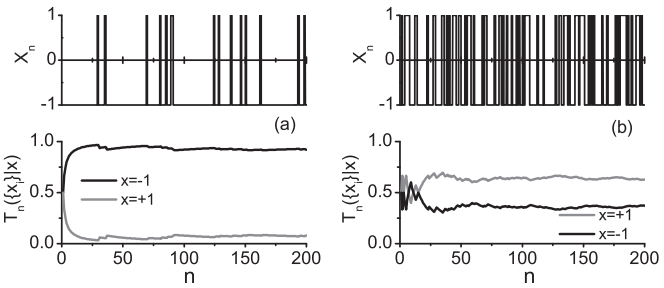

FIG. 1. Two realizations [(a) and (b)] for a set of classical spin variables $\{x_\mu\} = \{+1, -1\}$ obtained from Eq. (12) ($M = 2$). The lower panels correspond to the transition probabilities $T_n(\{x_i\}|x)$ ($x = 1$ gray line, $x = -1$ black line). The parameters are $q_+ = q_- = 1/2$ and $\lambda = 2$.

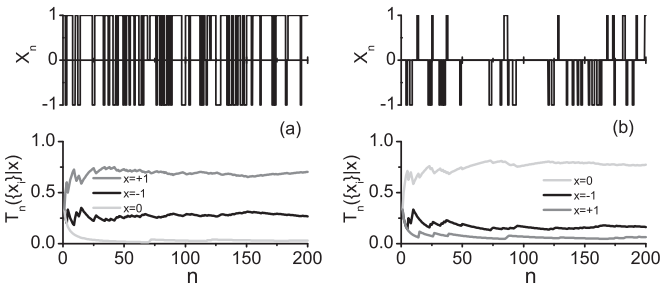

FIG. 2. Two realizations [(a) and (b)] for a set of three-state random variables $\{x_\mu\} = \{+1,0,-1\}$. They follow from the transition probability $T_n(\{x_i\}|x)$ Eq. (12) with $M = 3$, which is plotted in the lower panels ($x = 1$ gray line, $x = -1$ black line, $x = 0$ light gray line). The parameters are $q_+ = q_- = q_0 = 1/3$ and $\lambda = 3$.

Given the previous history, each curve defines the probability for the next variable. Therefore, they are random objects.

In agreement with Eq. (13), we found that for increasing n the transition probabilities always saturate to stationary values [Eq. (14)]. Therefore, when this regime is achieved, each realization is equivalent to that of independent random variables. Nevertheless, the stationary values reached by the transition probabilities are different for each realization.

In the realization of Fig. 1(a) the stationary transition probability for the state -1 is larger than for the state $+1$. Consistently, the state -1 is taken much more frequently, feature clearly visible in the upper panel. In Fig. 1(b) the difference between the stationary values is much smaller, inducing a more "noisy" realization.

In Fig. 2 we plot a set of realizations obtained from the transition probability Eq. (12) for random variables characterized by three states, $M = 3$, with $x_\mu = +1,0,-1$. Similar to the case of two-level variables, for increasing n the transition probabilities reach stationary values, which are different and random for each realization. Therefore, in this regime the realizations are also equivalent to that of identical independent random variables. In Fig. 2(a) the random variables almost always assume the values $x = \pm 1$. This happens because the stationary value of the transition probability corresponding to the state $x = 0$ is much smaller than the other two, $x = \pm 1$. Instead, in Fig. 2(b) the state $x = 0$ has the larger stationary transition probability. Consistently, the states $x = \pm 1$ appear sparsely.

B. Composed Pólya urn scheme

Here, we introduce a generalization of the previous urn scheme that is also consistent with interchangeability. We consider *nondiscrete* random variables with *arbitrary* probability density $P_1(x)$. The domain Ω of each variable X_i, that is, the domain of $P_1(x)$, is split in a finite set of disjoint subdomains $\{\Omega_\mu\}_{\mu=1}^M$ such that the total domain is their union, $\Omega = \cup\Omega_\mu$. To each Ω_μ we associate a probability density $p_\mu(x)$ that assumes values on Ω and is normalized as $\int_\Omega p_\mu(x)dx = 1$. Under these definitions, we propose the transition probability density,

$$T_n(\{x_i\}|x) = \frac{\lambda P_1(x) + \sum_{\mu=1}^{M} p_\mu(x) \sum_{i=1}^{n} \int_{\Omega_\mu} dy\delta(y - x_i)}{n + \lambda}. \quad (15)$$

As before, λ is a free parameter and $\{x_i\} = x_1, x_2, \ldots x_n$ is the previous trajectory.

The integral contributions,

$$n_\mu \equiv \sum_{i=1}^{n} \int_{\Omega_\mu} dy\delta(y - x_i), \quad \sum_{\mu=1}^{M} n_\mu = n, \quad (16)$$

give the number of times the variables $\{x_i\}_{i=1}^n$ fell in the subdomain Ω_μ. Therefore, we can write

$$T_n(\{x_i\}|x) = \frac{\lambda P_1(x) + \sum_{\mu=1}^{M} p_\mu(x) n_\mu}{n + \lambda}. \quad (17)$$

The correlation dynamics induced by Eq. (15) is then clear. With probability $\lambda/(n + \lambda)$ the next variable, independently of the previous history, is chosen in agreement with $P_1(x)$. On the other hand, with probabilities $n_\mu/(n + \lambda)$, the next value is chosen in agreement with the arbitrary density $p_\mu(x)$.

It is simple to check that the transition probability density Eq. (15), for arbitrary subdomains $\{\Omega_\mu\}_{\mu=1}^M$ and densities $\{p_\mu(x)\}_{\mu=1}^M$, is normalized $\int_\Omega dx T_n(\{x_i\}|x) = 1$, and positive defined $T_n(\{x_i\}|x) \geqslant 0$. Nevertheless, in order to be consistent with the similarity and interchangeability symmetries the densities $\{p_\mu(x)\}_{\mu=1}^M$ *cannot* be arbitrary. In fact, these symmetries are valid only when

$$p_\mu(x) = P_1(x) \frac{\theta_{\Omega_\mu}(x)}{\int_{\Omega_\mu} P_1(x')dx'}, \quad (18)$$

where we defined the region indicator

$$\theta_{\Omega_\mu}(x) \equiv \begin{cases} 1 & \text{if } x \in \Omega_\mu \\ 0 & \text{if } x \notin \Omega_\mu \end{cases}. \quad (19)$$

These expressions follow after imposing over Eq. (15) the validity of the hierarchical relations Eq. (7), which lead to the conditions $\sum_{\mu=1}^M p_\mu(x) \int_{\Omega_\mu} P_1(y)dy = P_1(x)$ and $\int_{\Omega_\mu} p_{\mu'}(x)dx = \delta_{\mu\mu'}$. The former constraint tells us that the set $\{p_\mu(x)\}_{\mu=1}^M$, under appropriate weights, recover the density $P_1(x)$, while the last one implies that $p_\mu(x)$ is not null only on the subdomain Ω_μ. Equation (18) satisfies both conditions.

Interestingly, from the previous solutions for $\{p_\mu(x)\}_{\mu=1}^M$, Eq. (18), we can write the probability density of each variable as

$$P_1(x) = \sum_{\mu=1}^{M} q_\mu p_\mu(x), \quad (20)$$

where the positive weights are

$$q_\mu \equiv \int_{\Omega_\mu} P_1(x)dx, \quad (21)$$

and fulfill $\sum_{\mu=1}^M q_\mu = 1$. In consequence, by using Eq. (17), the transition probability Eq. (15) becomes

$$T_n(\{x_i\}|x) = \sum_{\mu=1}^{M} \frac{\lambda q_\mu + n_\mu}{n + \lambda} p_\mu(x). \quad (22)$$

This final expression is the main result of this section.

Equation (22) can be read as an independent statistical *composition* of the Pólya urn scheme, Eq. (12), and the set

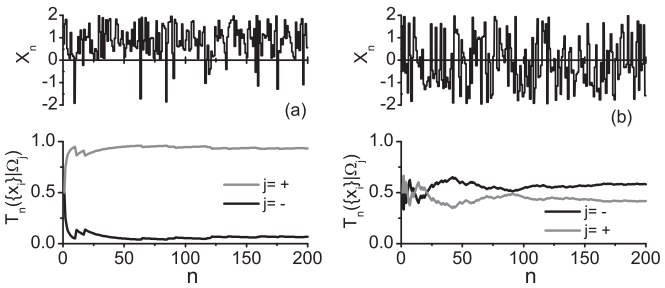

FIG. 3. Two realizations [(a) and (b)] of the composed Pólya urn scheme [Eq. (22)] defined by Eqs. (23) and (24). The lower panels correspond to the transition probabilities $T_n(\{x_i\}|\Omega_\pm)$ associated to the subdomains Ω_+ (gray line) and Ω_- (black line), each one having weights $q_+ = q_- = 1/2$ (see text). The parameter is $\lambda = 2$.

of probability densities $\{p_\mu(x)\}_{\mu=1}^M$. In fact, both expressions are related by the replacements $\delta(x - x_\mu) \leftrightarrow p_\mu(x)$, while the weights in the single density Eq. (11) here follows from Eqs. (20) and (21). Hence, each subdomain Ω_μ can be associated to the states x_μ [Eq. (11)]. Nevertheless, instead of the value x_μ, here the next variable assumes a random value distributed over the subdomain Ω_μ with probability density $p_\mu(x)$.

As an example, we take the uniform probability density,

$$P_1(x) = \frac{1}{4}, \quad -2 \leqslant x \leqslant 2, \tag{23}$$

and $P_1(x) = 0$ if $x \notin [-2,2]$. Therefore, each variable only assumes random values over the real interval $\Omega = [-2,2]$. The composed urn scheme is completely characterized after defining the subdomains $\{\Omega_\mu\}_{\mu=1}^M$. We consider only two, Ω_+ and Ω_-, defined as $\Omega_+ = [0,2]$ and $\Omega_- = [-2,0)$, respectively. Notice that $\Omega = \Omega_+ \cup \Omega_-$. The associated probability densities, from Eqs. (18) and (19), becomes

$$p_+(x) = \frac{1}{2}, \quad 0 \leqslant x \leqslant 2, \tag{24a}$$

$$p_-(x) = \frac{1}{2}, \quad -2 \leqslant x < 0. \tag{24b}$$

Notice that the underlying discrete process that decides which probability density is chosen, $p_+(x)$ or $p_-(x)$, is equivalent to that plotted in Fig. 1. In fact, from Eq. (21) if follows $q_+ = q_- = 1/2$.

In Fig. 3 we plot a set of realizations corresponding to the previous definitions. In contrast to the previous figures, here the random variables assume values over the real interval $[-2,2]$. In the lower panels we plot the underlying transition probability $T_n(\{x_i\}|\Omega_j)$, which governs which subdomain ($j = \pm$) is occupied in the next step. Consistently, its behavior is similar to that of Fig. 1. In Fig. 3(a) the subdomain Ω_+ has a higher stationary probability and, consistently, the realization take most of its values in the interval $[0,2]$. In Fig. 3(b) both subdomains have similar stationary values. Hence, the realization looks like a random signal in the full domain $[-2,2]$. Additionally, by averaging over realizations we checked that the probability density of each variable $\{X_i\}_{i=1}^n$ is given by $P_1(x)$, Eq. (23).

IV. STATISTICS OF THE SUM VARIABLE

In the previous section (and in the Appendices) we described different memory mechanism and statistics consistent with the required symmetries. Here, we study departures with respect to the standard CLT when considering such kinds of globally correlated variables.

We consider the normalized random variable defined by the limit

$$W = \lim_{n\to\infty} W_n = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{n} X_i \right). \tag{25}$$

Notice that in contrast with the standard CLT [1–4], instead of $1/\sqrt{n}$, here the normalization is $1/n$. We choose this factor because all studied models, depending on their characteristic parameters, are able to reach a full correlated regime where all variables $\{X_i\}_{i=1}^n$ assume the same random value. Hence, in that regime the normalization $1/n$ is the only one that delivers a random variable (W) that (asymptotically) does not depend on n.

The probability density $P(w)$ of W can be written as the following limit, $P(w) = \lim_{n\to\infty} P(w_n)$,

$$P(w) = \lim_{n\to\infty} \int dx_1 \dots dx_n \delta\left(w - \frac{1}{n} \sum_{i=1}^{n} x_i \right) P_n(\{x_i\}), \tag{26}$$

where $P_n(\{x_i\}) \equiv P_n(x_1, \dots x_n)$ is the n-joint probability density. $P(w)$ is denoted as the *attractor* associated to the sum Eq. (25). The sequence of random variables $X_1, X_2, \dots X_n, X_{n+1}, \dots$ is characterized by a given correlation mechanism (like urn schemes), which in turn is univocally defined by their n-joint probability densities. Therefore, the attractor can also be associated to a given set $\{P_n(\{x_i\})\}_{n=1}^\infty$.

For (statistically) different sequences of random variables, the normalized sum [Eq. (25)] may be characterized by the *same* attractor or probability density. The *basin of attraction* of $P(w)$ corresponds to all different sets $\{P_n(\{x_i\})\}_{n=1}^\infty$, which lead to the same probability density (attractor). This definition is formulated in the space of joint probability densities, which can also be defined in terms of conditional objects, Eq. (4). Notice that each set $\{P_n(\{x_i\})\}_{n=1}^\infty$ is a point in the previous space. Equivalently, the basin of attraction can be read as the set of all different sequences of random variables whose normalized sum, Eq. (25), is characterized by the same attractor.

By introducing the Fourier representation of the δ Dirac function, $\delta(x) = (1/2\pi) \int_{-\infty}^{+\infty} e^{-ikx} dk$, the characteristic function $G_w(k)$ of $P(w)$,

$$G_w(k) = \int_{-\infty}^{+\infty} dw e^{ikw} P(w), \tag{27}$$

can be written as $G_w(k) = \lim_{n\to\infty} G_{w_n}(k)$, where

$$G_{w_n}(k) = \int dx_1 \dots dx_n \exp\left(i \frac{k}{n} \sum_{i=1}^{n} x_i \right) P_n(\{x_i\}). \tag{28}$$

In terms of the the multiple Fourier transform of $P_n(\{x_i\})$, that is,

$$G_n(\{k_i\}) = \int dx_1 \dots dx_n \exp\left(i \sum_{i=1}^{n} k_i x_i\right) P_n(\{x_i\}),$$

it follows

$$G_w(k) = \lim_{n\to\infty} G_{w_n}(k) = \lim_{n\to\infty} G_n\left(\left\{k_i = \frac{k}{n}\right\}\right). \tag{29}$$

Below we treat the different cases introduced previously. For clarifying the derivation and understanding of some results, the well-known case of independent variables is reviewed first.

A. Statistical independent variables

Assume the set $\{X_i\}_{i=1}^n$ are independent random variables with probability density $P_1(x)$. Therefore, $P_n(\{x_i\}) = \prod_{i=1}^n P_1(x_i)$. From Eq. (28), it follows

$$G_{w_n}(k) = \left[G_x\left(\frac{k}{n}\right)\right]^n, \tag{30}$$

where $G_x(k)$ is the Fourier transform of $P_1(x)$.

Assuming that the mean value $\bar{x} = \int_{-\infty}^{+\infty} dx\, x P_1(x)$, and standard deviation $\sigma^2 = \int_{-\infty}^{+\infty} dx (x-\bar{x})^2 P_1(x)$, of the density $P_1(x)$ are finite, to leading order in n, it follows $\lim_{n\to\infty} G_{w_n}(k) = e^{ik\bar{x}} \lim_{n\to\infty} \exp\left[-\frac{1}{2}\frac{\sigma^2 k^2}{n}\right]$. After Fourier inversion, $P(w)$ is given by

$$P(w) = \delta(w - \bar{x}). \tag{31}$$

Therefore, the random variable W deterministically assumes the value $\bar{x}$. This result, which can be read as the well-known law of large numbers [1–4], follows from the normalization $1/n$ in Eq. (25). In fact, defining the variable $\sqrt{n}W$, one gets a Gaussian probability density, which in turn corresponds to the standard CLT. The basin or domain of attraction of the normal distribution corresponds to all joint probability densities $\{P_n(\{x_i\})\}_{n=1}^{\infty} = \{\prod_{i=1}^n P_1(x_i)\}_{n=1}^{\infty}$, where the arbitrary density $P_1(x)$ has finite first and second moments. Notice that each set $\{P_n(\{x_i\})\}_{n=1}^{\infty}$ can be parametrized by $P_1(x)$.

Using the same Fourier techniques, we showed that departure with respect to Eq. (31) arises from (correlated) Gaussian variables [see Eq. (B16)] and also in the de Finetti representation [see Eq. (C5)]. In fact, the possibility of achieving a fully correlated regime is enough for warranting departure from a δ distribution.

B. Pólya urn scheme

For increasing n, the transition probability of the Pólya urn scheme converge to that of identical independent random variables, Eqs. (13) and (14). This result was characterized previously in the literature [26,27]. In the previous section it was explicitly shown through numerical simulations [Figs. 1 and 2].

Given that the asymptotic values $\{F_\mu\}_{\mu=1}^M$ [Eq. (14)] are different for each realization, that is, their are random variables, one can define their probability density $D(\{f_\mu\}|\{\lambda_\mu\})$, which depends on the characteristic parameters of the problem, here defined as

$$\lambda_\mu \equiv \lambda q_\mu. \tag{32}$$

Due to the normalization of the weights $\{q_\mu\}_{\mu=1}^M$, it follows $\lambda = \sum_{\mu=1}^M \lambda_\mu$. It is known [26,27] that $D(\{f_\mu\}|\{\lambda_\mu\})$ is a *Dirichlet distribution*,

$$D(\{f_\mu\}|\{\lambda_\mu\}) \equiv \frac{\Gamma(\lambda)}{\prod_{\mu'=1}^M \Gamma(\lambda_{\mu'})} \prod_{\mu=1}^M f_\mu^{\lambda_\mu - 1}, \tag{33}$$

where $\Gamma(x)$ is the Gamma function. $D(\{f_\mu\}|\{\lambda_\mu\})$ is positive for all values of $\{f_\mu\}_{\mu=1}^M$, and normalized as $\int_\Lambda df_1 \dots df_{M-1}\ D(\{f_\mu\}|\{\lambda_\mu\}) = 1$, where Λ is the region defined by $\sum_{\mu=1}^M f_\mu = 1$. In addition, the relation $q_\nu = \int_\Lambda df_1 \dots df_{M-1}\ D(\{f_\mu\}|\{\lambda_\mu\}) f_\nu$ is fulfilled for all $\nu = 1, \dots M$.

With Eq. (33) it is possible to obtain the probability density $P(w)$ of the sum variable W, Eq. (25). Given that asymptotically each realization is equivalent to that of independent random variables, one can associate the probability density $\delta(w - \bar{X}_f)$ to each realization [see Eq. (31)], where $\bar{X}_f = \sum_{\mu=1}^M F_\mu x_\mu$. Now, the final structure of $P(w)$ arises after averaging over realizations. Given that the random variables F_μ obeys the statistics given by Eq. (33), it follows

$$P(w) = \int_\Lambda df_1 \dots df_{M-1}\ \delta(w - \bar{x}_f) D(\{f_\mu\}|\{\lambda_\mu\}), \tag{34}$$

where $\bar{x}_f$ is defined by

$$\bar{x}_f \equiv \sum_{\mu=1}^M f_\mu x_\mu. \tag{35}$$

From the result Eq. (34), in the limit of $\lambda \to \infty$ we consistently recover the independent random variables case, $\lim_{\lambda\to\infty} P(w) = \delta(w - \bar{x}_q)$, where $\bar{x}_q \equiv \sum_{\mu=1}^M q_\mu x_\mu$. In the limit $\lambda \to 0$, which corresponds to the fully correlated case, it follows $\lim_{\lambda\to 0} P(w) = \sum_{\mu=1}^M q_\mu \delta(w - x_\mu)$ [see Eq. (12)].

The final expression, Eq. (34), allow us to characterize the CLT for the Pólya urn scheme. It is valid for any value of M and arbitrary discrete distributions, Eq. (11). For example, for classical spin variables, $x_\mu = \pm 1$, after integration we get ($\lambda_\pm \equiv \lambda q_\pm$)

$$P(w) = \frac{1}{\mathcal{N}}(1+w)^{\lambda_+ - 1}(1-w)^{\lambda_- - 1}, \tag{36}$$

where $\mathcal{N} \equiv 2^{\lambda_+ + \lambda_- - 1}\Gamma(\lambda_+)\Gamma(\lambda_-)/\Gamma(\lambda_+ + \lambda_-)$.

Equation (36) can be read as a Beta [2] or asymmetric q-Gaussian distribution ($q < 1$) [19]. In the symmetric case $\lambda_+ = \lambda_-$, this result was derived previously in the context of a nonextensive thermodynamics approach [13] (see also Refs. [2,25]). We notice that in the present context (correlated random variables) no special role is played by this case ($M = 2$). Criticisms along this line have been introduced previously [31].

In Fig. 4 we obtained numerically $P(w)$ by averaging a set of realizations such as those presented in Fig. 1. Results for different values of λ are presented, while $q_+ = q_- = 1/2$. Independently of the parameter values, we find that Eq. (36) fits the numerical results.

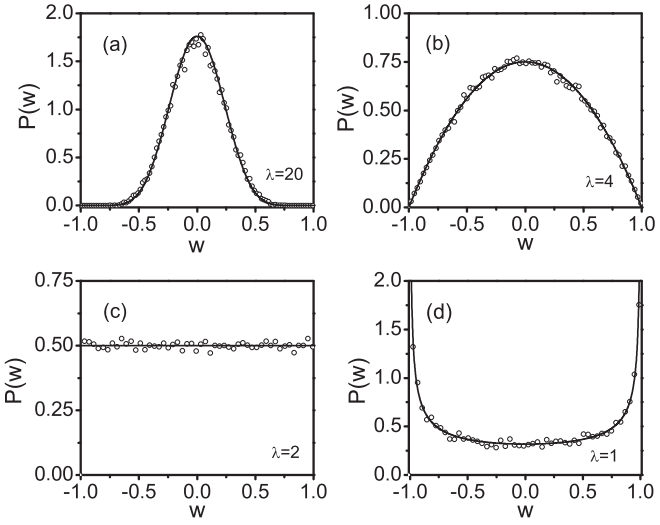

FIG. 4. Probability density of the sum variable W, Eq. (25). The random contributions are classical spin variables, $x_\mu = \pm 1$, defined by the transition probability Eq. (12). The full lines correspond to the analytical expression Eq. (36), while the circles correspond to numerical simulations with $n = 300$ terms and 5×10^5 realizations (see Fig. 1). In (a) $\lambda = 20$, (b) $\lambda = 4$, (c) $\lambda = 2$, and (d) $\lambda = 1$. In all cases, $q_+ = q_- = 1/2$.

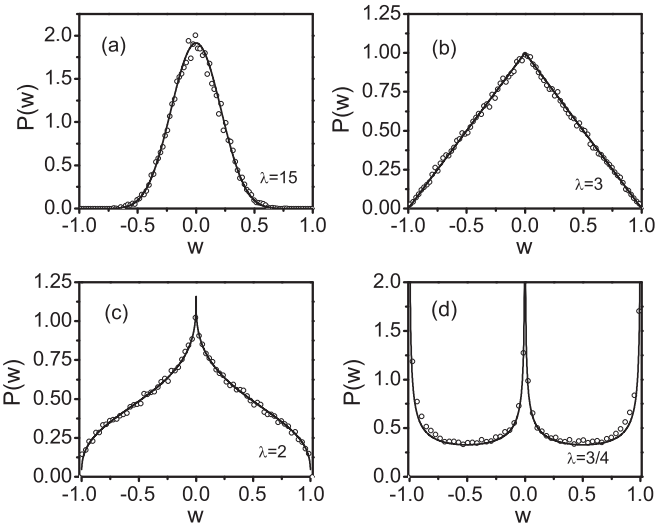

FIG. 5. Probability density $P(w)$ for random variables with three discrete states, $x_\mu = +1, 0, -1$, obtained from the transition probability Eq. (12). The full lines correspond to the analytical expression Eq. (37), while the circles correspond to a numerical simulation with $n = 300$ terms and 5×10^5 realizations (see Fig. 2). In (a) $\lambda = 15$, (b) $\lambda = 3$, (c) $\lambda = 2$, and (d) $\lambda = 3/4$. In all cases, $q_+ = q_0 = q_- = 1/3$.

For three-states variables with $\{x_\mu\} = \{+1,0,-1\}$, the parameters are $\{q_\mu\} = \{q_+,q_0,q_-\}$ and λ. They can be parametrized as $\{\lambda_\mu\} = \{\lambda q_\mu\} = \{\lambda_+,\lambda_0,\lambda_-\}$. By taking into account that $\bar{x}_f = f_+ - f_-$ [Eq. (35)], from Eq. (34) we get

$$P(w) = \begin{cases} g_+[w] & w > 0 \\ g_-[w] & w < 0 \end{cases}, \tag{37}$$

where each contribution is defined as

$$g_+[w] = \int_0^{\frac{1-w}{2}} df\ f^{\lambda_- -1}(1-w-2f)^{\lambda_0-1}(f+w)^{\lambda_+-1},$$

$$g_-[w] = \int_0^{\frac{1+w}{2}} df\ (f-w)^{\lambda_- -1}(1+w-2f)^{\lambda_0-1} f^{\lambda_+-1}.$$

These integrals can be solved in terms of the hypergeometric function ${}_2F_1[a,b,c,z]$ as

$$g_+[w] = \frac{(1-w)^{\lambda_- +\lambda_0-1} w^{\lambda_+-1}}{2^{\lambda_-}\Gamma^{-1}(\lambda)\Gamma(\lambda_+)\Gamma(\lambda_0+\lambda_-)} {}_2F_1\left[\lambda_-, 1-\lambda_+, \lambda_0+\lambda_-, \frac{w-1}{2w}\right],$$

and similarly,

$$g_-[w] = \frac{(1+w)^{\lambda_+ +\lambda_0-1}(-w)^{\lambda_- -1}}{2^{\lambda_+}\Gamma^{-1}(\lambda)\Gamma(\lambda_-)\Gamma(\lambda_0+\lambda_+)} {}_2F_1\left[1-\lambda_-, \lambda_+, \lambda_0+\lambda_+, \frac{w+1}{2w}\right].$$

The hypergeometric function is defined by ${}_2F_1[a,b,c,z] = \sum_{k=0}^{\infty}(a)_k(b)_k(c)_k z^k/k!$ with $(x)_k = \prod_{j=0}^{k-1}(x+j)$. Simpler expressions can be found in the particular case $\lambda_+ = \lambda_{-1} = 1$, where Eq. (37) reduces to

$$P(w) = \frac{1}{2}(1+\lambda_0)(1-|w|)^{\lambda_0}, \quad \lambda_+ = \lambda_{-1} = 1. \tag{38}$$

In Fig. 5 we show a set of plots corresponding to $P(w)$. The realizations, over which the probability densities are obtained, are those shown in Fig. 2. We found that the density Eq. (37) fits the numerical results. The case shown in Fig. 5(b) corresponds to Eq. (38) with $\lambda_0 = 1$.

C. Composed Pólya urn scheme

The previous results with the Pólya urn scheme (see also the Appendices) demonstrate that the sum variable Eq. (25), depending on the underlying correlation mechanism, may adopt very different statistics. In contrast to independent random variables, these probabilities do not have associated a basin of attraction. Here, we show that sum of sequences of random variables obtained from the composed Pólya urn scheme lead to the same probability densities (attractors) as in the standard scheme. From this result, it is possible to conclude that attractors (probabilities densities) associated to the Pólya urn scheme have a basin of attraction that is as wide as in the standard CLT. This is the main result of this section.

For the composed Pólya urn scheme, the probability density of the sum variable W [Eq. (25)] reads

$$P(w) = \int_\Lambda df_1 \dots df_{M-1}\ \delta(w - \bar{x}_f) D(\{f_\mu\}|\{\lambda_\mu\}). \tag{39}$$

Here, $\bar{x}_f$ is given by

$$\bar{x}_f = \sum_{\mu=1}^{M} f_\mu \bar{x}_\mu, \qquad \bar{x}_\mu = \int_{\Omega_\mu} dx x p_\mu(x), \tag{40}$$

while $D(\{f_\mu\}|\{\lambda_\mu\})$, Eq. (33), is defined with the parameters [see Eq. (21)]

$$\lambda_\mu = \lambda q_\mu = \lambda \int_{\Omega_\mu} P_1(x)dx. \tag{41}$$

This result is demonstrated below. Notice that Eq. (39) is similar to that of the standard scheme, Eq. (34). In fact, they can be related by the replacements $x_\mu \leftrightarrow \bar{x}_\mu$ [Eqs. (35) and (40)], while here the weights $\{q_\mu\}_{\mu=1}^M$ are defined by Eq. (21).

The composed urn scheme is completely defined after given $P_1(x)$ and the subdomains $\{\Omega_\mu\}_{\mu=1}^M$. Taking into account the previous three equations, we realize that *different* $P_1(x)$ that lead to the same weights $\{q_\mu\}_{\mu=1}^M$ [Eq. (21)] and mean values $\{\bar{x}_\mu\}_{\mu=1}^M$ [Eq. (40)] lead to the same attractor $P(w)$ [Eq. (26)] for the normalized sum W, Eq. (25). The basin of attraction of the probabilities density Eq. (39) correspond to all joint probabilities densities $\{P_n(\{x_i\})\}_{n=1}^\infty$ associated to the composed scheme. From the previous comments, it follows that these conditional objects can be parametrized by the density $P_1(x)$. Therefore, we conclude that the basin of attraction of the Pólya urn attractors [Eq. (39)] is as wide as in the standard CLT (independent random variables).

For demonstrating the validity of Eq. (39) we use that the composed Pólya urn scheme consist of two independent random processes: the randomness introduced by the probability densities $p_\mu(x)$ associated to each subdomain Ω_μ and the underlying Pólya urn process that select each subdomain. Therefore, the joint probability density of the random variables $\{X_i\}_{i=1}^n$ reads

$$P_n(\{x_i\}) = \langle p_{\mu_1}(x_1)\dots p_{\mu_n}(x_n)\rangle_{\{\mu\}}. \tag{42}$$

Here, each index $\mu_i = 1\dots M$ runs over the set of subdomains $\{\Omega_\mu\}_{\mu=1}^M$. Furthermore, $\langle\dots\rangle_{\{\mu\}}$ denotes and average over the ensemble of realizations associated with the underlying Pólya urn scheme. From Eqs. (28) and (29) we get $G_w(k) = \lim_{n\to\infty} G_w^{(n)}(k)$, with

$$G_w^{(n)}(k) = \left\langle G_{\mu_1}\left(\frac{k}{n}\right)\dots G_{\mu_n}\left(\frac{k}{n}\right)\right\rangle_{\{\mu\}}, \tag{43}$$

where $G_\mu(k)$ is the Fourier transform of $p_\mu(x)$. By indexing the realizations by the number of times n_μ that each subdomain Ω_μ is selected, we can write

$$G_w^{(n)}(k) = \left\langle \left[G_1\left(\frac{k}{n}\right)\right]^{n_1}\dots\left[G_M\left(\frac{k}{n}\right)\right]^{n_M}\right\rangle_{\{n\}}. \tag{44}$$

When $n\to\infty$ the set of occurrences also diverge, $\{n_\mu\}\to\infty$. Thus, maintaining the leading order in n, in that limit each factor in the previous expression can be written as

$$\lim_{n\to\infty}\left[G_\mu\left(\frac{k}{n}\right)\right]^{n_\mu} = \lim_{n\to\infty}\exp\left(ik\bar{x}_\mu\frac{n_\mu}{n}\right)\exp\left(-\frac{\sigma_\mu^2k^2}{2n}\frac{n_\mu}{n}\right), \tag{45}$$

where $\bar{x}_\mu$ is the mean value defined in Eq. (40) while $\sigma_\mu^2 = \int_{\Omega_\mu} dx(x-\bar{x}_\mu)^2 p_\mu(x)$. Notice that Eq. (45) relies on the validity of the law of large numbers for each density $p_\mu(x)$.

In the previous expression, the argument n_u/n, in the asymptotic limit, can be associated with the random variables F_μ, Eq. (14). Therefore, $n_u/n \to F_\mu$, which from Eqs. (44) and (45) leads to

$$G_w(k) = \left\langle \exp ik\sum_\mu F_\mu\bar{x}_\mu\right\rangle_{\{F\}}. \tag{46}$$

The average over the random set of weights $\{F\}$ is governed by the Dirichlet distribution Eq. (33). Therefore, after Fourier inversion we recover Eq. (39) with Eq. (40). This finishes the demonstration.

As an example of the previous result we take a composed Pólya urn scheme [Eq. (22)] defined with two subdomains $\Omega_\pm$ with densities $p_\pm(x)$. We get [Eq. (44)]

$$G_w^{(n)}(k) = \left\langle\left[G_+\left(\frac{k}{n}\right)\right]^{n_+}\left[G_-\left(\frac{k}{n}\right)\right]^{n_-}\right\rangle_{\{n\}}, \tag{47}$$

where $n_\pm$ are the number of times that each subdomain $\Omega_\pm$ was chosen, and $G_\pm(k) = \int_{-\infty}^{+\infty} dwe^{ikw}p_\pm(w)$. Using that $n_+ + n_- = n$, it follows

$$G_w^{(n)}(k) = \sum_{n_+=0}^{n} P_n(n_+)\left[G_+\left(\frac{k}{n}\right)\right]^{n_+}\left[G_-\left(\frac{k}{n}\right)\right]^{n-n_+}, \tag{48}$$

where $P_n(n_+)$ is the probability of the random variable n_+. This object, after some algebra and by using the properties of Gamma functions, can be obtained from Eqs. (4) and (12). Alternatively, it can be obtained directly from de Finetti representation theorem [see. Eq. (C12)]. It reads

$$P_n(n_+) = \frac{1}{\mathcal{N}_n}\binom{n}{n_+}\frac{\Gamma(n_+ + \lambda_+)}{\Gamma(\lambda_+)}\frac{\Gamma(n-n_+ + \lambda_-)}{\Gamma(\lambda_-)}, \tag{49}$$

where $\mathcal{N}_n = \Gamma(n+\lambda)/\Gamma(\lambda)$, and $\lambda_\pm = \lambda q_\pm$ [Eq. (21)].

The previous two expressions give an exact analytical expression for $G_w^{(n)}(k)$. For the example defined by Eq. (23), the random variables have a uniform probability density for $X\in[-2,2]$. The probabilities of each subdomain are defined by Eq. (24). Their Fourier transform reads

$$G_\pm(k) = [\sin(k)/k]e^{\pm ik}. \tag{50}$$

In order to check these results, in Fig. 6 we show a set of probability distributions obtained by averaging the realizations of the composed scheme (Fig. 3). For each $n = 1, 2, 10, 300$, the numerical results follows after averaging 5×10^5 realizations. For $n = 1$ it is recovered Eq. (23). For higher n we find that the (numerical) inverse Fourier transform of Eq. (48) evaluated with Eq. (50) fits very well the numerical results (circles). Consistently with the previous analysis, at $n = 300$ the density $P(w_n)$ is almost indistinguishable from the corresponding attractor, that is, $P(w_n)$ in Fig. 6(d) is very well fitted by the density $P(w)$ corresponding to the standard scheme, Eq. (36), which in turn is plotted in Fig. 4(b). This follows because the average values $\{\bar{x}_\mu\}$ [Eq. (40)] are $\bar{x}_\pm = \pm1$ and also the weights $\{q_\mu\}$ [Eq. (21)] are $q_\pm = 1/2$, which correspond to the parameters of Fig. 4. We also checked that for all values of λ the attractors correspond to those shown in that figure.

For arbitrary distributions $P_1(x)$ the sum variable, associated to the composed urn scheme with two subdomains, is characterized by the attractor Eq. (36). In general, the random variables can assume values over the entire real line. For

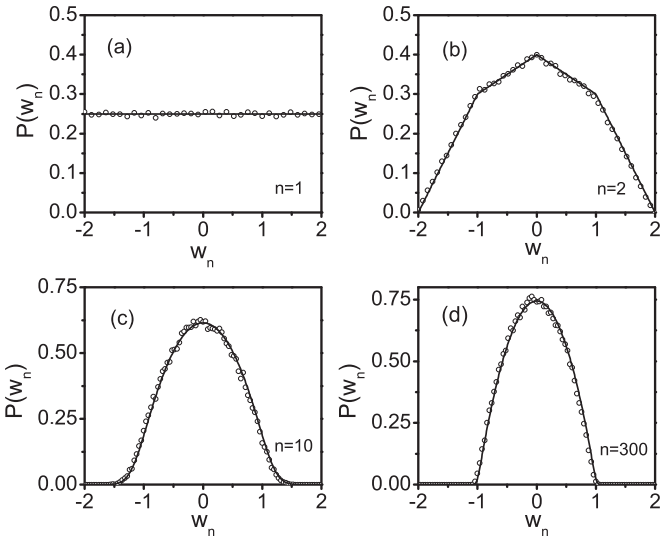

FIG. 6. Probability density $P(w_n)$ of the variable $W_n = (1/n)\sum_{i=1}^{n} X_i$, where each random variable X_i follows from the composed Pólya urn scheme defined by Eqs. (22) and (23), with $\lambda = 4$. The weights [Eq. (21)] are $q_+ = q_- = 1/2$. The solid line follows from the inverse Fourier transform of Eq. (48) defined with Eq. (50). The circles correspond to numerical results obtained by averaging 5×10^5 realizations. In (a) $n = 1$, (b) $n = 2$, (c) $n = 10$, and (d) $n = 300$.

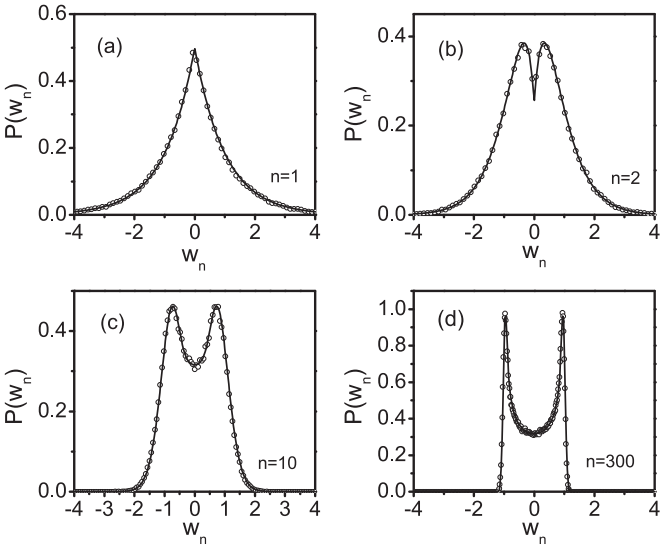

FIG. 7. Probability density $P(w_n)$ of the variable $W_n = (1/n)\sum_{i=1}^{n} X_i$, where each random variable X_i follows from the composed Pólya urn scheme defined by Eqs. (22) and (51), with $\lambda = 1$. The weights [Eq. (21)] are $q_+ = q_- = 1/2$. The solid line follows from the inverse Fourier transform of Eq. (48) defined with Eq. (52). The circles correspond to numerical results obtained by averaging 5×10^5 realizations. In (a) $n = 1$, (b) $n = 2$, (c) $n = 10$, and (d) $n = 300$.

example, we take

$$P_1(x) = (1/2)\exp(-|x|), \tag{51}$$

with subdomains $\Omega_\pm = x \gtrless 0$. Then, the Fourier transforms of $p_\pm(x) = \exp(\mp x)$ read

$$G_\pm(k) = \frac{1}{1 \mp ik}. \tag{52}$$

In Fig. 7 we show a set of probability distributions for the sum variable for this alternative single statistics. As in the previous case, the analytical expressions in the Fourier domain fit the numerical results. Notice that even when the single variables assume values over the real line, their normalized sum is characterized by an (probability density) attractor that is not null only in the interval $(-1,1)$ [see Fig. 4(d)]. This property is induced by the global correlation effects.

For an urn model with three states, similar results can be obtained. For example, by maintaining $P_1(x)$ given by Eq. (23), taking the subdomains $\Omega_+ = (1/3,5/3)$, $\Omega_- = (-5/3, -1/3)$, and $\Omega_0 = (-2, -5/3) \cup (5/3,2)$, leads to the attractors shown in Fig. 5. A model with exponential distributed variables can also be written.

V. SUMMARY AND CONCLUSIONS

Beyond statistically independent variables, there exist very few generalizations of the CLT. Here, we studied this problem for globally correlated random variables that are similar and interchangeable. In order to characterize these symmetries we derived a hierarchical set of equations that the transition probability densities must to satisfy, Eq. (7). These integral equations provide a tool for constructing correlation mechanisms that satisfy the required properties.

Different correlation mechanisms lead to transitions probability densities that fulfill the demanded symmetries, such as globally correlated Gaussian variables, de Finetti representation (see Appendices), and urn schemes. In this last context, we introduced a generalization of Pólya urn scheme, where the values assumed by the random variables are split in different subdomains over the real line, each one being endowed with a probability density. Each subdomain is chosen in agreement with the standard Pólya urn scheme, while the associated probability density delivers the next random value (Fig. 3). The transition probability of this composed scheme, Eq. (22), fulfills the required symmetries.

The sum variable that defines the CLT, Eq. (25), here is defined with a different normalization because the studied random variables may achieve a fully correlated regime. Thus, the case of independent variables leads to a delta Dirac distribution, a fact related to the validity of the law of large numbers. In general, global correlations consistent with the demanded symmetries lead to very different statistics of the sum variable. The Pólya urn scheme, depending on its number of states and characteristic parameters, delivers different probability densities, Eq. (34) (see Figs. 4 and 5). For the particular case of two states, the attractor is defined by an asymmetric q-Gaussian density ($q < 1$), Eq. (36). More complex expressions arise for more states.

Given the diversity of possible attractors, a very difficult task is to characterize their basin of attraction. We solved this problem in a constructive way. We demonstrated that the sum of random variables generated via the composed Pólya urn scheme leads to the same attractors as in the standard scheme (see Figs. 6 and 7). Therefore, their basin of attraction, in the space of joint probability densities, can be parametrized by the probability density of each random variable. In fact,

there exist infinite single probability distributions that with a specific splitting of their domain lead to the same attractor [see Eqs. (20) and (21)]. In contrast with the standard CLT, here the mechanism that guarantees this result is the validity of the law of large numbers for the probability density of each subdomain as well as the convergence to stationary values of the transition probability of the standard Pólya urn scheme.

The basin of attraction of the Pólya urn attractors can be extended after raising up the interchangeability symmetry [Eq. (18)] in the composed scheme. In addition, the same attractors arise, for example, by introducing correlations between the random variables in such a way that the law of large numbers remains valid in each subdomain. The present results also lead us to ask about physical systems characterized by dynamical correlations able to induce attractors that take values on a subdomain of the underlying random process (variables).

In conclusion, we developed a consistent approach for dealing with globally correlated similar interchangeable random variables, which in turn allowed us to characterize different attractors of the CLT as well as their basin of attraction.

ACKNOWLEDGMENT

This work was supported by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

APPENDIX A: INTERCHANGEABILITY CONDITION FOR THE CONDITIONAL PROBABILITIES

Here, we derive the hierarchical set of conditions defined by Eq. (7). Assuming that interchangeability is valid for $P_n(x_1,\dots x_n)$, we determine the conditions under which $P_{n+1}(x_1,\dots x_{n+1})$ also fulfill the symmetry. These functions are related as $P_{n+1}(x_1,\dots x_{n+1}) = P_n(x_1,\dots x_n)\, T_n(x_1,\dots x_n|x_{n+1})$. Therefore, $T_n(x_1,\dots x_n|x_{n+1})$ must also be symmetric in the $x_1,\dots x_n$ arguments. The interchangeability for $P_{n+1}(x_1,\dots x_{n+1})$ is valid when x_{n+1} can be interchanged with an arbitrary x_k, with $k = 1,\dots n$. Written in an explicit way, this requirement reads

$$P_{n+1}(x_1,\dots,x_k,\dots x_{n+1}) = P_{n+1}(x_1,\dots,x_{n+1},\dots x_k). \tag{A1}$$

By using Bayes rule, these objects can be written as

$$\begin{aligned} P_{n+1}(x_1,\dots,x_k,\dots x_{n+1}) &= P_{k-1}(x_1\dots x_{k-1}) \\ &\times T_{k-1}(x_1\dots x_{k-1}|x_k) \\ &\times T_k(x_1,\dots x_k|x_{k+1}) \\ &\times T_{k+1}(x_1,\dots,x_k,x_{k+1}|x_{k+2}) \\ &\cdots \times T_{n-1}(x_1,\dots x_{n-1}|x_n) \\ &T_n(x_1,\dots x_n|x_{n+1}), \end{aligned}$$

and also

$$\begin{aligned} P_{n+1}(x_1,\dots,x_{n+1},\dots x_k) &= P_{k-1}(x_1\dots x_{k-1}) \\ &\times T_{k-1}(x_1\dots x_{k-1}|x_{n+1}) \\ &\times T_k(x_1,\dots,x_{k-1},x_{n+1}|x_{k+1}) \\ &\times T_{k+1}(x_1,\dots,x_{k-1},x_{n+1},x_{k+1}|x_{k+2}) \\ &\cdots \times T_{n-1}(x_1,\dots,x_{k-1},x_{n+1},x_{k+1}\dots x_{n-1}|x_n) \\ &\times T_n(x_1,\dots,x_{k-1},x_{n+1},x_{k+1}\dots x_n|x_k), \end{aligned}$$

where now $k = 2,\dots n$. Performing the integrals $\int dx_k dx_{k+1}\dots dx_n$ to both objects, using the normalization condition $\int dx_j T_i(x_1,\dots x_i|x_j) = 1$, and simplifying the factor $P_{k-1}(x_1\dots x_{k-1})$, from Eq. (A1) it follows the condition

$$\begin{aligned} &T_{k-1}(x_1\dots x_{k-1}|x_{n+1}) \\ &\quad = \int dx_k\dots dx_n T_{k-1}(x_1\dots x_{k-1}|x_k) \\ &\qquad \times T_k(x_1,\dots x_k|x_{k+1}) T_{k+1}(x_1,\dots,x_k,x_{k+1}|x_{k+2}) \\ &\qquad \cdots \times T_{n-1}(x_1,\dots x_{n-1}|x_n) T_n(x_1,\dots x_n|x_{n+1}). \end{aligned} \tag{A2}$$

For $k = n$, this equation reduces to

$$\begin{aligned} &T_{n-1}(x_1\dots x_{n-1}|x_{n+1}) \\ &\quad = \int dx_n T_{n-1}(x_1\dots x_{n-1}|x_n) T_n(x_1,\dots x_n|x_{n+1}). \end{aligned} \tag{A3}$$

For $k = n-1$, after using the validity of Eq. (A3), Eq. (A2) leads to

$$\begin{aligned} &T_{n-2}(x_1\dots x_{n-2}|x_{n+1}) \\ &\quad = \int dx_{n-1} T_{n-2}(x_1\dots x_{n-2}|x_{n-1}) T_{n-1}(x_1,\dots x_{n-1}|x_{n+1}). \end{aligned}$$

Notice that this equation has the same structure as Eq. (A3). Hence, it is simple to realize that Eq. (A2) is satisfied if

$$\begin{aligned} &T_{n-j}(x_1\dots x_{n-j}|x_{n+1}) \\ &\quad = \int dx_{n-j+1} T_{n-j}(x_1\dots x_{n-j}|x_{n-j+1}) \\ &\qquad \times T_{n-j+1}(x_1,\dots x_{n-j+1}|x_{n+1}). \end{aligned}$$

where $j = 1,\dots n-(k-1)$. This last equation, after a straightforward change of indexes, recovers Eq. (7).

APPENDIX B: ADDITIVE MEMORY CASE

The symmetry of the transition probability $T_n(x_1\dots x_n|x_{n+1})$ on the previous conditional values $x_1\dots x_n$ is trivially fulfilled by assuming that it depends on the addition of these values. Then, we write

$$T_n(x_1\dots x_n|x_{n+1}) = \mathcal{T}_n(x_1+x_2\cdots+x_n|x_{n+1}), \tag{B1}$$

where $\mathcal{T}_n(x'|x)$ is a set of equivalent functions that only depends on two arguments. For random variables $\{X_i\}_{i=1}^n$ with a finite domain, $X \in [x_<,x_>]$, the variable x' in $\mathcal{T}_n(x'|x)$ runs in the interval $[nx_<,nx_>]$.

From Eq. (7), it follows that the functions $\mathcal{T}_n(x'|x)$ must to satisfy the recursive relations

$$\mathcal{T}_{n-1}(x'|x) = \int dy \mathcal{T}_{n-1}(x'|y)\mathcal{T}_n(x'+y|x). \tag{B2}$$

Below we show that the additive assumption allows us to find a complete solution of the hierarchy Eq. (7) after assuming different statistics for each single variable.

1. Gaussian random variables

For the single statistics of each random variable, let us assume a Gaussian distribution

$$P_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right], \tag{B3}$$

which satisfies $\int dx\, P_1(x) = 1$. The width σ^2 is a free parameter. Given that $\mathcal{T}_1(x'|x) = T_1(x'|x)$, in order to fulfill Eq. (5) we assume that $\mathcal{T}_1(x'|x)$ is a Gaussian distribution in both variables x' and x. Hence, $\mathcal{T}_1(x'|x) \approx \exp[-(\frac{x^2}{2\rho^2} + \frac{x'^2}{2\mu^2} + \frac{xx'}{\nu})]$. The undetermined free parameters (ρ, μ, ν) are constrained by the normalization condition $\int dx \mathcal{T}_1(x'|x) = 1$ and Eq. (5). After imposing these constraints, we obtain

$$\mathcal{T}_1(x'|x) = \frac{1}{\sqrt{2\pi\rho^2}} \exp\left[-\frac{1}{2\rho^2}(x - \lambda x')^2\right], \tag{B4}$$

where the real parameter λ is

$$\lambda \equiv \sqrt{1 - \frac{\rho^2}{\sigma^2}}. \tag{B5}$$

ρ remains as a free parameter and satisfies $\rho^2 \leqslant \sigma^2$. Notice that when $\lambda = 0$, that is $\rho = \sigma$, we get independent variables, $\mathcal{T}_1(x'|x) = P_1(x)$. On the other hand, for $\lambda = 1, \rho \to 0$, it follows $\mathcal{T}_1(x'|x) = \delta(x - x')$. This is the maximal correlated case, where $x = x'$. Hence, after the first random value, the next one is equal to the previous one.

Higher transition probabilities can be obtained from Eq. (B2) and the solution Eq. (B4). Proposing a Gaussian structure for higher objects, we get

$$\mathcal{T}_n(x'|x) = \frac{1}{\sqrt{2\pi\rho_n^2}} \exp\left[-\frac{1}{2\rho_n^2}(x - \lambda_n x')^2\right], \tag{B6}$$

where the coefficients satisfy the recursive relations

$$\lambda_n = \frac{\lambda_{n-1}}{1 + \lambda_{n-1}}, \quad \rho_n^2 = \left[1 - \left(\frac{\lambda_{n-1}}{1 + \lambda_{n-1}}\right)^2\right]\rho_{n-1}^2 \tag{B7}$$

$(n \geqslant 2)$, with $\lambda_1 \equiv \lambda$ and $\rho_1 \equiv \rho$. Their solution is

$$\lambda_n = \frac{\lambda}{1 + (n-1)\lambda}, \quad \rho_n^2 = \left[1 + \frac{(n-1)\lambda^2}{1 + n\lambda}\right]^{-1} \rho^2. \tag{B8}$$

The joint probability density $P_n(x_1, \ldots x_n)$ can be obtained from the set of transition probabilities [Eq. (4)]. For example, the joint probability $P_2(x_1, x_2)$, from Eqs. (B3) and (B4), reads

$$P_2(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma^2\rho^2}} \exp\left[-\frac{1}{2\rho^2}\left(x_1^2 + x_2^2 - 2\lambda x_1 x_2\right)\right], \tag{B9}$$

which consistently is symmetric in x_1 and x_2. For arbitrary $n \geqslant 2$, we get

$$P_n(x_1, \ldots x_n) = \sqrt{\frac{\det[A^{(n)}]}{(2\pi)^n}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{n} x_i \mathrm{A}_{ij}^{(n)} x_j\right], \tag{B10}$$

where the matrix elements are

$$A_{ii}^{(n)} = \frac{1}{\rho_{n-1}^2}, \quad A_{ij}^{(n)} = -\frac{\lambda_{n-1}}{\rho_{n-1}^2}, \quad i \neq j. \tag{B11}$$

Here, ρ_n and λ_n are defined by Eq. (B8). The determinant of the matrix $A_{ij}^{(n)}$ reads

$$\det[A^{(n)}] = \left\{[1 + (n-1)\lambda]\sigma^2\left(\frac{\rho^2}{1+\lambda}\right)^{n-1}\right\}^{-1}. \tag{B12}$$

The validity of Eq. (B10) can be probed by using the mathematical principle of induction and the recursive relations Eq. (B7).

We remark that Eq. (B10) was derived over the basis of the conditional probabilities densities Eq. (B6), which in turn are a solution of the hierarchy Eq. (B2) after assuming the Gaussian statistics defined by Eq. (B3). Clearly, due to the symmetry of the covariance matrix Eq. (B11), the multidimensional Gaussian density Eq. (B10) is compatible with the interchangeability symmetry.

Now we obtain the probability density of W [Eq. (25)] for a set of random variables $\{X_i\}_{i=1}^n$ correlated in agreement with the Gaussian distribution Eq. (B10), which in turn is related to the transition probability Eq. (B6). The (multiple) Fourier transform of Eq. (B10) reads

$$G_k(k_1, \ldots k_k) = \exp\left[-\frac{1}{2}\sum_{i,j=1}^{k} k_i (1/A^{(k)})_{ij} k_j\right], \tag{B13}$$

where $(1/A^{(k)})$ is the matrix inverse of $A^{(k)}$ [Eq. (B11)]. It can be written as

$$(1/A^{(k)})_{ii} = \sigma^2, \quad (1/A^{(k)})_{ij} = \sigma^2\lambda, \quad i \neq j, \tag{B14}$$

where $\lambda = (1 - \rho^2/\sigma^2)^{1/2}$ [Eq. (B5)]. Hence, from Eqs. (28) and (B13), we get

$$G_{w_n}(k) = \exp\left\{-\frac{1}{2}\sigma^2\lambda k^2\left[1 + \frac{1}{n}(\lambda^{-1} - 1)\right]\right\}. \tag{B15}$$

After taking the limit $n \to \infty$, it follows

$$P(w) = \sqrt{\frac{1}{2\pi\sigma^2\lambda}} \exp\left[-\frac{1}{2}\frac{w^2}{\sigma^2\lambda}\right]. \tag{B16}$$

Contrarily to the case of independent variables, here the probability density of W is not a delta Dirac distribution, Eq. (31). This departure has its origin in the correlations between the random variables, which are tuned by the parameter λ. In fact, in the limit $\lambda \to 0$ we recover Eq. (31) with $\bar{x} = 0$, that is, independent variables. On the other hand, for maximally correlated variables, $\lambda \to 1$, we recover the Gaussian distribution $P_1(x)$ [Eq. (B3)]. This result, which gives the maximal departure with respect to independent variables, follows after noting that all random variables assume the same value [see the transition probabilities Eqs. (B4) and (B6)].

2. Linear additive memory case

Here, we search another class of solution that in addition assumes that the transition probabilities $\mathcal{T}_n(x'|x)$ depend linearly on the argument x'. In the following results, the structure of $P_1(x)$ is arbitrary.

Given $P_1(x)$, and given the linear dependence of $\mathcal{T}_1(x'|x)$ on x', the relation defined by Eq. (5) becomes

$$\mathcal{T}_1(\langle X\rangle|x) = P_1(x), \quad \langle X\rangle \equiv \int dx\, P_1(x) x. \tag{B17}$$

Given $P_1(x)$, any transition probability density $T_1(x'|x)$ satisfying this equation is a valid one. Assuming that all transition probability densities depend linearly on x', the conditions

Eq. (B2) can be written as

$$\mathcal{T}_n(x' + \langle X\rangle_{n-1,x'}|x) = \mathcal{T}_{n-1}(x'|x), \tag{B18}$$

where the conditional average $\langle X\rangle_{n-1,x'}$ is defined as

$$\langle X\rangle_{n-1,x'} \equiv \int dx \mathcal{T}_{n-1}(x'|x)x. \tag{B19}$$

By evaluating the previous two expressions in $x' = \langle X\rangle$, the relation

$$\mathcal{T}_n(n\langle X\rangle|x) = P_1(x) \tag{B20}$$

follows, which generalizes that defined by Eq. (B17).

From Eq. (B19), we realize that $\langle X\rangle_{n-1,x'}$ is also a linear function of x'. In particular, it is possible to write

$$\langle X\rangle_{1,x'} = \int dx \mathcal{T}_1(x'|x)x = ax' + b. \tag{B21}$$

This equation defines the constants a and b, the former being a dimensionless one, while the last one has units of x. Multiplying the previous expression by $P_1(x')$ and integrating in x', the relation $\langle X\rangle = b/(1-a)$ follows.

From Eq. (B21), the solution of Eq. (B18) for $n = 2$ is $\mathcal{T}_2(x'(1+a)+b|x) = \mathcal{T}_1(x'|x)$, which can be rewritten as

$$\mathcal{T}_2(x'|x) = \mathcal{T}_1\left(\frac{x'-b}{1+a}\middle|x\right). \tag{B22}$$

In a similar form, an explicit expression for $\mathcal{T}_3(x'|x)$ can be obtained. For arbitrary n, as a solution of Eq. (B18) we propose the expression

$$\mathcal{T}_n(x'|x) = \mathcal{T}_1\left(\frac{x'-(n-1)b}{1+(n-1)a}\middle|x\right). \tag{B23}$$

The validity of this result can be proven from Eq. (B18) by using the mathematical principle of induction.

Discrete distributions with finite domain

The set of functions defined by Eq. (B23) give a full solution to the hierarchical structure Eq. (B2). Nevertheless, it is not guaranteed that they are are positive functions. In order to check this issue, we consider discrete random variables defined by

$$P_1(x) = \sum_{\mu=1}^{M} q_\mu \delta(x - x_\mu), \tag{B24}$$

where the positive weights satisfy $\sum_{\mu=1}^{M} q_\mu = 1$.

The mean value, $\langle X\rangle = \int dx\, P_1(x)x$, reads $\langle X\rangle = \sum_{\mu=1}^{M} q_\mu x_\mu$. The first conditional density, given its linear dependence on x', is written as

$$\mathcal{T}_1(x'|x) = \frac{1}{\mathcal{N}}\sum_{\mu=1}^{M}(\alpha_\mu + \beta_\mu x')\delta(x - x_\mu), \tag{B25}$$

where (α_μ,β_μ) and $\mathcal{N}$ are arbitrary parameters. Using the normalization condition $\int dx \mathcal{T}_1(x'|x) = 1$, it follows $\mathcal{N} = \sum_{\mu=1}^{M}\alpha_\mu$, and

$$\sum_{\mu=1}^{M}\beta_\mu = 0. \tag{B26}$$

The condition $\mathcal{T}_1(\langle X\rangle|x) = P_1(x)$, leads to $\frac{1}{\mathcal{N}}(\alpha_\mu + \beta_\mu\langle X\rangle) = q_\mu$. Under the association $(\beta_\mu/\mathcal{N}) \to \beta_\mu$, we get

$$\mathcal{T}_1(x'|x) = \sum_{\mu=1}^{M}[q_\mu + \beta_\mu(x' - \langle X\rangle)]\delta(x - x_\mu). \tag{B27}$$

The first conditional average reads

$$\int dx \mathcal{T}_1(x'|x)x = \zeta x' + \langle X\rangle(1-\zeta) = ax' + b, \tag{B28}$$

where the constant ζ is

$$\zeta \equiv \sum_{\mu=1}^{M} x_\mu\beta_\mu. \tag{B29}$$

From Eq. (B23), higher objects read

$$\mathcal{T}_n(x'|x) = \sum_{\mu=1}^{M}\left[q_\mu + \beta_\mu\frac{x' - n\langle X\rangle}{1+(n-1)\zeta}\right]\delta(x - x_\mu). \tag{B30}$$

We remark that this set of equations provides a solution to the full hierarchy of conditional probabilities under the interchangeability symmetry. Nevertheless, the positivity of these objects must be checked.

The constants β_μ should be chosen such that the positivity of $\mathcal{T}_n(x'|x)$ is guaranteed for all n and $x' \in (nx_<, nx_>)$, where $x_<$ and $x_>$ define, respectively, the minimal and maximal values of the set $\{x_\mu\}_{\mu=1}^{M}$. Hence, for $n = 1$, it follows that

$$q_\mu + \beta_\mu(x - \langle X\rangle) \geqslant 0, \tag{B31}$$

while in the limit $n \to \infty$, we get

$$q_\mu + \frac{\beta_\mu(x - \langle X\rangle)}{\sum_{\nu=1}^{M} x_\nu\beta_\nu} \geqslant 0. \tag{B32}$$

In both inequalities, x assumes values over the set $\{x_u\}$. In the case of two states, $M = 2$, from these inequalities we obtain $\beta \leqslant |x_2 - x_1|^{-1}$, where $\beta_1 = -\beta_2 = \beta$, and $\{x_\mu\} = \{x_1, x_2\}$. Hence, positivity can always be guaranteed in this case.

In general, for $M \geqslant 3$, there is not a solution for the set $\{\beta_\mu\}$ that guarantees the validity of the previous two inequalities. In fact, from Eq. (B32), we deduce that

$$|\beta_\mu(x - \langle X\rangle)| \leqslant q_\mu\left|\sum_{\nu=1}^{M} x_\nu\beta_\nu\right|. \tag{B33}$$

Taking $x \to x_\mu$, and adding in the μ index, $\sum_{\mu=1}^{M}$, it follows that

$$\sum_{\nu=1}^{M}|\beta_\mu x_\mu| \leqslant \left|\sum_{\nu=1}^{M} x_\mu\beta_\mu\right|. \tag{B34}$$

Hence, we deduce that $x_\mu\beta_\mu > 0$, and then $\sum_{\mu=1}^{M} x_\mu\beta_\mu > 0$. Therefore, Eqs. (B31) and (B32) are equivalents, in the sense that one of them always implies the other. Taking one of them and the previous one, it follows $M(M-1) - 1$ equations, while the number of variables is $M - 1$. Thus, a consistent solution (positive transition probabilities) is *only* available when $M = 2$.

For classical spin variables $x_\mu = \pm 1$, parametrizing $\beta = (1/2)(1+\lambda)^{-1} \leqslant |x_+ - x_-|^{-1} = 1/2$, from Eq. (B30) we get

$(\lambda_\pm = \lambda q_\pm)$

$$\mathcal{T}_n(x'|x) = \left(\frac{\lambda_+ + \frac{n+x'}{2}}{n+\lambda}\right)\delta(x-1) + \left(\frac{\lambda_- + \frac{n-x'}{2}}{n+\lambda}\right)\delta(x+1). \tag{B35}$$

This expression gives a positive solution consistent with the required symmetries. Nevertheless, it is simple to realize that the quantities $\frac{n+x'}{2}$ and $\frac{n-x'}{2}$ give the number of times n_+ and n_- that the previous variables assumed the values ± 1, respectively. Therefore, Eq. (B35) recovers the transition probability corresponding to the Pólya urn scheme, Eq. (12).

APPENDIX C: de FINETTI REPRESENTATION

de Finetti [20] introduced the concept of interchangeability and also defined a general representation structure for the joint probability density of a set of dichotomic interchangeable variables. The de Finetti representation can be generalized for arbitrary (nondichotomic) random variables. Given a set of interchangeable random variables $\{X_i\}_{i=1}^n$, their n-joint probability density is expressed as

$$P_n(x_1,\ldots,x_n) = \int_{\Omega_y} dy p(y) \prod_{i=1}^n p(y|x_i). \tag{C1}$$

Here, $p(y)$ is the probability density of an extra random variable Y, which assumes values in the domain Ω_y. Furthermore, $p(y|x_i)$ is a transition probability: it gives the probability density of X_i given the value y of the random variable Y.

The structure given by Eq. (C1) allows us to read the realizations of the correlated set $\{X_i\}_{i=1}^n$ as an average over realizations of a set of identical random variables with the joint probability density $\prod_{i=1}^n p(y|x_i)$. A similar interpretation can be obtained for the transition density $T_n(x_1,\ldots x_n|x_{n+1})$. In fact, by using that $T_n(x_1,\ldots x_n|x_{n+1}) = P_{n+1}(x_1,\ldots,x_{n+1})/P_n(x_1,\ldots,x_n)$, it can be written as

$$T_n(x_1,\ldots x_n|x_{n+1}) = \int_{\Omega_y} dy p_n(\{x_i\}|y)\; p(y|x_{n+1}), \tag{C2}$$

where $p(y|x_{n+1})$ was introduced previously while $p_n(\{x_i\}|y)$ is

$$p_n(\{x_i\}|y) = \frac{\prod_{i=1}^n p(y|x_i)}{\int_{\Omega_y} dy' p(y') \prod_{j=1}^n p(y'|x_j)} p(y). \tag{C3}$$

Therefore, $T_n(x_1,\ldots x_n|x_{n+1})$ is set by $p(y|x_{n+1})$, where now the probability density $p_n(\{x_i\}|y)$ of the random variable Y [see Eq. (C2)] depends on all previous values $\{X_i\}_{i=1}^n$. Hence, $p_n(x'|y)$ can be read as the conditional probability density of the random variable Y "given" the previous history defined by the set of values $\{x_i\}_{i=1}^n$. Moreover, it is simple to check that Eq. (C2) satisfies the hierarchical equations defined by Eq. (7).

The sum variable Eq. (25) can be straightforwardly characterized from Eqs. (C1) and (28). We get

$$G_{w_n}(k) = \int_{\Omega_y} dy p(y)[G(y|k/n)]^n, \tag{C4}$$

where $G(y|k) \equiv \int_{-\infty}^{+\infty} dx e^{ikx} p(y|x)$. In the asymptotic limit, Eq. (29), assuming valid the law of large numbers for the transition $p(y|x)$, from Eqs. (30) and (31) it follows that

$$P(w) = \int_{\Omega_y} dy p(y) \delta(w - \bar{x}_y), \tag{C5}$$

where the mean value $\bar{x}_y$ is a function of y,

$$\bar{x}_y \equiv \int dx p(y|x) x. \tag{C6}$$

In the case of dichotomic variables, $X_i = 0,1$, with transition probability $p(y|x_i) = y^{x_i}(1-y)^{1-x_i}$, the joint probability $P_n(x_1,\ldots,x_n)$, from Eq. (C1), becomes

$$P_n(x_1,\ldots,x_n) = \int_0^1 dy p(y) \prod_{i=1}^n y^{x_i}(1-y)^{1-x_i}. \tag{C7}$$

Noting that the dependence of the probability $P_n(x_1,\ldots,x_n)$ on the set $\{x_i\}_{i=1}^n$ can be written in terms of the variable $x' \equiv \sum_{i=1}^n x_i$ [Eq. (C7)], from Eq. (C2) it follows the presentation

$$\mathcal{T}_n(x'|x) = \int_0^1 dy p_n(x'|y)\; y^x(1-y)^{1-x}, \tag{C8}$$

where

$$p_n(x'|y) = \frac{y^{x'}(1-y)^{n-x'}}{\int_0^1 d\tilde{y} p(\tilde{y})\; \tilde{y}^{x'}(1-\tilde{y})^{n-x'}} p(y). \tag{C9}$$

Equation (C8) provides a representation for the transition probability $\mathcal{T}_n(x'|x)$ similar to that defined by Eq. (C7).

Given that Eq. (C6) leads to $\bar{x}_y = y$, from Eq. (C5) it follows that $P(w) = p(y)|_{y=w}$. Hence, *any* attractor can be obtained by choosing an arbitrary density $p(y)$.

If one chooses a Beta distribution,

$$p(y) = \frac{\Gamma(\alpha+\alpha')}{\Gamma(\alpha)\Gamma(\alpha')} y^{\alpha-1}(1-y)^{\alpha'-1}, \tag{C10}$$

where $\alpha > 1$ and $\alpha' > 1$ are real parameters, from Eqs. (C7) it is possible to obtain the joint probability densities. In particular, $P_1(x)$ can be written as $P_1(x) = [\alpha'\delta(x) + \alpha\delta(x-1)]/(\alpha+\alpha')$. Additionally, by rewriting Eq. (C8) as $\mathcal{T}_n(x'|x) = \delta(x)\int_0^1 dy p_n(x'|y)(1-y) + \delta(x-1)\int_0^1 dy p_n(x'|y) y$, the transition probability density explicitly reads

$$\mathcal{T}_n(x'|x) = \frac{(n-x'+\alpha')\delta(x) + (x'+\alpha)\delta(x-1)}{n+\alpha+\alpha'}. \tag{C11}$$

In deriving this expression we used the dichotomic property of the random variables.

By introducing the parameter $\lambda = \alpha + \alpha'$, the weights $q_0 = \alpha'/(\alpha+\alpha')$, $q_1 = \alpha/(\alpha+\alpha')$, and the numbers $n_0 = n - x'$, $n_1 = x'$, the transition probability Eq. (C11) can be written as a particular case of the Pólya urn scheme [see Eq. (12)]. In fact, n_0 and n_1 are the number of times that the random variables assumed the values 0 and 1, respectively.

The probability of the variable $X' \equiv \sum_{i=1}^n X_i$, from Eqs. (C7) and (C10), reads

$$P(x') = \binom{n}{x'} \frac{\Gamma(\alpha+\alpha')}{\Gamma(n+\alpha+\alpha')} \frac{\Gamma(n-x'+\alpha')}{\Gamma(\alpha')} \frac{\Gamma(x'+\alpha)}{\Gamma(\alpha)}, \tag{C12}$$

where $\Gamma(x)$ is the Gamma function. The factor $\binom{n}{x'}$ follows from all configurations that lead to the same value of x'.

[1] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Boston, 1968).

[2] W. Feller, *An Introduction to Probability Theory and Applications*, Vols. I and II (John Wiley & Sons, New York, 1967).

[3] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 2nd ed. (North-Holland, Amsterdam, 1992).

[4] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 2004).

[5] P. Lévy, *Théorie de lÁdditions des Variables Aléatoires* (Gauthier Villars, Paris, 1954).

[6] E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958); J.-P. Bouchaud and M. Mézard, Universality classes for extreme-value statistics, J. Phys. A **30**, 7997 (1997).

[7] E. Bertin, Global Fluctuations and Gumbel Statistics, Phys. Rev. Lett. **95**, 170601 (2005); E. Bertin and M. Clusel, Generalized extreme value statistics and sum of correlated variables, J. Phys. A **39**, 7607 (2006); M. Clusel and E. Bertin, Interplay between sum and extreme value statistics, Int. J. Mod. Phys. B **22**, 3311 (2008).

[8] F. Baldowin and A. L. Stella, Central limit theorem for anomalous scaling due to correlations, Phys. Rev. E **75**, 020101(R) (2007); A. L. Stella and F. Baldowin, Anomalous scaling due to correlations: Limit theorems and self-similar processes, J. Stat. Mech. (2010) P02018.

[9] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics* (Springer, Berlin, 2009).

[10] G. Miritello, A. Pluchino, and A. Rapisarda, Central limit behavior in the Kuramoto model at the "edge of chaos," Phys. A **388**, 4818 (2009); A. Pluchino, A. Rapisarda, and C. Tsallis, A closer look at the indications of q-generalized central limit theorem behavior in quasistationary states of the HMF model, *ibid.* **387**, 3121 (2008); A. Pluchino and A. Rapisarda, Nonergodicity and central limit behavior for long-range interactions, Europhys. Lett. **80**, 26002 (2007).

[11] U. Tirnakli, C. Beck, and C. Tsallis, Central limit behavior of deterministic dynamical systems, Phys. Rev. E **75**, 040106(R) (2007); U. Tirnakli, C. Tsallis, and C. Beck, Closer look at time averages of the logistic map at the edge of chaos, *ibid.* **79**, 056209 (2009); U. Tirnakli, H. J. Jensen, and C. Tsallis, Restricted random walk model as a new testing ground for the applicability of q-statistics, Europhys. Lett. **96**, 40008 (2011).

[12] J. Ruseckas, Probabilistic model of N correlated binary random variables and nonextensive statistical mechanics, Phys. Lett. A **379**, 654 (2015); D. O'Malley, V. V. Vesselinov, and J. H. Cushman, Diffusive mixing and Tsallis entropy, Phys. Rev. E **91**, 042143 (2015).

[13] R. Hanel, S. Thurner, and C. Tsallis, Limit distributions of scale-invariant probabilistic models of correlated random variables with q-Gaussian as an explicit example, Eur. Phys. J. B **72**, 263 (2009).

[14] A. Rodríguez, V. Schwämmle, and C. Tsallis, Strictly and asymptotically scale invariant probabilistic models of N correlated binary random variables having q-Gaussians as $N \to \infty$ limiting distributions, J. Stat. Mech. (2008) P09006; S. Umarov, C. Tsallis, and S. Steinberg, On a q-Central Limit Theorem consistent with nonextensive statistical mechanics, Milan J. Math. **76**, 307 (2008).

[15] C. Vignat and A. Plastino, Central limit theorem and deformed exponentials, J. Phys. A **40**, F969 (2007); Scale invariance and related properties of q-Gaussian systems, Phys. Lett. A **365**, 370 (2007).

[16] A. Rodríguez and C. Tsallis, A dimension scale-invariant probabilistic model based on Leibniz-like pyramids, J. Math. Phys. **53**, 023302 (2012).

[17] E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrígues, On a generalization of the binomial distribution and its Poisson-like limit, J. Stat. Phys. **146**, 264 (2012); H. Bergeron, E. M. F. Curado, J. P. Gazeau, and L. M. C. S. Rodrigues, Symmetric generalized binomial distributions, J. Math. Phys. **54**, 123301 (2013); G. A. Casas, F. D. Nobre, and E. M. F. Curado, Nonlinear Ehrenfest's urn model, Phys. Rev. E **91**, 042139 (2015).

[18] C. Beck, Dynamical Foundations of Nonextensive Statistical Mechanics, Phys. Rev. Lett. **87**, 180601 (2001); Non-additivity of Tsallis entropies and fluctuations of temperature, Europhys. Lett. **57**, 329 (2002); C. Beck and E. G. D. Cohen, Superstatistics, Phys. A **322**, 267 (2003); H. Touchette and C. Beck, Asymptotics of superstatistics, Phys. Rev. E **71**, 016131 (2005); S. Abe, C. Beck, and E. G. D. Cohen, Superstatistics, thermodynamics, and fluctuations, *ibid.* **76**, 031102 (2007).

[19] A. A. Budini, Extended q-Gaussian and q-exponential distributions from gamma random variables, Phys. Rev. E **91**, 052113 (2015).

[20] B. de Finetti, Funzione caratteristica di un fenomeno aleatorio, Memorie della R. Accademia Nazionale dei Lincei **6**, 86 (1930); La prévision: Ses lois logiques, ses sources subjectives, Ann. Inst. H. Poincaré **7**, 1 (1937).

[21] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, New York, 2003).

[22] E. Hewitt and L. J. Savage, Symmetric measures on Cartesian products, Trans. Am. Math. Soc. **80**, 470 (1955).

[23] D. Heath and W. Sudderth, de Finetti theorem on exchangeable variables, Am. Stat. **30**, 188 (1976).

[24] J. F. C. Kingman, Uses of exchangeability, Ann. Prob. **6**, 183 (1978).

[25] N. L. Johnson and S. Kotz, *Urn Models and Their Application*, (John Wiley & Sons, New York, 1977).

[26] D. Blackwell and J. B. MacQueen, Fergurson distributions via Pólya urn schemes, Ann. Stat. **1**, 353 (1973).

[27] J. Pitman, *Combinatorial Stochastic Processes* (Springer, Berlin, 2006).

[28] B. Bassetti, M. Zarei, M. C. Lagomarsino, and G. Bianconi, Statistical mechanics of the "Chinese restaurant process": Lack of self-averaging, anomalous finite-size effects, and condensation, Phys. Rev. E **80**, 066118 (2009).

[29] We choose the notation $T_{n-1}(x_1, \ldots x_{n-1}|x_n)$ instead of the reversed one $T_{n-1}(x_n|x_{n-1} \ldots x_1)$ for facilitating the writing of some expressions (see Appendix A).

[30] P. Lévy, Exemples de processus pseudo-Markoviens, Comptes Rendus Académie Sciences, Paris **228**, 2004 (1949); W. Feller, Non-Markovian processes with the semigroup property, Ann. Math. Stat. **30**, 1252 (1959).

[31] H. J. Hilhorst, Central limit theorems for correlated variables: Some critical remarks, Braz. J. Phys. **39**, 371 (2009); H. J. Hilhorst and G. Schehr, A note on q-Gaussian and non-Gaussians in statistical mechanics, J. Stat. Mech. (2007) P06003.