Weak ergodicity breaking induced by global memory effects

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 14 May 2016; published 8 August 2016)

We study the phenomenon of weak ergodicity breaking for a class of globally correlated random walk dynamics defined over a finite set of states. The persistence in a given state or the transition to another one depends on the whole previous temporal history of the system. A set of waiting time distributions, associated to each state, sets the random times between consecutive steps. Their mean value is finite for all states. The probability density of time-averaged observables is obtained for different memory mechanisms. This statistical object explicitly shows departures between time and ensemble averages. While the residence time in each state may have a divergent mean value, we demonstrate that this condition is in general not necessary for breaking ergodicity. Hence, we conclude that global memory effects are an alternative mechanism able to induce ergodicity breaking without involving power-law statistics. Analytical and numerical calculations support these results.

DOI: 10.1103/PhysRevE.94.022108

I. INTRODUCTION

Ergodicity plays a fundamental role in the formulation of statistical physics. This property is usually stated by saying that ensemble average and time average of observables are equals, the latter being taken in the long time limit. In contrast with thermodynamical systems, where the lack or ergodicity is induced by a spontaneous symmetry breaking [1], the disparity between ensemble and time averages can also be found as an emergent property of complex (nonequilibrium) systems dynamics. In this context, the discrepancy with respect to ergodicity is named weak ergodicity breaking (EB) [2]. It is known that dynamics able to induce power-law statistics for the relevant observables lead to weak EB [3].

In systems that admit a description through a stochastic dynamics, time averages in the presence of weak EB may remain random even in the long time limit. Their statistics, termed as weakly nonergodic statistical physics [4,5], define a still very active line of research. Continuous-time random walks form a natural frame where these effects were studied [4–9]. Furthermore, diverse kinds of complex anomalous diffusion processes are a natural partner of weak EB [10–20]. In addition to its theoretical interest, weak EB was also found in different physical and biological systems [21–36]. Weak EB can be studied in systems that have associated a stationary state, for example random walks on finite domains, and also in nonstationary systems such as unbounded diffusive ones. Different measures of EB are naturally associated to each case (see, for example, Refs. [4] and [28], respectively).

For continuous-time random walks leading to weak EB the times between (renewal) transition events are governed by waiting time distributions with power-law tails (see, for example, Refs. [4,5]). Independently of the system dimensionality and stochastic dynamics, weak EB is in general associated with or related to some underlying statistically self-similar (effective) process. One of the main goals of this paper is to show that there exist dynamics where this requisite is not necessary. We demonstrate that systems whose dynamics involves global memory effects may also develop EB. Furthermore, we establish that the lack of ergodicity induced by the memory effects may happen even in the absence of statistical properties (residence times) characterized by dominant power-law distributions.

Global memory (or correlation) effects refer to systems whose stochastic dynamics at a given time depends on their whole previous temporal history (trajectory). These kinds of (highly non-Markovian) dynamics have been studied previously [37–45], mainly as a mechanism that induces superdiffusion. In contrast, here we study random walk processes defined over a finite set of states where the persistence in a given state or the transition to another one depends on the previous system trajectory. The random times between consecutive steps are defined by a set of waiting time distributions with finite average times. In addition, our main results rely on alternative memory mechanisms. They are related to Pólya urn dynamics [46–50], which is one of the simplest models of a contagion process, being of interest in various disciplines [47].

We demonstrate that, in contrast to other global correlation mechanisms [37–39], the urnlike dynamics is able to induce weak EB. In fact, in the long time limit, its memory mechanism is equivalent to a memoryless dynamics with different random transition rates for each realization. This randomness leads to the discrepancy between time and ensemble averages. Interestingly, in contrast with renewal random walks [4,5], the departure from ergodicity arises even when the statistics of the residence times in each state are not dominated by power-law behaviors. Hence, the random times that the system spent in a given state before jumping to another one are characterized by finite averages.

We remark that there exist previous contributions where the interplay between memory effects and weak EB was analyzed [51,52]. Nevertheless, the memory mechanisms cannot be mapped with the present ones. Furthermore, the dynamics develops on unbounded domains whereas here it develops over a finite number of states, a situation that may be of interest in biophysical systems (finite domains).

The paper is organized as follows. In Sec. II, we introduce the globally correlated random walk model. The probability density of time-averaged observables is obtained in general. In Sec. III, we study three different global memory mechanisms: the elephant random walk model, a random walk driven by an urnlike dynamics, and an imperfect case of the last one. In Sec. IV, for all models, we obtain the probability density of the residence times. Section V is devoted to the conclusions. Analytical calculations that support the main results are presented in the appendices.

II. FINITE RANDOM WALK WITH GLOBAL MEMORY EFFECTS

In this section we introduce the globally correlated random walk model and study its properties. The probability density of time-averaged observables is also obtained.

A. Model

The system is characterized by a finite set of states $\mu = 1,\ldots,L$. To each state μ we assign a waiting time distribution $w_\mu(t)$, which gives the statistics of times between consecutive steps of the stochastic dynamics. We assume that all average times

$$\tau_\mu \equiv \int_0^\infty dt\, w_\mu(t) t \tag{1}$$

are finite ($\tau_\mu < \infty$).

The stochastic dynamics is as follows. At the beginning (initial time), each state is selected in agreement with a set of probabilities $\{p_\mu\}_{\mu=1}^L$, $0 \leqslant p_\mu \leqslant 1$, normalized as $\sum_{\mu=1}^L p_\mu = 1$. Given that a state μ is selected, the system remains in it during a random time selected in agreement with the waiting time distribution $w_\mu(t)$. After this step, the system may remain in the same state or jump to another one. Hence, it may *persist* in the same state, remaining an extra time interval chosen in agreement with the same waiting time distribution, or *jump* to a different state with a different waiting time distribution. This dynamic repeats itself in time after each step, where a step refers to the process of selecting the next state.

The state corresponding to the next step is chosen in agreement with a conditional probability $\mathcal{T}_n(\{n_1,n_2,\ldots,n_L\}|\mu)$ [denoted as $\mathcal{T}_n(\{n_\nu\}|\mu)$]. Here, n indicates the number of steps performed up to the present time, while n_ν gives the number of times that each state ν was chosen previously. Then, $n = \sum_{\nu=1}^L n_\nu$. The dependence of the process on the whole previous trajectory (global correlation) is given by the dependence of $\mathcal{T}_n(\{n_\nu\}|\mu)$ on the set $\{n_\nu\}_{\nu=1}^L$. The previous definitions completely characterize the stochastic dynamics in terms of the initial probabilities $\{p_\mu\}_{\mu=1}^L$, the waiting time distributions $\{w_\mu(t)\}_{\mu=1}^L$, and the conditional (or transition) probabilities $\mathcal{T}_n(\{n_\nu\}|\mu)$.

For the studied models [see Eqs. (15), (16), and (21)], as a consequence of the memory effects, the following property is observed. In the long time limit ($t \to \infty$), which also corresponds to a divergent number of steps ($n \to \infty$), the fractions

$$f_\mu = \lim_{n\to\infty} \frac{n_\mu}{n}, \tag{2}$$

$\sum_{\mu=1}^L f_\mu = 1$, may become random variables whose values depend on each particular realization. Their probability density is denoted by $\mathcal{P}(\{f_\mu\})$, which satisfies the normalization condition $\int_\Lambda df_1 \cdots df_{L-1} \mathcal{P}(\{f_\mu\}) = 1$. Here, Λ is the region defined by the condition $\sum_{\nu=1}^L f_\mu = 1$. The average of f_μ over an ensemble realizations, denoted by $\langle\cdots\rangle$, is

$$\langle f_\mu\rangle = \int_\Lambda df_1 \cdots df_{L-1}\, f_\mu \mathcal{P}(\{f_\nu\}). \tag{3}$$

At a given time t, with $P_\mu(t)$ we denote the (ensemble) probability $[\sum_{\mu=1}^L P_\mu(t) = 1]$ that the system is in the (arbitrary) state μ. This object is characterized in Appendix A from the dynamics defined previously. The stationary probability reads $P_\mu^{\text{st}} \equiv \lim_{t\to\infty} P_\mu(t)$. It can be written in terms of $\mathcal{P}(\{f_\nu\})$ as

$$P_\mu^{\text{st}} = \left\langle \frac{f_\mu \tau_\mu}{\sum_{\mu'=1}^L f_{\mu'}\tau_{\mu'}} \right\rangle, \tag{4}$$

where τ_μ is defined by Eq. (1). In Appendix A we also derive this result. Basically it says that in each realization the system reaches a (random) stationary state defined by the weights $(f_\mu\tau_\mu)/\sum_{\mu'=1}^L f_{\mu'}\tau_{\mu'}$. In consequence, P_μ^{st} depends on which memory mechanism drives the stochastic dynamics.

B. Time-averaged observables

To each state μ, we assign an observable with value $\mathcal{O}_\mu$. Hence, each realization of the random walk defines a corresponding trajectory $\mathcal{O}(t)$. In the stationary regime, its *ensemble average* $\langle\mathcal{O}\rangle_{\text{st}} \equiv \lim_{t\to\infty}\langle\mathcal{O}(t)\rangle = \lim_{t\to\infty}\sum_{\mu=1}^L P_\mu(t)\mathcal{O}_\mu$ is

$$\langle\mathcal{O}\rangle_{\text{st}} = \sum\nolimits_{\mu=1}^L P_\mu^{\text{st}}\mathcal{O}_\mu, \tag{5}$$

where the weights follow from Eq. (4). On the other hand, its *time average* is defined as $\mathcal{O} \equiv \lim_{t\to\infty}(1/t)\int_0^t dt'\mathcal{O}(t')$, which leads to

$$\mathcal{O} = \lim_{t\to\infty}\sum_{\mu=1}^L \left(\frac{t_u}{t}\right)\mathcal{O}_\mu. \tag{6}$$

Here, t_u is the total residence time in the state μ in the interval $(0,t)$. Hence, $\sum_{\mu=1}^L t_u = t$.

Even when a long time limit is present in the previous definition, the observable $\mathcal{O}$ may be a random object that depends on each particular realization. Its probability density can be written as

$$P(\mathcal{O}) = \lim_{t\to\infty}\left\langle \delta\left(\mathcal{O} - \sum\nolimits_{\mu=1}^L \frac{t_u}{t}\mathcal{O}_\mu\right)\right\rangle, \tag{7}$$

where, as before, $\langle\cdots\rangle$ denotes the average over an ensemble of realizations and $\delta(x)$ is the Dirac delta function. Now, our goal is to calculate this object for the dynamics defined previously.

Given that the waiting time distributions are characterized by a finite average time τ_μ, Eq. (1), after invoking the law of large numbers, in the long time limit the total residence time t_u in each state can be approximated as $t_u \simeq n_\mu\tau_\mu$. Consistently, the present time is $t \simeq \sum_{\mu=1}^L n_\mu\tau_\mu$. Hence, we can write

$$\lim_{t\to\infty}\frac{t_u}{t} \simeq \lim_{n\to\infty}\frac{n_\mu\tau_\mu}{\sum_{\mu'=1}^L n_{\mu'}\tau_{\mu'}} = \frac{f_\mu\tau_\mu}{\sum_{\mu'=1}^L f_{\mu'}\tau_{\mu'}}, \tag{8}$$

where the last relation follows from Eq. (2). Taking into account that the fractions $\{f_\mu\}_{\mu=1}^L$ are characterized by the distribution $\mathcal{P}(\{f_\mu\})$, Eq. (7) becomes

$$P(\mathcal{O}) = \int_\Lambda df_1 \cdots df_{L-1}\mathcal{P}(\{f_\mu\}) \times \delta\left(\mathcal{O} - \sum\nolimits_{\mu=1}^{L} \frac{f_\mu \tau_\mu}{\sum_{\mu'=1}^{L} f_{\mu'}\tau_{\mu'}}\mathcal{O}_\mu\right). \quad (9)$$

Therefore, $P(\mathcal{O})$ can be completely characterized after knowing the distribution $\mathcal{P}(\{f_\mu\})$. Notice that the specific structure of the waiting time distributions $\{w_\mu(t)\}_{\mu=1}^{L}$ only appears through the average times $\{\tau_\mu\}_{\mu=1}^{L}$, Eq. (1).

C. Ergodicity and localization

For an ergodic dynamics the fractions f_μ [Eq. (2)] must be characterized by their ensemble average, Eq. (3). Hence,

$$\mathcal{P}(\{f_\mu\}) = \delta(f_1 - \langle f_1\rangle)\delta(f_2 - \langle f_2\rangle)\cdots\delta(f_L - \langle f_L\rangle). \quad (10)$$

Inserting this expression into Eq. (9), it follows the distribution

$$P(\mathcal{O}) = \delta(\mathcal{O} - \langle\mathcal{O}\rangle_{\mathrm{st}}), \quad (11)$$

where $\langle\mathcal{O}\rangle_{\mathrm{st}}$ is given by Eq. (5) with the weights

$$P_\mu^{\mathrm{st}} = \frac{\langle f_\mu\rangle\tau_\mu}{\sum_{\mu'=1}^{L}\langle f_{\mu'}\rangle\tau_{\mu'}}. \quad (12)$$

From Eqs. (4) and (10), we note that these weights correspond to the stationary probabilities of each state μ in the ergodic case. Hence, time averages and ensemble averages do in fact coincide.

The maximal departure with respect to ergodicity happens when the dynamics localize; that is, the system remains in the initial condition. This case corresponds to

$$\mathcal{P}(\{f_\mu\}) = \sum_{\mu=1}^{L} p_\mu \delta(f_1)\cdots\delta(f_\mu - 1)\cdots\delta(f_L). \quad (13)$$

Hence, Eq. (9) becomes

$$P(\mathcal{O}) = \sum_{\mu=1}^{L} p_\mu \delta(\mathcal{O} - \mathcal{O}_\mu). \quad (14)$$

These limits are reached by the following memory mechanisms.

III. EXAMPLES

In the examples worked below, the stochastic dynamics may reach both the ergodic and localized regimes, Eqs. (11) and (14), respectively. The distribution $\mathcal{P}(\{f_\mu\})$ can be explicitly calculated and then the nonergodic properties characterized through Eq. (9).

A. Elephant random walk model

This correlation model has been studied extensively in the recent literature as a mechanism for inducing superdiffusion [37–39]. In the present context, it is defined by the transition probability

$$\mathcal{T}_n(\{n_\nu\}|\mu) = \varepsilon q_\mu + (1-\varepsilon)\frac{n_\mu}{n}. \quad (15)$$

The positive weights $0 < q_\mu < 1$ are extra parameters normalized as $\sum_{\mu=1}^{L} q_\mu = 1$. The parameter ε assumes values in the interval $[0,1]$. The stochastic dynamics can be read as follows. With probability ε, and independently of the previous history, the new state is chosen in agreement with the probabilities $\{q_\mu\}_{\mu=1}^{L}$. On the other hand, with probability $1-\varepsilon$ each state is chosen in agreement with the weights $\{n_\mu/n\}_{\mu=1}^{L}$, which in fact depend on the whole previous history of the process.

For $\varepsilon = 1$, the selection of the new state is completely random and independent of the previous history. Therefore, the system is ergodic in this case [Eq. (11)]. On the other hand, for $\varepsilon = 0$ the dynamics localize; that is, the system remains in the initial condition[Eq. (14)].

Even when the dynamics reaches the ergodic and localized regime, for intermediate values $0 < \varepsilon < 1$ the dynamics is ergodic. This property is demonstrated in Appendix B. In fact, the distribution $\mathcal{P}(\{f_\mu\})$ is δ distributed, Eq. (10), with $\langle f_\mu\rangle = q_\mu$.

B. Random walk driven by an urnlike dynamics

In the Pólya urn dynamics [46,47] (initially) an urn contains many balls that, for example, are characterized by L different possible colors. At each step, one determines the color of one ball taken at random and puts into the urn one extra ball of the same color. A similar process can be defined by starting the urn with only one ball [48–50] (Blackwell-MacQueen urn). Its dynamics is defined by the following conditional probability, which is taken as the driving memory mechanism.

For the random walk over the $\mu = 1,\ldots,L$ states, we take the conditional probability [48,49]

$$\mathcal{T}_n(\{n_\nu\}|\mu) = \frac{\lambda q_\mu + n_\mu}{n+\lambda}. \quad (16)$$

As before, the set of parameters $\{q_\mu\}_{\mu=1}^{L}$ is normalized to 1. Instead, λ is a positive free parameter. For $\lambda \to \infty$ the dynamics loses any dependence on the previous history achieving in consequence an ergodic regime, Eq. (11). On the other hand, for $\lambda = 0$, a localized regime is achieved, Eq. (14). Hence, the intermediate values of λ avoid this regime and in consequence one can define a nontrivial dynamics starting from $n = 1$.

For arbitrary values of λ, the probability density of the (asymptotic) fractions (2) is derived in Appendix C. It can be written as

$$\mathcal{P}(\{f_\mu\}) = \left\{\sum_{\nu=1}^{L}\frac{p_\nu}{q_\nu}f_\nu\right\}D(\{f_\mu\}|\{\lambda q_\mu\}), \quad (17)$$

where $D(\{f_\mu\}|\{\lambda_\mu\})$ is a Dirichlet distribution [48,49],

$$D(\{f_\mu\}|\{\lambda_\mu\}) \equiv \frac{\Gamma(\lambda)}{\prod_{\mu'}\Gamma(\lambda_{\mu'})}\prod_\mu f_\mu^{\lambda_\mu - 1}. \quad (18)$$

Here, $\lambda = \sum_{\mu=1}^{L}\lambda_\mu$. The (ensemble) average fraction reads $\langle f_\mu\rangle = (q_\mu\lambda + p_\mu)/(\lambda+1)$. When $p_\nu = q_\nu$, due to the normalization $\sum_{\nu=1}^{L} f_\nu = 1$, the first factor in Eq. (17) does not contribute, and $\langle f_\mu\rangle = q_\mu$.

We notice that $\mathcal{P}(\{f_\mu\})$ [Eq. (17)] depends on the initial conditions $\{p_\mu\}_{\mu=1}^{L}$. This property arises from the strong memory effects that drive the underlying stochastic dynamics. Nevertheless, this dependence is not able to cancel any of the stationary fractions. In consequence, the initial conditions are not relevant for breaking ergodicity or not. In fact, given that $\mathcal{P}(\{f_\mu\})$ departs from Eq. (10), this model leads to EB. The distribution $P(\mathcal{O})$ [Eq. (9)] can be evaluated from Eq. (17).

As an example, we consider a *two-level system*, where the observable is defined by $\{\mathcal{O}_\mu\} \to (\mathcal{O}_2,\mathcal{O}_1)$, with $\mathcal{O}_1 \leqslant \mathcal{O} \leqslant \mathcal{O}_2$. After integration, we get

$$P(\mathcal{O}) = \frac{1}{\mathcal{N}} \frac{[\omega_2(\mathcal{O}_2 - \mathcal{O})]^{\lambda_1 - 1}[\omega_1(\mathcal{O} - \mathcal{O}_1)]^{\lambda_2 - 1}}{[\omega_2(\mathcal{O}_2 - \mathcal{O}) + \omega_1(\mathcal{O} - \mathcal{O}_1)]^{\lambda_1+\lambda_2}}, \tag{19}$$

where for shortening the expression we introduced the parameters $\lambda_1 \equiv \lambda q_1$, $\lambda_2 \equiv \lambda q_2$, and the weights

$$\omega_1 \equiv \frac{\tau_1}{\tau_1 + \tau_2}, \qquad \omega_2 \equiv \frac{\tau_2}{\tau_1 + \tau_2}. \tag{20}$$

Here, τ_1 and τ_2 are the average times corresponding to the two waiting time distributions $w_1(t)$ and $w_2(t)$, respectively [Eq. (1)]. The normalization constant reads $\mathcal{N}^{-1} = (\mathcal{O}_2 - \mathcal{O}_1)\omega_1\omega_2\Gamma(\alpha_1 + \alpha_2)/\Gamma(\alpha_1)\Gamma(\alpha_2)$. For simplicity, in the previous expressions we assumed the initial condition $p_\mu = q_u$. The case $p_\mu \neq q_u$ can be recovered from these expressions [see Eqs. (17) and (18)].

The model (16) demonstrates that global memory effects may lead to EB. This result has a close relation with the breakdown of the standard central limit theorem for globally correlated random variables [50]. On the other hand, as shown in Sec. IV, depending on the values of λ, here EB arises because the average residence times in each state may be divergent; that is, their probability density is characterized by power-law tails. The next modified dynamics also develops EB but does not involve power-law statistics.

C. Imperfect urnlike model

Here, we consider a model that can be seen as an imperfect case of the previous one. We consider the possibility of having random state selections that do not depend on the previous system history. The transition probability reads

$$\mathcal{T}_n(\{n_\nu\}|\mu) = \varepsilon q_\mu + (1-\varepsilon)\frac{\lambda q_\mu + M_\mu}{M+\lambda}. \tag{21}$$

The set $\{q_\mu\}_{\mu=1}^L$ is normalized as before, $0 \leqslant \varepsilon \leqslant 1$, and $\lambda \geqslant 0$. Hence, with probability ε each state μ, independently of the previous trajectory, is chosen with weight q_μ. Complementarily, with probability $1-\varepsilon$ the state is chosen in agreement with the urn mechanism, Eq. (16). In fact, here M_μ is the number of times that the state μ was chosen *with* the urn dynamics. Furthermore, M is the number of times that the urn mechanism was applied: $M = \sum_{\mu=1}^L M_\mu$. In contrast with the elephant walk model [Eq. (15)], here the contribution proportional to ε can be thought of as an error in the application of the urn dynamics.

In order to clarify the stochastic dynamics induced by Eq. (21), in Fig. 1 we plot two realizations [Figs. 1(a) and 1(c)] for a two-level system with $\mathcal{O}_2 = 1$ and $\mathcal{O}_1 = -1$. Hence, the observable realizations switch between these two values. The waiting time distributions are exponential ones,

$$w_\mu(t) = \gamma_\mu \exp[-\gamma_\mu t], \tag{22}$$

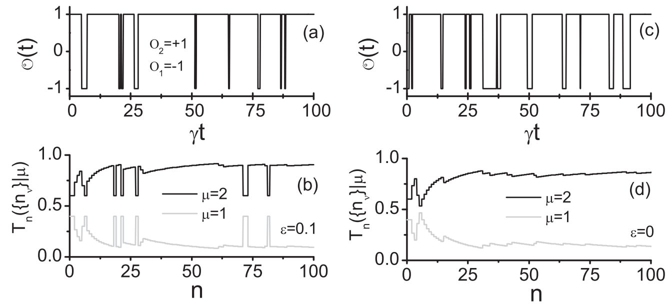

FIG. 1. Realizations of a two-level system (a, c) with observable $\{\mathcal{O}_2 = 1, \mathcal{O}_1 = -1\}$ driven by an urnlike dynamics, jointly with the corresponding conditional probabilities $\mathcal{T}_n(\{n_\nu\}|\mu)$ [Eqs. (16) and (21)] as a function of n (b, d). The parameters are $\lambda = 2$, $p_1 = q_1 = 0.4$, and $p_2 = q_2 = 0.6$. The waiting time distributions are exponential functions [Eq. (22)] with $\gamma_1 = \gamma_2 = \gamma$. In (a) and (b) $\varepsilon = 0.1$, while in (c) and (d) we take $\varepsilon = 0$.

with $\mu = 1,2$. In Figs. 1(b) and 1(d) we plotted the conditional probability $\mathcal{T}_n(\{n_\nu\}|\mu)$ as a function of n. For clarity, each value is continued in the real interval $(i-1,i)$. Figures 1(a) and 1(b) correspond to $\varepsilon = 0.1$, that is, Eq. (21), while Figs. 1(c) and 1(d) correspond to $\varepsilon = 0$, that is, Eq. (16). In both cases $\mathcal{T}_n(\{n_\nu\}|\mu)$ attains stationary values for increasing n [Eq. (2)]. Nevertheless, in the case $\varepsilon = 0.1$ at random values of n the conditional probability collapses to the value q_μ. This effect gives the error or imperfection with respect to the case $\varepsilon = 0$.

The probability distribution of the asymptotic fractions [Eq. (2)] associated to Eq. (21) is given by

$$\mathcal{P}(\{f_\mu\}) = \left\{\sum_{\nu=1}^L \frac{p_\nu}{q_\nu}\frac{f_\nu - \varepsilon q_\nu}{1-\varepsilon}\right\}\frac{1}{(1-\varepsilon)^{L-1}} \times D\left(\left\{\frac{f_\mu - \varepsilon q_\mu}{1-\varepsilon}\right\}\Big|\{\lambda q_\mu\}\right), \tag{23}$$

where $D(\{f_\mu\}|\{\lambda_\mu\})$ is the Dirichlet distribution [Eq. (18)]. Furthermore, each fraction is restricted to the domain

$$\varepsilon q_\mu \leqslant f_\mu \leqslant 1 - \varepsilon(1-q_\mu). \tag{24}$$

In this case, the average fraction reads

$$\langle f_\mu \rangle = \frac{q_\mu(\lambda+\varepsilon) + p_\mu(1-\varepsilon)}{(\lambda+1)}. \tag{25}$$

Equation (23) is related to Eq. (17) by the change of variables $f_\mu \to \varepsilon q_\mu + (1-\varepsilon)f_\mu$. This relation follows by considering the asymptotic limits of Eqs. (21) and (16), and by using that the law of large numbers applies to the error mechanism. For $\varepsilon = 0$ the previous expressions recover the previous case, Eq. (17). Interestingly, the effect of introducing the imperfect mechanism is to reduce the domain of each fraction f_μ, Eq. (24).

From Eqs. (9) and (23) we can calculate the distribution of the time-averaged observable. Below we consider a *two-level system* with $\mathcal{O}_1 < \mathcal{O} < \mathcal{O}_2$ and initial condition $p_\mu = q_u$. This case straightforwardly allows us to reconstruct the case

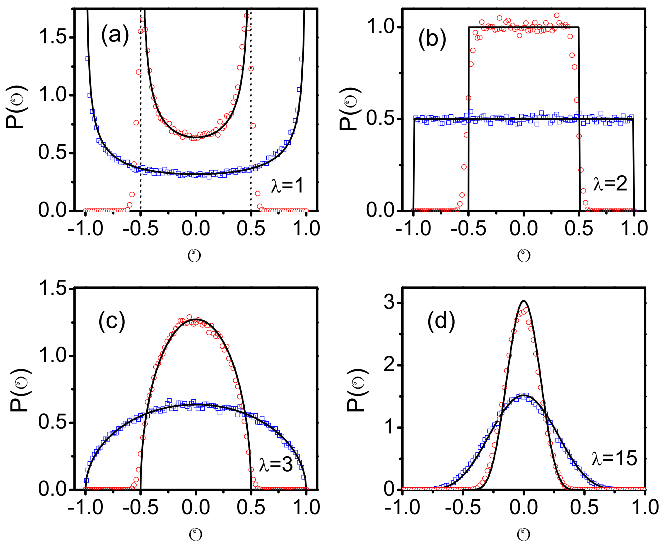

FIG. 2. Probability density of the time-averaged observable $\mathcal{O}$. We take a two-level system driven by an urnlike dynamics with different values of λ and ε. The solid lines correspond to the analytical expressions Eqs. (19) and (26). The waiting time distributions are exponential functions [Eq. (22)] with $\gamma_1 = \gamma_2 = \gamma$. In all plots we take $p_1 = p_2 = q_1 = q_2 = 1/2$. The (red) circles ($\varepsilon = 0.5$) and (blue) squares ($\varepsilon = 0$) correspond to numerical simulations. λ is indicated in each plot.

$p_\mu \neq q_u$. We get

$$P(\mathcal{O}) = \frac{1}{\mathcal{N}_\varepsilon}\big[\omega_2(\mathcal{O}_2 - \mathcal{O}) - \omega_1^\varepsilon(\mathcal{O} - \mathcal{O}_1)\big]^{\lambda_1 - 1} \times\big[\omega_1(\mathcal{O} - \mathcal{O}_1) - \omega_2^\varepsilon(\mathcal{O}_2 - \mathcal{O})\big]^{\lambda_2 - 1} \times\frac{1}{[\omega_2(\mathcal{O}_2 - \mathcal{O}) + \omega_1(\mathcal{O} - \mathcal{O}_1)]^{\lambda_1+\lambda_2}}. \tag{26}$$

The possible values of the time-averaged observable are restricted to the domain $\mathcal{O}_{\min} \leqslant \mathcal{O} \leqslant \mathcal{O}_{\max}$, where

$$\mathcal{O}_{\max} \equiv \frac{\mathcal{O}_2 + \mathcal{O}_1\omega_1^\varepsilon\omega_2^{-1}}{1 + \omega_1^\varepsilon\omega_2^{-1}}, \qquad \mathcal{O}_{\min} \equiv \frac{\mathcal{O}_1 + \mathcal{O}_2\omega_2^\varepsilon\omega_1^{-1}}{1 + \omega_2^\varepsilon\omega_1^{-1}}. \tag{27}$$

Furthermore, we introduced the parameters

$$\omega_1^\varepsilon \equiv \omega_1\frac{\varepsilon q_1}{1 - \varepsilon q_1}, \quad \omega_2^\varepsilon \equiv \omega_2\frac{\varepsilon q_2}{1 - \varepsilon q_2}, \tag{28}$$

while the normalization constant is $\mathcal{N}_\varepsilon^{-1} = (\mathcal{O}_2 - \mathcal{O}_1)\omega_1\omega_2[\Gamma(\lambda)/\Gamma(\lambda_1)\Gamma(\lambda_2)](1-\varepsilon)^{-(\lambda-1)}(1-\varepsilon q_1)^{\lambda_1 - 1}(1 - \varepsilon q_2)^{\lambda_2 - 1}$. Consistently, for $\varepsilon = 0$, Eq. (26) recovers the previous case, Eq. (19). From the previous expression it becomes clear that the error mechanism introduced in Eq. (21) leads to a shrinking of the probability density of the time-averaged observable.

In order to check these results, in Fig. 2, we plot the distribution (26) for a two-level system where as before we take $\mathcal{O}_2 = 1$, $\mathcal{O}_1 = -1$, and the exponential waiting time distributions (22). For each value of λ, we plot the cases $\varepsilon = 0.5$ [Eqs. (21) and (26)] and $\varepsilon = 0$ [Eqs. (16) and (19)]. Consistently, a higher ε leads to a shrinking of the density $P(\mathcal{O})$, which confirms that for $\varepsilon \to 1$ an ergodic regime is achieved, $P(\mathcal{O}) = \delta(\mathcal{O})$. The same happens for increasing λ. On the other hand, the plots show that $P(\mathcal{O})$ may develop different forms such as U and bell shapes, or even uniform ones. Similar dependencies arise when studying renewal random walks with divergent average trapping times [4].

In all cases, the numerical simulations (circles and squares) follow from a time average performed on a time interval with $n = 1 \times 10^3$ steps and 1×10^5 realizations. The theoretical results fit very well the numerical ones.

IV. PROBABILITY DENSITY OF RESIDENCE TIMES

In contrast to the elephant random walk model, the previous urn models develop weak EB. Here, we explore if this property is induced, or not, by a power-law statistics. In fact, for continuous-time random walks with renewal events, EB is induced by the divergence of the average residence time in each state [4]. The residence times are the random times that the system stays or remains in a given state before jumping to another one (see Fig. 1). Here, for the models introduced previously, we calculate their probability density. The calculations are valid for an arbitrary number of states, L.

We consider a single trajectory in the *long time limit*, such that the fractions $\{f_\mu\}_{\mu=1}^L$ [Eq. (2)] can be described by their associated probability density $\mathcal{P}(\{f_\mu\})$ [see Eqs. (17) and (23)]. At the beginning of the residence in a given state μ the first time interval is chosen in agreement with its waiting time distribution $w_\mu(t)$. In each step, the system remains in the same state with probability f_μ, which adds a new random time interval also defined from $w_\mu(t)$. The residence time ends when a different state $\nu \neq \mu$ is chosen. This change occurs with probability $1 - f_\mu$. Therefore, the probability $W_\mu(\{f\}|\tau)\,d\tau$ of leaving the state μ after a residence time τ can be written in the Laplace domain [$g(s) = \int_0^\infty d\tau g(\tau)e^{-s\tau}$] as

$$W_\mu(\{f\}|s) = (1 - f_\mu)w_\mu(s)\sum_{n=0}^\infty f_\mu^n\; w_\mu^n(s). \tag{29}$$

Here, $w_\mu(s)$ is the Laplace transform of the waiting time distribution $w_\mu(t)$ associated to the state μ. The previous expression takes into account all possible ways of leaving the state μ after a given number of steps. It can be rewritten as

$$W_\mu(\{f\}|s) = (1 - f_\mu)\frac{w_\mu(s)}{1 - f_\mu w_\mu(s)}. \tag{30}$$

The density $W_\mu(\{f\}|\tau)$ is a conditional object. In fact, it is defined for a particular realization with random values of the fraction f_μ. Therefore, the probability density of the residence time $W_\mu(t)$ is obtained after averaging over realizations, $W_\mu(t) = \langle W_\mu(\{f\}|t)\rangle$, which is equivalent to an average over the distribution $\mathcal{P}(\{f_\nu\})$ of the set of fractions $\{f_\nu\}_{\nu=1}^L$. Therefore, we get

$$W_\mu(\tau) = \int_\Lambda df_1\cdots df_{L-1}\mathcal{P}(\{f_\nu\})W_\mu(\{f\}|\tau), \tag{31}$$

where $W_\mu(\{f\}|\tau)$ follows from Eq. (30) after Laplace inversion. The average residence time T_μ is defined by

$$T_\mu \equiv \int_0^\infty d\tau\,\tau\,W_\mu(\tau). \tag{32}$$

The previous two expressions can be evaluated for arbitrary waiting time distributions and memory models. For an exponential waiting time distribution $w_\mu(t) = \gamma_\mu \exp[-\gamma_\mu t]$ [Eq. (22)] with mean value $\tau_\mu = 1/\gamma_\mu$ [Eq. (1)], it follows that $w_\mu(s) = \gamma_\mu/(s+\gamma_\mu)$. From Eq. (30), we get $W_\mu(\{f\}|s) = (1-f_\mu)\gamma_\mu/[s+(1-f_\mu)\gamma_\mu]$, which can be inverted as

$$W_\mu(\{f\}|\tau) = (1-f_\mu)\gamma_\mu \exp[-(1-f_\mu)\gamma_\mu \tau]. \tag{33}$$

For an ergodic system, characterized by the probability density $\mathcal{P}(\{f_\nu\})$ given by Eq. (10), from Eq. (31) we get

$$W_\mu(\tau) = (1-\langle f_\mu\rangle)\gamma_\mu \exp[-(1-\langle f_\mu\rangle)\gamma_\mu \tau]. \tag{34}$$

This result is consistent with the definition of the underlying stochastic process that in each step allows the persistence in the same state. In fact, the average residence time is $T_\mu = 1/[\gamma_\mu(1-\langle f_\mu\rangle)]$, indicating an increasing of the average residence time with an increasing of the weight $\langle f_\mu\rangle$. On the other hand, in the localized regime [Eq. (13)], due to the absence of transitions, it is not possible to define $W_\mu(\tau)$.

Taking exponential waiting time distributions Eq. (22), for a *two-level system* [$\mu = 1,2$] characterized by the conditional probability (21) (imperfect urn model), after a simple change of variables, Eqs. (23) and (31) deliver

$$W_\mu^\varepsilon(\tau) = \frac{1}{\mathcal{N}} \int_0^1 df \varphi_\varepsilon \exp[-\varphi_\varepsilon\, \tau] c_\mu\, f^{\lambda_\mu -1}(1-f)^{\lambda_{\mu'}-1}, \tag{35}$$

where the superscript denotes the dependence on the parameter ε. $\lambda_\mu = \lambda q_\mu$ [$\mu = 1,2$] while $\lambda_{\mu'}$ [$\mu' = 2,1$] corresponds to the other system state, $\lambda_{\mu'} = \lambda q_{\mu'} = \lambda(1-q_\mu)$. The initial conditions appear through the contribution

$$c_\mu \equiv \frac{p_\mu}{q_\mu} f + \frac{1-p_\mu}{1-q_\mu}(1-f). \tag{36}$$

The decay rate φ_ε is

$$\varphi_\varepsilon \equiv \gamma_\mu[1-\varepsilon q_\mu - (1-\varepsilon)f], \tag{37}$$

while the normalization constant reads $\mathcal{N}^{-1} = \Gamma(\lambda_1+\lambda_2)/\Gamma(\lambda_1)\Gamma(\lambda_2)$. Straightforwardly, the average residence time, $T_\mu^\varepsilon = \int_0^\infty d\tau\,\tau\, W_\varepsilon(\tau)$, from Eq. (35) can then be written as

$$T_\mu^\varepsilon = \frac{1}{\mathcal{N}} \int_0^1 df \frac{1}{\varphi_\varepsilon} c_\mu f^{\lambda_\mu -1}(1-f)^{\lambda_{\mu'}-1}. \tag{38}$$

Consistently, for $\varepsilon = 1$ Eq. (35) recovers Eq. (34) with $\langle f_\mu\rangle = q_\mu$ [Eq. (25)]. Hence, $\gamma_\mu T_\mu^1 = 1/(1-q_\mu)$. The same results arise when $\lambda \to \infty$. For arbitrary ε and λ, from Eqs. (35) and (38) explicit analytical expressions can be found for both $W_\mu^\varepsilon(\tau)$ and T_μ^ε [see Appendix D].

Interestingly, for $0 < \varepsilon \leqslant 1$ (and any initial condition) the average residence time T_μ^ε is finite [see Eq. (D4)]. This is the main result of this section. In fact, this result demonstrates that weak EB may arise even in the absence of power-law statistical distributions with divergent average residence times. On the other hand, for the case $\varepsilon = 0$, that is, the dynamics defined by the conditional probabilities (16), the average residence time T_μ^0, depending on the parameter values, may be finite or infinite. From Eqs. (38) and (D4) we get

$$\gamma_\mu T_\mu^0 = \frac{\lambda - \left(\frac{1-p_\mu}{1-q_\mu}\right)}{\lambda(1-q_\mu)-1}, \quad \lambda > \frac{1}{(1-q_\mu)} > 1. \tag{39}$$

Consistently, for increasing λ this expression recovers the ergodic case [Eq. (34)], $\lim_{\lambda\to\infty} \gamma_\mu T_\mu^0 = 1/(1-q_\mu)$. In the complementary region of possible values of λ, the average residence time is divergent:

$$\gamma_\mu T_\mu^0 = \infty, \quad \lambda \leqslant \frac{1}{(1-q_\mu)}. \tag{40}$$

This last regime indicates that the density $W_\mu^\varepsilon(\tau)$ develops power-law tails. In fact, for long residence times, $\gamma_\mu\tau \gg 1$, from Eqs. (35) and (D1) it can be approximated as

$$W_\mu^0(\tau) \approx \gamma_\mu C_\mu^0 \left(\frac{1}{\gamma_\mu \tau}\right)^{\lambda(1-q_\mu)+1}, \tag{41}$$

which defines the previous finite and infinite average time regimes. The dimensionless constant reads $C_\mu^0 = (p_\mu/q_\mu)(1-q_\mu)\Gamma(1+\lambda)/\Gamma(q_\mu\lambda)$. When $p_\mu = 0$ ($p_{\mu'} = 1$) the asymptotic behavior becomes $W_\mu^0(\tau) \approx (1/\gamma_\mu\tau)^{\lambda(1-q_\mu)+2}$, while for $W_{\mu'}^0(\tau)$ it is given by Eq. (41). We remark that in general $W_\mu^\varepsilon(\tau)$ ($\varepsilon > 0$) may also develop power-law behaviors. Nevertheless, a multiplicative exponential factor always leads to finite average times [see, for example, Eq. (42) below].

For particular values of the characteristic parameters, the integral results defined by Eqs. (35) and (38) lead to simple analytical expressions. Taking $p_1 = q_1 = 1/2$, $p_2 = q_2 = 1/2$, and $\lambda = 2$ [Fig. 2(b)] the density of residence times becomes

$$W_\mu^\varepsilon(\tau) = \frac{\exp(-\gamma_\varepsilon^+\tau)(1+\gamma_\varepsilon^+\tau) - \exp(-\gamma_\varepsilon^-\tau)(1+\gamma_\varepsilon^-\tau)}{\gamma_\mu\tau^2(1-\varepsilon)}, \tag{42}$$

where for shortening the expression we introduced the rates $\gamma_\varepsilon^+ \equiv \gamma_\mu\varepsilon/2$ and $\gamma_\varepsilon^- \equiv \gamma_\mu(1-\varepsilon/2)$. In the case $\varepsilon = 1$ (ergodic dynamics), we get $W_\mu^\varepsilon(\tau) = (\gamma_\mu/2)T_\mu^1 \exp[-(\gamma_\mu/2)\tau]$. Hence, $T_\mu^1 = 2/\gamma_\mu$. In the case $\varepsilon = 0$ it reduces to

$$W_\mu^0(\tau) = \frac{1}{\gamma_\mu\tau^2}[1-(1+\gamma_\mu\tau)\exp(-\gamma_\mu\tau)], \tag{43}$$

which explicitly shows the presence of dominant power-law tails. The average residence time [Eq. (38)] for arbitrary ε reads

$$\gamma_\mu T_\mu^\varepsilon = \frac{2\operatorname{arctanh}(1-\varepsilon)}{(1-\varepsilon)} = \frac{\ln\left(\frac{2-\varepsilon}{\varepsilon}\right)}{(1-\varepsilon)}, \tag{44}$$

where $\operatorname{arctanh}[x] = \ln\sqrt{\frac{1+x}{1-x}}$ for $x \in (-1,1)$. Thus, T_μ^ε is finite for $0 < \varepsilon \leqslant 1$. Consistently with Eqs. (40) and (43), it diverges for $\varepsilon = 0$, $T_\mu^0 = \lim_{\varepsilon\to 0} T_\mu^\varepsilon = \infty$.

In order to check the previous results we determined the distribution $W_\mu^\varepsilon(\tau)$ from a set of realizations such as those shown in Fig. 1. For the same system as in Fig. 2, the results are shown in Fig. 3. Furthermore, we take $w_1(t) = w_2(t) = \gamma\exp(-\gamma t)$, which implies $W_1^\varepsilon(\tau) = W_2^\varepsilon(\tau)$. Consistently with the previous analytical results [Eq. (42)], for $\varepsilon = 0.5$ [Fig. 3(a)] asymptotically the density of residence times $W_\mu^\varepsilon(\tau)$ is not dominated by power-law behaviors. Instead for

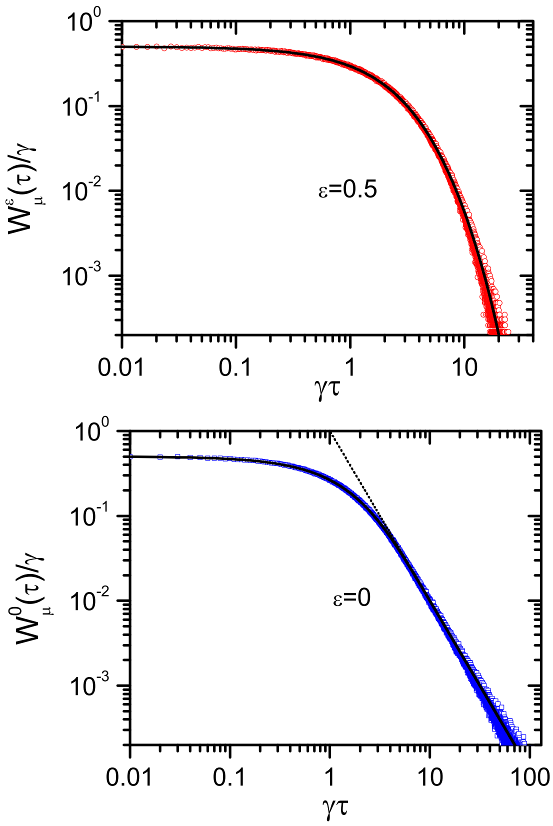

FIG. 3. Probability distribution $W^{\varepsilon}_{\mu}(\tau)$ [$\mu = 1,2$] of the residence times for a two-level system. The solid lines correspond to the analytical result of Eq. (42). The waiting time distributions are exponential functions [Eq. (22)] with $\gamma_1 = \gamma_2 = \gamma$. In both curves, $p_1 = p_2 = q_1 = q_2 = 1/2$, and $\lambda = 2$. The (red) circles correspond to a numerical simulation with $\varepsilon = 0.5$, while the (blue) squares correspond to $\varepsilon = 0$. The dotted line is the asymptotic power-law behavior (41) of Eq. (43).

$\varepsilon = 0$ [Fig. 3(b)] an asymptotic power-law behavior is clearly observed [Eq. (43)]. The numerical and theoretical results are consistent between them.

The numerical probability densities of Fig. 3 were obtained from a set of equally sampled realizations. This means that the same number of data for the random residence times is taken from each realization. We took 5×10^3 realizations with a total length of $n = 5 \times 10^5$ steps. Furthermore, after running the dynamics during 1×10^3 steps (long time limit), 5×10^3 random residence times were taken from each realization.

V. SUMMARY AND CONCLUSIONS

We have introduced a random walk dynamics characterized by global memory mechanisms. Given a finite set of states, in each step the system may remain in the same state or jump to another one. These alternative events are chosen from a conditional probability that depends on the whole previous history of the system. The time between consecutive steps is determined by a set of waiting time distributions, all of them characterized by a finite average time.

We focused the analysis on the ergodic properties of the stochastic dynamics. Hence, we characterized the probability density of time-averaged observables [Eq. (9)]. By analyzing different memory mechanisms, we conclude that global correlations are not a sufficient condition for breaking ergodicity, such as, for example, in the elephant random walk model [Eq. (15)]. On the other hand, alternative urnlike memory mechanisms [Eqs. (16) and (21)] do in fact break ergodicity. In these cases, considering a two-level dynamics, the distribution of time-averaged observables can be found in an explicit analytical way [Eqs. (19) and (26)].

For random walk dynamics over a finite set of states, EB may be induced by a divergent average residence time in each state. In order to check this possibility for the present models, we calculated the probability density of the residence times [Eq. (31)] and the corresponding average residence time [Eq. (32)]. In general, the distributions do not develop asymptotic power-law behaviors consistent with a divergent average residence time. Hence, we conclude that global memory effects are in fact an alternative mechanism that leads to EB without involving power-law statistics. This main conclusion was explicitly checked for two-level dynamics [Eqs. (35) and (38)]. Only for a particular set of values do the residence times have a divergent average. All previous results were confirmed by numerical simulations [see Figs. 2 and 3].

In conclusion, we established that weak EB may also arise in (finite dimensional) systems characterized by global memory effects. This property may emerge even when the relevant variables are not characterized by power-law statistical behaviors.

ACKNOWLEDGMENT

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

APPENDIX A: ENSEMBLE PROBABILITIES AND STATIONARY STATE

Here, we obtain the ensemble probabilities $\{P_\mu(t)\}_{\mu=1}^{L}$ and their corresponding long time limit, Eq. (4).

From the dynamics defined in Sec. II, the probability $P_\mu(t)$ that the system is in the (arbitrary) state μ at time t can be written in the Laplace domain [$g(s) = \int_0^\infty d\tau g(\tau) e^{-s\tau}$] as

$$P_\mu(s) = P_1(\mu)\Phi_\mu(s) + \sum_{n=1}^{\infty} \sum_{\mu_1,\cdots\mu_n} P_{n+1}(\mu_1, \ldots, \mu_n, \mu) \times w_{\mu_1}(s) \cdots w_{\mu_n}(s)\Phi_\mu(s), \quad \text{(A1)}$$

where $\Phi_\mu(s) = [1 - w_\mu(s)]/s$ is the Laplace transform of the survival probability $\Phi_\mu(t) = 1 - \int_0^t dt' w_\mu(t')$. Furthermore, $P_n(\mu_1, \ldots, \mu_n)$ is the probability of obtaining, after n steps, the states $\{\mu_1, \ldots, \mu_n\}$ from the globally correlated mechanism. Hence, $P_1(\mu) = p_\mu$.

Equation (A1) can be seen as an addition over the ensemble realizations, where each term gives the weight of all realizations with n selection events. Taking into account that the variables $\mu_1, \ldots, \mu_{n-1}$ run over the domain of possible states $1,2, \ldots, L$, Eq. (A1) can also be written as

$$P_\mu(s) = p_\mu \Phi_\mu(s) + \sum_{n=1}^{\infty} \sum_{\{n_\nu\}} P_n(n_1, \ldots, n_L) \times \mathcal{T}_n(\{n_\nu\}|\mu)\, w_1^{n_1}(s) \cdots w_L^{n_L}(s)\Phi_\mu(s). \quad \text{(A2)}$$

Here, $P_n(n_1,\ldots,n_L)$ is the joint probability of getting n_ν times the state ν after n random steps, $\nu=1,\ldots,L$. Therefore, the sum $\sum_{\{n_\nu\}}$ is restricted to the condition $\sum_{\nu=1}^{L} n_\nu = n$.

The expression (A2) is exact. Now, we perform a set of approximations for getting the stationary state $P_\mu^{\mathrm{st}} = \lim_{t\to\infty} P_\mu(t)$. In the long time regime, $t \gg \{\tau_\nu\}_{\nu=1}^{L}$, in the Laplace domain we can approximate [46] the waiting time distribution as $w_\nu(s) \simeq 1-\tau_\nu s$, where τ_ν is the average time defined by Eq. (1). Therefore, $\Phi_\mu(s)\simeq\tau_\mu$ and also $w_1^{n_1}(s)\cdots w_L^{n_L}(s) = \prod_{\nu=1}^{L} w_\nu^{n_\nu}(s) \simeq \exp[-s\sum_{\nu=1}^{L}\tau_\nu n_\nu]$, which in the time domain leads to a Dirac δ function, $\delta(t-\sum_{\nu=1}^{L}\tau_\nu n_\nu)$.

In the long time regime, n increases unbounded. For the studied models, the conditional probability can then be approximated as $\mathcal{T}_n(\{n_\nu\}|\mu) \simeq n_\mu/n \simeq f_\mu$ [Eq. (2)]. Consequently, Eq. (A2) leads to the approximation

$$P_\mu(t) \simeq \sum_{n=1}^{\infty}\sum_{\{n_\nu\}} P_n(n_1,\ldots,n_L)\tau_\mu \frac{n_\mu}{n}\delta\left(t-\sum_{\nu=1}^{L}\tau_\nu n_\nu\right). \tag{A3}$$

By writing the δ contribution as $\delta(t-\sum_{\nu=1}^{L}\tau_\nu n_\nu) = \delta(t - n\sum_{\nu=1}^{L}\tau_\nu f_\nu)$, we realize that in the sum over n the dominant term is that with $n \simeq t/\sum_{\nu=1}^{L}\tau_\nu f_\nu$. Using the properties of the δ distribution, $\delta(t-n\sum_{\nu=1}^{L}\tau_\nu f_\nu) = (1/\sum_{\nu=1}^{L}\tau_{\nu'} f_{\nu'})\delta(n - t/\sum_{\nu=1}^{L}\tau_\nu f_\nu)$, and after the change of variables $n_\nu \to f_\nu$, Eq. (A3) leads to the stationary state

$$P_\mu^{\mathrm{st}} = \int_\Lambda df_1\cdots df_{L-1}\frac{\tau_\mu f_\mu}{\sum_{\nu=1}^{L}\tau_\nu f_\nu}\mathcal{P}(\{f_\nu\}), \tag{A4}$$

which in fact recovers Eq. (4). This result was also checked by numerical calculations for the memory models introduced in Sec. III.

APPENDIX B: ERGODICITY OF THE ELEPHANT RANDOM WALK

The elephant random walk is defined by the transition probability (15),

$$\mathcal{T}_n(\{n_\nu\}|\mu) = \varepsilon q_\mu + (1-\varepsilon)\frac{n_\mu}{n}. \tag{B1}$$

Here, we demonstrate that the fractions defined in Eq. (2), $f_\mu = \lim_{n\to\infty}(n_\mu/n)$, converge to q_μ; that is, the distribution of the fractions is given by Eq. (10) with $\langle f_\mu\rangle = q_\mu$ $(0<\varepsilon\leqslant 1)$.

At a given stage, the numbers n_μ can be split as follows:

$$n_\mu = m_\mu^{(1)} + M_\mu^{(1)}. \tag{B2}$$

Here, $m_\mu^{(1)}$ gives the number of times that, with probability ε, the state μ was chosen with probabilities $\{q_\mu\}_{\mu=1}^{L}$. Complementarily, $M_\mu^{(1)}$ gives the number of times that, with probability $1-\varepsilon$, the state μ was chosen with probabilities $\{n_\mu/n\}$. In the limit of a diverging number of selections (steps), the law of large numbers gives $\lim_{n\to\infty} m_\mu^{(1)}/n = \varepsilon q_\mu$. Thus, asymptotically we can approximate

$$\mathcal{T}_n(\{n_\nu\}|\mu) \simeq \varepsilon q_\mu + (1-\varepsilon)\left[\varepsilon q_\mu + \frac{M_\mu^{(1)}}{n}\right]. \tag{B3}$$

Now, we can split $M_\mu^{(1)}$ in the same way as follows:

$$M_\mu^{(1)} = m_\mu^{(2)} + M_\mu^{(2)}. \tag{B4}$$

Here, $m_\mu^{(2)}$ is the number of times that, with probability $(1-\varepsilon)\varepsilon$, the state μ was chosen with probabilities $\{q_\mu\}_{\mu=1}^{L}$. Similarly, $M_\mu^{(2)}$ gives the number of times that, with probability $(1-\varepsilon)(1-\varepsilon)$, the state μ was chosen with probabilities $\{M_\mu^{(1)}/n\}$. By using that $\lim_{n\to\infty} m_\mu^{(2)}/n = (1-\varepsilon)\varepsilon q_\mu$, it follows the approximation

$$\mathcal{T}_n(\{n_\nu\}|\mu) \simeq \varepsilon q_\mu + (1-\varepsilon)\left[\varepsilon q_\mu + \varepsilon q_\mu(1-\varepsilon) + \frac{M_\mu^{(2)}}{n}\right]. \tag{B5}$$

Performing the same splitting, at an arbitrary order we can write

$$M_\mu^{(k-1)} = m_\mu^{(k)} + M_\mu^{(k)}, \tag{B6}$$

where the law of large numbers gives $\lim_{n\to\infty} m_\mu^{(k)}/n = (1-\varepsilon)^{k-1}\varepsilon q_\mu$. Therefore, we get

$$\mathcal{T}_n(\{n_\nu\}|\mu) \simeq \varepsilon q_\mu + (1-\varepsilon)\varepsilon q_\mu \sum_{k=0}^{\infty}(1-\varepsilon)^k = q_\mu. \tag{B7}$$

This argument shows that in the asymptotic limit the memory on the previous states is lost. Hence, the finite random walk becomes ergodic, Eq. (10) with $\langle f_\mu\rangle = q_\mu$. Numerical simulations confirm this result. Notice that the previous argument does not apply to the urn models in Eqs. (16) and (21). On the other hand, we checked that for $\varepsilon\to 0$ the rate of convergence to the regime defined by Eq. (B7) is smaller, being infinite for $\varepsilon = 0$, that is, in the localized regime. We remark that this result does not contradict previous results for unbounded diffusion processes [37–39].

APPENDIX C: FRACTION PROBABILITY DENSITY OF THE URNLIKE DYNAMICS

For the urn dynamics defined by Eq. (16), here we obtain the probability density of the stationary fractions, Eq. (2).

By using Bayes rule, the joint probability $P_n(\mu_1,\ldots,\mu_n)$ of obtaining the values $\mu_1,\ldots,\mu_n$ with the dynamics defined by Eq. (16) can be written as

$$P_n(\mu_1,\ldots,\mu_n) = P_1(\mu_1)\mathcal{T}_1(\{n_{\nu_1}\}|\mu_2)\cdots\mathcal{T}_{n-1}(\{n_{\nu_{n-1}}\}|\mu_n).$$

By writing this expression in an explicit way, we realize that the joint probability $P_n(n_1,\ldots,n_L)$ of getting n_μ times the state μ after n random steps can be written as

$$P_n(n_1,\ldots,n_L) = \sum_{\nu=1}^{L}\frac{(n-1)!}{n_1!\cdots(n_\nu-1)!\cdots n_L!} \times p_\nu\frac{\Gamma(\lambda)}{\Gamma(n+\lambda)}\frac{1}{q_\nu}\prod_{\mu=1}^{L}\frac{\Gamma(n_\mu+\lambda_\mu)}{\Gamma(\lambda_\mu)}, \tag{C1}$$

where $\lambda_\mu = \lambda q_\mu$. Each term in the sum $\sum_{\nu=1}^{L}$ corresponds to all realizations with the same initial condition, which leads to the weight p_ν. The contributions proportional to the Γ functions follows straightforwardly from the product of successive conditional probabilities $\mathcal{T}_k(\{n_k\}|\mu_{k+1})$ and the property $\Gamma(n+x)/\Gamma(x)=x(1+x)(2+x)\cdots(n-1+x)$. Furthermore, in the first line the multinomial factor takes into account all realizations with the same numbers $\{n_\mu\}_{\mu=1}^L$. Equation (C1) can be rewritten as

$$P_n(n_1,\ldots,n_L)=\sum_{\nu=1}^{L}\frac{p_\nu}{q_\nu}\frac{n_\nu}{n}D_n(n_1,\ldots,n_L), \tag{C2}$$

where

$$D_n(n_1,\ldots,n_L)\equiv\frac{n!}{n_1!\cdots n_L!}\frac{\Gamma(\lambda)}{\Gamma(n+\lambda)}\prod_{\mu=1}^{L}\frac{\Gamma(n_\mu+\lambda_\mu)}{\Gamma(\lambda_\mu)}. \tag{C3}$$

In the limit $x\to\infty$ it is valid the Stirling approximation $\Gamma(x)\approx\sqrt{2\pi/x}e^{-x}x^x$. Hence, in the same limit, it follows that $\Gamma(x+\alpha)/\Gamma(x)\approx x^\alpha$. Using that $n!=\Gamma(n+1)$, and applying the previous approximations to Eq. (C3), in the limit $n\to\infty$ it follows that

$$D_n(n_1,\ldots,n_L)\approx\frac{\Gamma(\lambda)}{n^{\lambda-1}}\prod_{\mu=1}^{L}\frac{n_\mu^{\lambda_\mu-1}}{\Gamma(\lambda_\mu)}. \tag{C4}$$

By performing the change of variables $n_\mu\to nf_\mu$, and by using that (due to normalization) there are $L-1$ independent variables f_μ, the previous expression straightforwardly leads to the Dirichlet distribution $D(\{f_\mu\}|\{\lambda_\mu\})$, Eq. (18). Therefore, in the same limit, Eq. (C2) trivially recovers Eq. (17).

APPENDIX D: EXACT ANALYTICAL RESULTS FOR TWO-LEVEL SYSTEMS

For two-level systems driven by the imperfect urn dynamics, the integral expressions for the probability density of residence times [Eq. (35)] and the average residence time [Eq. (38)] can be explicitly evaluated. $W_\mu^\varepsilon(\tau)$ reads

$$\begin{aligned}W_\mu^\varepsilon(\tau)&=\gamma_\mu\exp[-\gamma_\mu\tau(1-\varepsilon q_\mu)]\\&\quad\times\{a_\mu(\tau)\,{}_1F_1[\lambda q_\mu;\lambda;(1-\varepsilon)\gamma_\mu\tau]\\&\quad+b_\mu(\tau)\,{}_1F_1[\lambda q_\mu;\lambda+1;(1-\varepsilon)\gamma_\mu\tau]\}.\end{aligned} \tag{D1}$$

The Kummer confluent hypergeometric function is ${}_1F_1[a;b;z]=\sum_{k=0}^{\infty}(a)_k(b)_k z^k/k!$ with $(x)_k=\prod_{j=0}^{k-1}(x+j)=\Gamma(x+k)/\Gamma(x)$. The auxiliary function $a_\mu(\tau)$ is

$$a_\mu(\tau)\equiv\frac{p_\mu}{q_\mu}(1-q_\mu)\varepsilon+\frac{(p_\mu-q_\mu)\lambda}{q_\mu\gamma_\mu\tau}, \tag{D2}$$

while $b_\mu(\tau)$ is

$$b_\mu(\tau)\equiv\left(1-\frac{p_\mu}{q_\mu}\varepsilon\right)-(1-\varepsilon)p_\mu+(p_\mu-q_\mu)\left(\varepsilon-\frac{\lambda}{q_\mu\gamma_\mu\tau}\right). \tag{D3}$$

Similarly, the average residence time is

$$\begin{aligned}\gamma_\mu T_\mu^\varepsilon&=a_\mu\,{}_2F_1\left[1;\lambda q_\mu;\lambda;\frac{1-\varepsilon}{1-\varepsilon q_\mu}\right]\\&\quad+b_\mu\,{}_2F_1\left[1;1+\lambda q_\mu;1+\lambda;\frac{1-\varepsilon}{1-\varepsilon q_\mu}\right].\end{aligned} \tag{D4}$$

Here, the hypergeometric function is defined by ${}_2F_1[a;b;c;z]=\sum_{k=0}^{\infty}(a)_k(b)_k(c)_k z^k/k!$, while the coefficients are

$$a_\mu\equiv\frac{1-p_\mu}{(1-q_\mu)(1-\varepsilon q_\mu)},\quad b_\mu\equiv\frac{p_\mu-q_\mu}{(1-q_\mu)(1-\varepsilon q_\mu)}. \tag{D5}$$

[1] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Perseus Books, Massachusetts, 1992).

[2] J. P. Bouchaud, Weak ergodicity breaking and aging in disordered systems, J. Phys. I France **2**, 1705 (1992).

[3] E. Lutz, Power-Law Tail Distributions and Nonergodicity, Phys. Rev. Lett. **93**, 190602 (2004).

[4] A. Rebenshtok and E. Barkai, Weakly non-ergodic statistical physics, J. Stat. Phys. **133**, 565 (2008); Distribution of Time-Averaged Observables for Weak Ergodicity Breaking, Phys. Rev. Lett. **99**, 210601 (2007).

[5] G. Margolin and E. Barkai, Nonergodisity of a time series obeying Lévy statistics, J. Stat. Phys. **122**, 137 (2006).

[6] G. Bel and E. Barkai, Weak Ergodicity Breaking in the Continuous-Time Random Walk, Phys. Rev. Lett. **94**, 240602 (2005); A random walk to a non-ergodic equilibrium concept, Phys. Rev. E **73**, 016125 (2006); J. H. P. Schulz and E. Barkai, Fluctuations around equilibrium laws in ergodic continuous-time random walks, *ibid.* **91**, 062129 (2015).

[7] A. Saa and R. Venegeroles, Ergodic transitions in continuous-time random walks, Phys. Rev. E **82**, 031110 (2010).

[8] T. Albers and G. Radons, Subdiffusive continuous time random walks and weak ergodicity breaking analyzed with the distribution of generalized diffusivities, Europhys. Lett. **102**, 40006 (2013).

[9] M. Dentz, A. Russian, and P. Gouze, Self-averaging and ergodicity of subdiffusion in quenched random media, Phys. Rev. E **93**, 010101(R) (2016).

[10] A. G. Cherstvy, A. V. Chechkin, and R. Metzler, Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes, New J. Phys. **15**, 083039 (2013); A. G. Cherstvy and R. Metzler, Non-ergodicity, fluctuations, and criticality in heterogeneous diffusion processes, Phys. Rev. E **90**, 012134 (2014).

[11] F. Kindermann, A. Dechant, M. Hohmann, T. Lausch, D. Mayer, F. Schmidt, E. Lutz, and A. Widera, Nonergodic diffusion of single atoms in a periodic potential, arXiv:1601.06663 (2016); M. Khoury, A. M. Lacasta, J. M. Sancho, and K. Lindenberg, Weak Disorder: Anomalous Transport and Diffusion Are Normal Yet Again, Phys. Rev. Lett. **106**, 090602 (2011).

[12] P. Massignan, C. Manzo, J. A. Torreno-Pina, M. F. García-Parajo, M. Lewestein, and G. J. Lapeyre, Jr., Nonergodic Subdiffusion from Brownian Motion in an Inhomogeneous Medium, Phys. Rev. Lett. **112**, 150603 (2014).

[13] O. Peters, Ergodicity Breaking in Geometric Brownian Motion, Phys. Rev. Lett. **110**, 100603 (2013).

[14] H. Safdari, A. G. Cherstvy, A. V. Chechkin, F. Thiel, I. M. Sokolov, and R. Metzler, Quantifying the non-ergodicity of scaled Brownian motion, J. Phys. A **48**, 375002 (2015); H. Safdari, A. V. Chechkin, G. R. Jafari, and R. Metzler, Aging scaled Brownian motion, Phys. Rev. E **91**, 042107 (2015).

[15] A. Godec, A. V. Chechkin, E. Barkai, H. Kantz, and R. Metzler, Localization and universal fluctuations in ultraslow diffusion processes, J. Phys. A **47**, 492002 (2014); A. S. Bodrova, A. G. Cherstvy, A. V. Chechkin, and R. Metzler, Ultraslow scaled Brownian motion, arXiv:1503.08125.

[16] A. Godec and R. Metzler, Finite-Time Effects and Ultraweak Ergodicity Breaking in Superdiffusive Dynamics, Phys. Rev. Lett. **110**, 020603 (2013).

[17] G. Bel and I. Nemenman, Ergodic and non-ergodic anomalous diffusion in coupled stochastic processes, New. J. Phys. **11**, 083009 (2009).

[18] Y. Meroz, I. M. Sokolov and J. Klafter, Subdiffusion of mixed origins: When ergodicity and nonergodicity coexist, Phys. Rev. E **81**, 010101(R) (2010); F. Thiel and I. M. Sokolov, Weak ergodicity breaking in an anomalous diffusion process of mixed origins, *ibid.* **89**, 012136 (2014).

[19] A. Fulinski, Anomalous diffusion and weak nonergodicity, Phys. Rev. E **83**, 061140 (2011).

[20] A. Dechant, E. Lutz, D. A. Kessler, and E. Barkai, Fluctuations of Time Averages for Langevin Dynamics in a Binding Force Field, Phys. Rev. Lett. **107**, 240603 (2011).

[21] G. Bel and E. Barkai, Ergodicity breaking in a deterministic dynamical system, Europhys. Lett. **74**, 15 (2006).

[22] T. Albers and Günter Radons, Weak Ergodicity Breaking and Aging of Chaotic Transport in Hamiltonian Systems, Phys. Rev. Lett. **113**, 184101 (2014).

[23] A. Figueiredo, T. M. Rocha Filho, M. A. Amato, Z. T. Oliveira, Jr., and R. Matsushita, Truncated Lévy flights and weak ergodicity breaking in the Hamiltonian mean-field model, Phys. Rev. E **89**, 022106 (2014).

[24] T. Akimoto, Distributional Response to Biases in Deterministic Superdiffusion, Phys. Rev. Lett. **108**, 164101 (2012).

[25] X. Brokmann, J.-P. Hermier, G. Messin, P. Desbiolles, J.-P. Bouchaud, and M. Dahan, Statistical Aging and Nonergodicity in the Fluorescence of Single Nanocrystals, Phys. Rev. Lett. **90**, 120601 (2003).

[26] G. Margolin and E. Barkai, Nonergodicity and Blinking Nanocrystals and Other Lévy-Walk Processes, Phys. Rev. Lett. **94**, 080601 (2005).

[27] C. Manzo, J. A. Torreno-Pina, P. Massignan, G. J. Lapeyre, Jr., M. Lewenstein, and M. F. G. Parajo, Weak Ergodicity Breaking of Receptor Motion in Living Cells Stemming from Random Diffusivity, Phys. Rev. X **5**, 011021 (2015).

[28] Y. He, S. Burov, R. Metzler, and E. Barkai, Random Time-Scale Invariant Diffusion and Transport Coefficients, Phys. Rev. Lett. **101**, 058101 (2008).

[29] A. Lubelski, I. M. Sokolov, and J. Klafter, Nonergodicity Mimics Inhomogeneity in Single Particle Tracking, Phys. Rev. Lett. **100**, 250602 (2008).

[30] S. Burov, J.-H. Jeon, R. Metzler, and E. Barkai, Single particle tracking in systems showing anomalous diffusion: The role of weak ergodicity breaking, Phys. Chem. Chem. Phys. **13**, 1800 (2011).

[31] J. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sorensen, L. Oddershede, and R. Metzler, *InVivo* Anomalous Diffusion and Weak Ergodicity Breaking of Lipid Granules, Phys. Rev. Lett. **106**, 048103 (2011).

[32] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking, Proc. Natl. Acad. Sci. USA **108**, 6438 (2011).

[33] B. J. West, E. L. Geneston, and P. Grigolini, Maximizing information exchange between complex networks, Phys. Rep. **468**, 1 (2008), and references therein.

[34] N. Piccinini, D. Lambert, B. J. West, M. Bologna, and P. Grigolini, Non-ergodic complexity management, Phys. Rev. E **93**, 062301 (2016); E. Geneston, R. Tuladhar, M. T. Beig, M. Bologna, and P. Grigolini, Ergodicity breaking and localization, arXiv:1601.02879.

[35] L. Silvestri, L. Fronzoni, P. Grigolini, and P. Allegrini, Event-Driven Power-Law Relaxation in Weak Turbulence, Phys. Rev. Lett. **102**, 014502 (2009).

[36] S. Bianco, M. Ignaccolo, M. S. Rider, M. J. Ross, P. Winsor, and P. Grigolini, Brain, music and non-Poisson renewal processes, Phys. Rev. E **75**, 061911 (2007).

[37] G. M. Schütz and S. Trimper, Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk, Phys. Rev. E **70**, 045101(R) (2004).

[38] H. Kim, Anomalous diffusion induced by enhancement of memory, Phys. Rev. E **90**, 012103 (2014).

[39] R. Kürsten, Random recursive trees and the elephant random walk, Phys. Rev. E **93**, 032111 (2016).

[40] J. C. Cressoni, M. A. A. da Silva, and G. M. Viswanathan, Amnestically Induced Persistence in Random Walks, Phys. Rev. Lett. **98**, 070603 (2007); A. S. Ferreira, J. C. Cressoni, G. M. Viswanathan, and M. A. Alves da Silva, Anomalous diffusion in non-Markovian walks having amnestically induced persistence, Phys. Rev. E **81**, 011125 (2010); J. C. Cressoni, G. M. Viswanathan, and M. A. A. da Silva, Exact solution of an anisotropic 2D random walk model with strong memory correlations, J. Phys. A **46**, 505002 (2013).

[41] V. M. Kenkre, Analytic formulation, exact solutions, and generalizations of the elephant and the Alzheimer random walks, arXiv:0708.0034.

[42] N. Kumar, U. Harbola, and K. Lindenberg, Memory-induced anomalous dynamics: Emergence of diffusion, subdiffusion, and superdiffusion from a single random walk model, Phys. Rev. E **82**, 021101 (2010).

[43] D. Boyer and J. C. Romo-Cruz, Solvable random-walk model with memory and its relations with Markovian models of anomalous diffusion, Phys. Rev. E **90**, 042136 (2014).

[44] F. N. C. Paraan and J. P. Esguerra, Exact moments in a continuous time random walk with complete memory of its history, Phys. Rev. E **74**, 032101 (2006).

[45] R. Hanel and S. Thurner, Generalized (c, d)-entropy and aging random walks, Entropy **15**, 5324 (2013).

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

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

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

[49] D. Blackwell and J. B. MacQueen, Fergurson distributions via Pólya urn schemes, Ann. Stat. **1**, 353 (1973).
[50] A. A. Budini, Central limit theorem for a class of globally correlated random variables, Phys. Rev. E **93**, 062114 (2016).
[51] M. Magdziarz, R. Metzler, W. Szczotka, and P. Zebrowski, Correlated continuous-time random walks in external force fields, Phys. Rev. E **85**, 051103 (2012).
[52] V. Tejedor and R. Metzler, Anomalous diffusion in correlated continuous time random walks, J. Phys. A **43**, 082002 (2010).