

Solvable random-walk model with memory and its relations with Markovian models of anomalous diffusion

D. Boyer^{1,2,*} and J. C. R. Romo-Cruz¹¹*Instituto de Física, Universidad Nacional Autónoma de México, D.F. 04510, México*²*Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, D.F. 04510, México*

(Received 22 May 2014; published 22 October 2014)

Motivated by studies on the recurrent properties of animal and human mobility, we introduce a path-dependent random-walk model with long-range memory for which not only the mean-square displacement (MSD) but also the propagator can be obtained exactly in the asymptotic limit. The model consists of a random walker on a lattice, which, at a constant rate, stochastically relocates at a site occupied at some earlier time. This time in the past is chosen randomly according to a memory kernel, whose temporal decay can be varied via an exponent parameter. In the weakly non-Markovian regime, memory reduces the diffusion coefficient from the bare value. When the mean backward jump in time diverges, the diffusion coefficient vanishes and a transition to an anomalous subdiffusive regime occurs. Paradoxically, at the transition, the process is an anticorrelated Lévy flight. Although in the subdiffusive regime the model exhibits some features of the continuous time random walk with infinite mean waiting time, it belongs to another universality class. If memory is very long-ranged, a second transition takes place to a regime characterized by a logarithmic growth of the MSD with time. In this case the process is asymptotically Gaussian and effectively described as a scaled Brownian motion with a diffusion coefficient decaying as $1/t$.

DOI: [10.1103/PhysRevE.90.042136](https://doi.org/10.1103/PhysRevE.90.042136)

PACS number(s): 05.40.Fb, 89.75.Fb, 82.39.Rt

I. INTRODUCTION

“True” self-interacting random walks (as opposed to static models like the self-avoiding walk of polymer physics) are an important class of non-Markovian kinetic processes with long-range memory [1]. Reinforced random walks are self-attracting processes where, typically, a random walker tends to preferentially revisit the nearest-neighbor sites that it has visited before (or, in a variant, recross the edges already crossed). In the past few decades, reinforced random walks have received attention not only for the mathematical challenges they raise but also for their applications to biology [2,3]. These walks have been used for the description of the displacements of ants or bacteria [2]. In ecology, they can also represent simple models of “site fidelity,” a behavior observed in many animals in the wild [4,5]. Many reinforced walk models are defined through transition probabilities that depend on the number of visits (or crossings) received by the sites (or edges) and the resulting dynamics is thus strongly path dependent.

Some rigorous results have been obtained for certain reinforcement rules, showing that different dynamical behaviors can emerge depending on the strength of memory and the spatial dimension. A single walker may asymptotically become localized (keeping oscillating between a few sites), or diffusive, in which the range of its position X_t is infinite and the origin visited infinitely often in one dimension (1D) [1,2,6]. Numerical simulations actually show that a variety of models seems to exhibit a phase transition at finite reinforcement between a localized and a diffusive regime [7–10]. Interestingly, in the diffusive regime, the same studies have presented evidence that diffusion is anomalous, namely subdiffusive. This type of motion, widely studied in

Markovian contexts [11,12], is characterized by a mean-square displacement (MSD) of the particle which does not follow the Smoluchowski-Einstein law of Brownian motion, but a slower one, of the form $\langle X_t^2 \rangle \propto t^\mu$ with μ an exponent < 1 .

Few random-walk models with long-range memory are analytically tractable [13–15]. For a single particle, path-dependent diffusion is notoriously difficult to formalize and many results on reinforced random walks are based on numerical simulations. Usually, such processes cannot be described by a master equation for the single-time occupation probability (or propagator). Instead, they require the introduction of multiple-time distribution functions that are related to each other via a hierarchy of relations. Path-integral approaches [16,17] or approximate scaling arguments [18,19] have been developed to overcome these difficulties. Nonlinear integrodifferential Fokker-Planck equations can also be derived within a mean-field approximation [10]. This latter approach can give a fairly good picture of the phase diagram but remains limited for a precise dynamical description.

The recent ecological literature reveals a regain of interest for reinforced random walks. Thanks to spectacular advances in tracking technology, the positions of individual animals [20–22], including humans [19,23–25], can be recorded with a high resolution and during long periods of time. Recently, random-walk models with memory have successfully explained some features of the trajectories of humans [19], monkeys [26], or bison [27], which all exhibit strong recurrence, an anomalous slow diffusion, and a heterogeneous occupation of space. Other theoretical studies have identified memory as a key factor for the emergence of home ranges, that is, a restricted space use [28–32]. Reinforced random walks thus offer a promising alternative to Markovian random walks, which remain the dominant modeling paradigm in movement ecology [33–35]. However, most models developed so far are computational and we still lack a basic understanding of the

*boyer@fisica.unam.mx

effects of memory on mobility patterns, even in simple cases. For this purpose, it is desirable, in parallel with ecologically realistic computational approaches, to gain knowledge from the mathematical analysis of very simple models.

Here, we exactly solve in the asymptotic time limit a reinforced model that generalizes the one presented in Ref. [26], where a random walker intermittently relocates to sites visited in the past. We obtain not only the behavior of the MSD, but also the properties of the propagator or diffusion front, a quantity which is almost unknown for single reinforced walks. In contrast with usual reinforced models, here the walker's steps are not necessarily directed to nearest-neighbor sites. This property has two advantages: It agrees with the empirical fact that many animals and humans actually perform long, intermittent commuting bouts to visit places that are beyond their perception range [36–39], and it also greatly simplifies the mathematical analysis since an exact master equation can be written in this case.

Depending on a memory parameter, we find that motion can be Brownian or subdiffusive. Two subdiffusive regimes are identified, with power-law and logarithmic dynamics, respectively. Somewhat paradoxically, a Lévy-like distribution for the length of the steps emerges at the onset of subdiffusion. We thus suggest that a mechanism based on memory could be at the origin of the Lévy flight patterns observed in many animals [21,23,35,40–42]. In addition, the availability of the propagator allows us to discuss in a formal way some analogies and differences between the present model and well-known, essentially Markovian, models of anomalous diffusion: notably, the continuous time random walk (CTRW) of Montroll and Weiss [43] and the scaled Brownian motion [44].

II. MODEL AND BASIC QUANTITIES

Let us consider a one-dimensional lattice with unit spacing and a walker with position X_t at time t . Time is discrete ($t = 0, 1, 2, \dots$) and the walker starts at the origin, $X_0 = 0$. Let q be a constant parameter, $0 < q < 1$. At each time step, $t \rightarrow t + 1$, the walker chooses one of the two movement modes: (i) With probability $1 - q$, the walker performs a random step to a nearest-neighbor site, like in the simple symmetric random walk, and (ii) with the complementary probability q , the walker chooses a random integer t' in the interval $[0, t]$ according to a probability distribution $p_t(t')$, which is given *a priori*. Then the walker directly relocates at the site that it occupied at time t' , i.e., $X_{t+1} = X_{t'}$.

These rules are depicted in Fig. 1. If, for instance, $p_t(t')$ is uniform, i.e., $p_t(t') = \frac{1}{t+1}$, one recovers the preferential visit model studied in Ref. [26] (see also Refs. [4,5]). In this case, the memory rule (ii) is equivalent to revisiting an already-visited site, say, n , with probability proportional to the total amount of time spent by the walker at this site since $t = 0$. Hence, the more visits site n receives, the more likely it will be chosen for future visits in the memory movement mode. This is the principle of reinforced random walks. In a different context, similar rules characterize network growth models with preferential attachment [45–47]. An important difference with usual reinforced random walks is that, here, any of the previously visited sites is susceptible to receive the

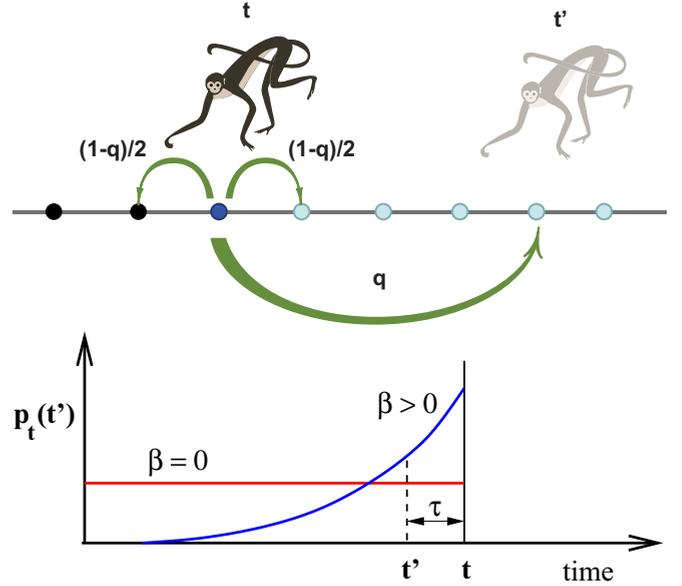


FIG. 1. (Color online) Top: At time t , the model walker performs a nearest-neighbor random-walk step with probability $1 - q$ or relocates, with probability q , to the site it occupied at some earlier time t' . Bottom: Probability distribution of t' for the uniform preferential visit case [$\beta = 0$ in Eq. (4)] or in a case with memory decay ($\beta > 0$).

next visit and not just the nearest neighbors of the current position of the walker. This property considerably simplifies the analysis, which still is not trivial. Note that a Markovian limiting case of this model is the random walk with stochastic resetting to the origin [48], which corresponds to $p_t(t') = \delta_{t',0}$.

We focus on the propagator $P(n, t)$ of the single particle, which is the probability that $X_t = n$ given that $X_0 = 0$. As formally shown in Appendix A, despite the fact that our process is highly non-Markovian, the evolution of $P(n, t)$ is exactly described through a single master equation as follows:

$$P(n, t + 1) = \frac{1 - q}{2} P(n - 1, t) + \frac{1 - q}{2} P(n + 1, t) + q \sum_{t'=0}^t p_t(t') P(n, t'). \quad (1)$$

The last term is the probability to choose site n using the memory mode. This term can also be interpreted as proportional to the weighted number of previous visits received by site n , where the weight of a visit received at time t' is $p_t(t')$. In this study, we will restrict ourselves to the cases where the probability distribution of t' is of the form $p_t(t') \propto F(t - t')$. Hence, the memory kernel $F(\tau)$ depends on the time elapsed between the remembered event and the present time, $\tau \equiv t - t'$. The normalization condition $\sum_{t'=0}^t p_t(t') = 1$ at each t imposes that

$$p_t(t') = \frac{F(t - t')}{C(t)}, \quad (2)$$

where

$$C(t) = \sum_{t'=0}^t F(t - t'). \quad (3)$$

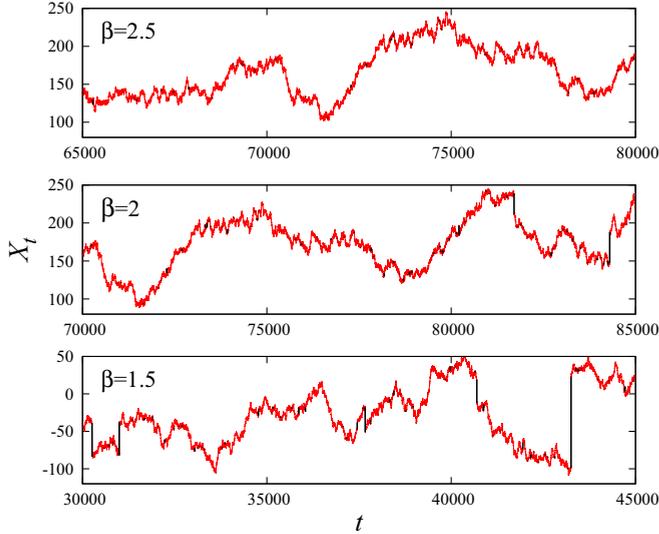


FIG. 2. (Color online) Trajectories obtained from numerical simulations of the model in the normally diffusive regime (top), at the subdiffusive threshold (middle) and in the first subdiffusive regime (bottom). The thick (black) vertical lines indicate the relocation steps in the memory mode. The rate of memory use is $q = 0.01$ in these cases.

In the uniform preferential visit model, $dF(\tau)/d\tau = 0$, i.e., memory does not decay. Here, we are interested in cases where $dF(\tau)/d\tau < 0$, i.e., where recent visits count more than visits performed further in the past, as represented in Fig. 1 (bottom). We will consider a particularly interesting case, namely scale-free memory decays,

$$F(\tau) = (\tau + 1)^{-\beta}, \quad 0 \leq \tau \leq t, \quad (4)$$

with β an exponent. If β is large, the early time trajectory ($t' \ll t$, or τ large) tends to be completely forgotten. Hence, when the walker uses its memory, it will most likely decide to revisit positions occupied at times proximate to the present time t . But if β is small enough, memory becomes long range and may drastically affect the diffusion process, like in the preferential visit model ($\beta = 0$) [26].

At a given time t , it is natural to define the mean backward jump in time as $\langle \tau \rangle_t \equiv \langle t - t' \rangle$, or

$$\langle \tau \rangle_t = \frac{\sum_{\tau=0}^t \tau F(\tau)}{\sum_{\tau=0}^t F(\tau)}. \quad (5)$$

If $\beta > 2$ in the kernel (4), both sums in Eq. (5) converge and $\langle \tau \rangle_t$ tends to a constant at large t ,

$$\langle \tau \rangle_\infty = \frac{\sum_{\tau=0}^{\infty} \tau (1 + \tau)^{-\beta}}{\zeta(\beta)}, \quad (6)$$

where $\zeta(\beta) = \sum_{\tau=0}^{\infty} (1 + \tau)^{-\beta}$ is the Riemann ζ function. In this case, memory has a finite range and should not affect in an essential way the normal diffusion process. Figure 2 displays examples of walks generated numerically in 1D in three different cases: $\beta > 2$, $\beta = 2$, and $\beta < 2$.

III. SUMMARY OF THE MAIN RESULTS

The analytical results presented below are derived in Sec. IV using the following methodology. Instead of solving Eq. (1), we have considered the moments of $P(n, t)$, defined by the ensemble averages $\langle X_t^{2p} \rangle$. The moments obey simpler equations and carry information on the behavior of the distribution itself. These equations can be solved in the asymptotic time limit. We first study the second moment ($p = 1$), i.e., the mean-square displacement (MSD). For $\beta > 1$, we assume that the leading asymptotic term of the MSD is of the form $K t^\mu$ and calculate the constants K and μ . For the case $\beta < 1$, the correct ansatz is $\langle X_t^2 \rangle \simeq K \ln t$ and we determine K . We next assume that $P(n, t)$ obeys a scaling form at large time in each case, see Eq. (11) below, which enables us to obtain the asymptotic behavior of all the higher-order moments for any positive integer $p \geq 1$. The MSD ansatz and the scaling assumption yield consistent results, which are also checked numerically by simulations or exact time integration of the equations.

A. Mean-square displacement $\langle X_t^2 \rangle$

We summarize the exact asymptotic results obtained for the MSD, which is the second moment of $P(n, t)$,

$$M_2(t) \equiv \sum_{n=-\infty}^{\infty} n^2 P(n, t) = \langle X_t^2 \rangle. \quad (7)$$

If $\beta > 2$, we have the following:

$$M_2(t) \simeq \left(\frac{1 - q}{1 + q \langle \tau \rangle_\infty} \right) t. \quad (8)$$

This result actually applies to any kernel $F(\tau)$ with finite first moment. Thus, diffusion is normal and the main effect of memory is to decrease the diffusion constant compared to that of the memoryless random walk ($q = 0$). The larger the q (frequent memory use) and the larger the $\langle \tau \rangle_\infty$ (better memory), the slower the diffusion.

If $1 < \beta < 2$, Eq. (8) above no longer holds since $\langle \tau \rangle_\infty = \infty$, making the diffusion constant vanish. Instead, motion is asymptotically subdiffusive as follows:

$$M_2(t) \simeq K t^{\beta-1} \quad \text{with} \quad K = \frac{(1 - q)\zeta(\beta)}{q \int_0^1 du \frac{1 - u^{\beta-1}}{(1 - u)^\beta}}. \quad (9)$$

If $\beta < 1$, diffusion is even slower, logarithmic in time,

$$M_2(t) \simeq K \ln t \quad \text{with} \quad K = \frac{1 - q}{q(1 - \beta) \int_0^1 du \frac{-\ln u}{(1 - u)^\beta}}. \quad (10)$$

In particular, by setting $\beta = 0$ in (10) one recovers the result of the uniform preferential visit model [26], $M_2(t) \simeq \frac{1 - q}{q} \ln t$.

B. Scaling functions

In each case above, we make a scaling hypothesis for the probability density in the long-time limit as follows:

$$P(n, t) \simeq \frac{1}{\sqrt{M_2(t)}} g\left(\frac{n}{\sqrt{M_2(t)}}\right), \quad (11)$$

where, by construction, $g(x)$ is a normalized scaling function [$\int_{-\infty}^{\infty} g(x) dx = 1$] of unit variance [$\int_{-\infty}^{\infty} x^2 g(x) dx = 1$].

Although $g(x)$ may not be always obtained explicitly, its properties can be inferred from the asymptotic behavior of the even moments, defined as

$$M_{2p}(t) = \sum_{n=-\infty}^{\infty} n^{2p} P(n,t), \quad (12)$$

with p being a positive integer [$M_{2p+1}(t) = 0$ by symmetry]. Assuming that Eq. (11) holds, it is easy to see that in the long-time limit the moments are given by

$$M_{2p}(t) \simeq a_p [M_2(t)]^p, \quad (13)$$

where a_p is a constant given by $a_p = \int_{-\infty}^{\infty} dx x^{2p} g(x)$. Obviously, $a_0 = a_1 = 1$. The constants a_p are calculated in Sec. IV B, and we find the following.

For $\beta > 2$, the scaling function $g(x)$ is Gaussian, i.e.,

$$\frac{a_p}{a_{p-1}} = 2p - 1, \quad (14)$$

for any positive integer p .

For $\beta < 1$ (logarithmic diffusion case), the process is also Gaussian asymptotically, namely $a_p/a_{p-1} = 2p - 1$ as above, despite the fact that memory is very long ranged and diffusion strongly anomalous. This result is nontrivial: The mechanism that makes $g(x)$ Gaussian is strongly driven by memory and differs from the one that leads to Gaussianity in the central limit theorem (previous case).

In the intermediate case $1 < \beta < 2$ (anomalous diffusion as $t^{\beta-1}$), $g(x)$ deviates from the Gaussian form. Its moments satisfy

$$\frac{a_p}{a_{p-1}} = (2p - 1) \frac{p \mathcal{I}_1(\beta)}{\mathcal{I}_p(\beta)}, \quad (15)$$

where

$$\mathcal{I}_p(\beta) = \int_0^1 du \frac{1 - u^{p(\beta-1)}}{(1 - u)^\beta}. \quad (16)$$

Note that (14), (15), and (16) are *universal* relations, i.e., the corresponding limiting distributions $P(n,t)$ do not depend on q nor on the details of the memory kernel but only on the fact that $F(\tau) \sim \tau^{-\beta}$ at large τ .

Alternatively, expressions (15) and (16) can be written using special functions. In this paper we define

$$\mu \equiv \beta - 1. \quad (17)$$

By integrating (16) by parts, we have $\mathcal{I}_p(\beta) = -\frac{1}{\mu} + pB[p\mu, 1 - \mu]$, with $0 < \mu < 1$ and where $B(x, y)$ is the Beta function. Using properties of the Γ function [49], Eqs. (15) and (16) can be rewritten as follows:

$$\frac{a_p}{a_{p-1}} = (2p - 1) \frac{p \left[\frac{\pi \mu}{\sin(\pi \mu)} - 1 \right]}{\Gamma(1 - \mu) \frac{\Gamma(p\mu + 1)}{\Gamma(p\mu + 1 - \mu)} - 1}. \quad (18)$$

In the two limiting cases $\mu \rightarrow 1^-$ (or $\beta \rightarrow 2^-$) and $\mu \rightarrow 0^+$ (or $\beta \rightarrow 1^+$), Eq. (18) gives back $a_p/a_{p-1} \rightarrow 2p - 1$, namely the two Gaussian cases previously mentioned.

IV. DERIVATION OF THE RESULTS

A. Second moment

By taking the second moment of Eq. (1), we obtain a recursive relation for the mean-square displacement as follows:

$$M_2(t+1) = 1 - q + (1 - q)M_2(t) + \frac{q}{C(t)} \sum_{t'=0}^t F(t-t')M_2(t'). \quad (19)$$

This equation may not be exactly solvable for all t , but we investigate its asymptotic behavior. To show (8)–(10), we write Eq. (19) in the following form:

$$M_2(t+1) - M_2(t) - (1 - q) + qM_2(t) = \frac{q}{C(t)} \mathcal{F}\{M_2(t)\} \quad (20)$$

with $C(t) = \sum_{\tau=0}^t (1 + \tau)^{-\beta}$ and

$$\mathcal{F}\{M_2(t)\} \equiv \sum_{t'=0}^t \frac{M_2(t')}{(t-t'+1)^\beta}. \quad (21)$$

We notice that a diverging $M_2(t)$ in the infinite time limit implies that the terms with large t' dominate in the sum (21). Therefore, one may approximate $M_2(t')$ in the sum by its leading asymptotic form, even at small t' .

If $\beta > 2$, we make the asymptotic ansatz $M_2(t) \simeq Kt$, obtaining the following:

$$\mathcal{F}\{Kt\} = Kt \sum_{\tau=0}^t (1 + \tau)^{-\beta} - K \sum_{\tau=0}^t \tau (1 + \tau)^{-\beta}, \quad (22)$$

$$\simeq Kt \sum_{\tau=0}^{\infty} (1 + \tau)^{-\beta} - K \sum_{\tau=0}^{\infty} \tau (1 + \tau)^{-\beta}, \quad (23)$$

since the two sums above converge. Substituting (23) into (20) and noting that $M_2(t+1) - M_2(t) \simeq K$, one obtains an equation for K which is easily solved as $K = (1 - q)/[1 + q\langle \tau \rangle_\infty]$. This is result (8), which is displayed in Fig. 3 (top). This panel also shows the MSD obtained by iterating the exact Eq. (19) numerically from the initial condition $M_2(0) = 0$, as well as the MSD obtained from Monte Carlo simulations (both are identical within numerical errors). Figure 3 (bottom) actually shows that the time derivative of the exact MSD tends to the calculated diffusivity K . Note that higher-order corrections to the linear behavior become more important at finite t as $\beta \rightarrow 2$, and convergence is slower. At $\beta = 2$, $\langle \tau \rangle_\infty = \infty$ and the diffusivity K vanishes, suggesting a change of temporal behavior.

To examine the case $1 < \beta < 2$, we make the general ansatz $M_2(t) \simeq Kt^\nu$, where ν and K are unknown. Similarly to (23), we seek to expand the memory term as $\mathcal{F}\{Kt^\nu\} \simeq Kt^\nu(c_1 + c_2t^{-\beta+1})$ for large t . As shown in Appendix B,

$$\mathcal{F}\{Kt^\nu\} \simeq Kt^\nu \left\{ \zeta(\beta) - t^{-\beta+1} \left[\int_0^1 du \frac{1 - (1 - u)^\nu}{u^\beta} + \frac{1}{\beta - 1} \right] \right\}. \quad (24)$$

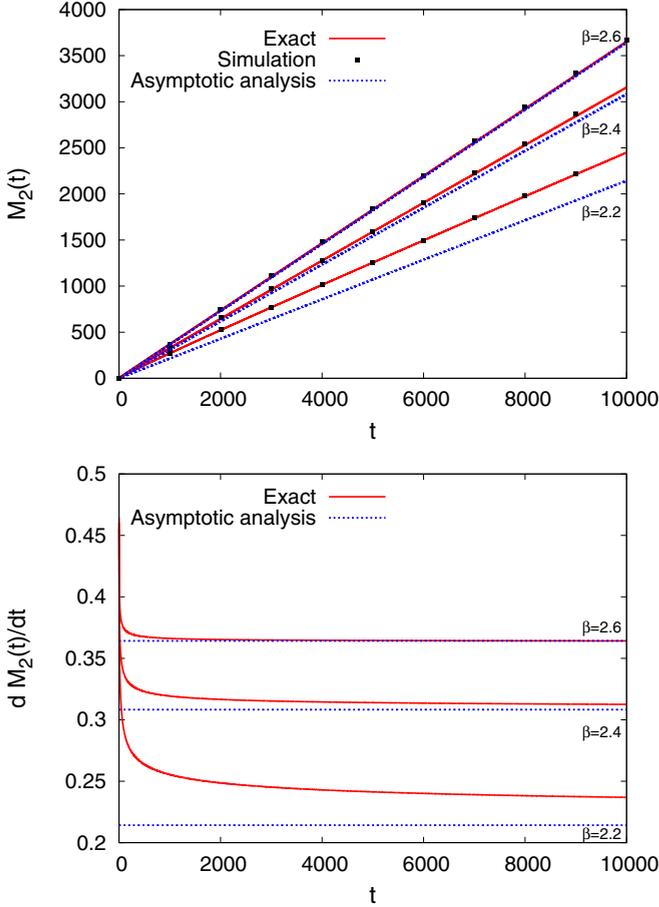


FIG. 3. (Color online) Comparison between the exact MSD obtained from solving Eq. (19) numerically (solid red lines), and the asymptotic analytical results [dotted blue lines, from Eq. (8)], in the normal diffusive case $\beta > 2$. Top: $M_2(t)$ vs t . The black dots correspond to Monte Carlo simulations. Bottom: The same curves, represented as $dM_2(t)/dt$ vs t . Higher-order corrections, which are not calculated in the theory, become more important at finite t when β approaches the subdiffusive transition point $\beta = 2$. In all cases, $q = 0.5$.

The normalization constant $C(t)$ can be written as follows:

$$C(t) = \sum_{\tau=0}^t \frac{1}{(1+\tau)^\beta} \simeq \zeta(\beta) - \frac{t^{-\beta+1}}{\beta-1}. \quad (25)$$

Assuming that $\nu < 1$ (subdiffusive behavior), so we can neglect $M_2(t+1) - M_2(t) \simeq \dot{M}_2(t)$ in (20) compared to the constant $-(1-q)$, and substituting (24) and (25) in (20), we obtain the following:

$$-(1-q) = qKt^{\nu-\beta+1} \frac{1}{\zeta(\beta)} \int_0^1 du \frac{u^\nu - 1}{(1-u)^\beta} + O(t^{-\beta+1}). \quad (26)$$

The exponent of t in the right-hand side of (26) must be 0 for consistency, which gives $\nu = \beta - 1 \equiv \mu$. Consequently, this equation can be solved for the constant K and formula (9) is obtained. Figure 4 (top) shows that the exact numerical MSD obtained from iterating (19) tends to a linear behavior with

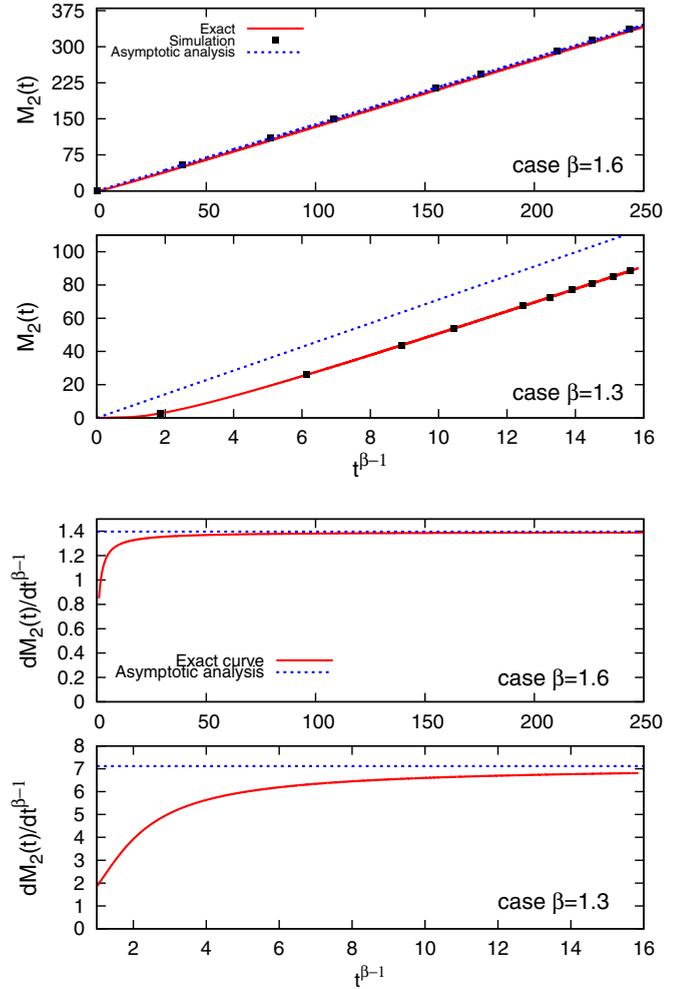


FIG. 4. (Color online) Top panels: $M_2(t)$ as a function of $t^{\beta-1}$ for two values of β in the range (1,2) and with $q = 0.5$. The solid red lines are obtained from exact resolution of Eq. (19) and the dotted blue lines are given by Eq. (9). The black dots correspond to Monte Carlo simulations of the model. Bottom panels: The same curves, represented as $dM_2(t)/dt^{\beta-1}$ vs $t^{\beta-1}$.

respect to the variable $t^{\beta-1}$, as predicted by our result. Note that at large but finite time the exact MSD would be better approximated by the leading asymptotic result plus a constant, which, unfortunately, cannot be calculated by our method (this situation occurs for $\beta < 1$ as well, see Fig. 5 below). However, such a constant becomes negligible as $t \rightarrow \infty$. For a more precise comparison, we have displayed in Fig. 4 (bottom) the derivative of the MSD with respect to $t^{\beta-1}$; this quantity tends to a constant in very good agreement with the calculated K .

In the case $\beta < 1$, we follow a similar route, now with a logarithmic ansatz $M_2(t) \simeq K \ln t$, as suggested by the exact result for the preferential visit model, i.e., $\beta = 0$ [26]. Since $\sum_{\tau=0}^t (1+\tau)^{-\beta}$ does not converge, one can use directly the Euler-Maclaurin expansion as follows:

$$C(t) \simeq \int_0^t d\tau (1+\tau)^{-\beta} \simeq \frac{t^{1-\beta}}{1-\beta}, \quad (27)$$

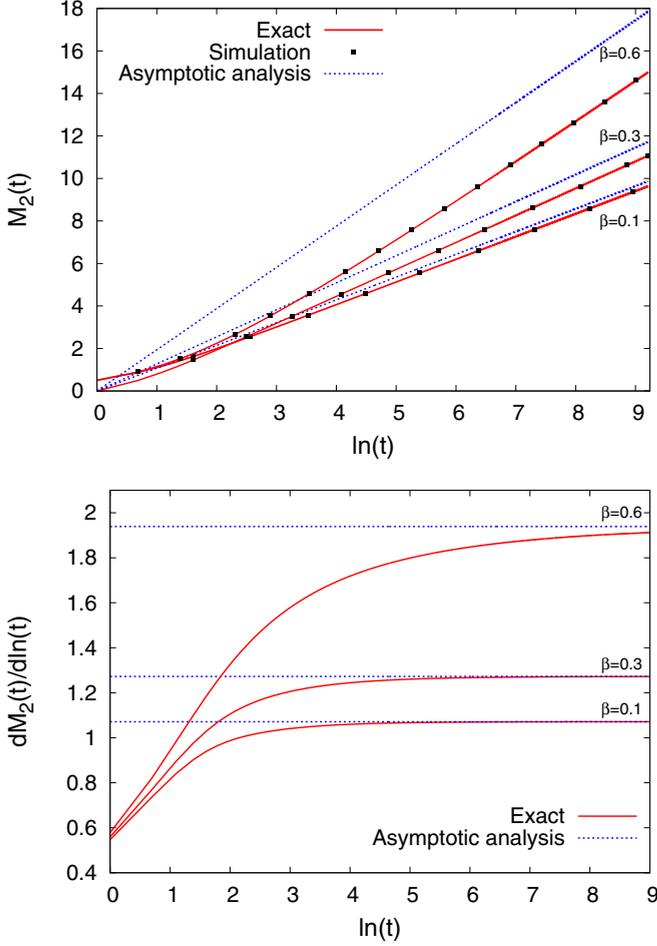


FIG. 5. (Color online) Top: $M_2(t)$ as a function of $\ln t$ for three values of $\beta < 1$ and with $q = 0.5$. The solid red lines are obtained from exact resolution of Eq. (19) and the dotted blue lines are given by Eq. (10). The black dots correspond to Monte Carlo simulations. Bottom: The same curves, represented as $dM_2(t)/d(\ln t)$ vs $\ln t$.

at large t . Similarly,

$$\mathcal{F}\{M_2(t)\} \simeq \int_0^t dt' \frac{M_2(t')}{(t-t'+1)^\beta} \simeq K \int_A^t dt' \frac{\ln t'}{(t-t'+1)^\beta}, \quad (28)$$

where A is some constant. Making the change $t' = ut$ and noticing that the functions $\ln u/(1-u)^\beta$ and $1/(1-u)^\beta$ are integrable both in 0 and 1 when $\beta < 1$, we obtain, for the case when t is large,

$$\mathcal{F}\{M_2(t)\} \simeq K t^{1-\beta} \int_0^1 du \frac{\ln u + \ln t}{(1-u)^\beta}. \quad (29)$$

By substituting (27) and (29) into (20) and neglecting $M_2(t+1) - M_2(t)$ as before, we obtain the following:

$$-(1-q) = qK(1-\beta) \int_0^1 du \frac{\ln u}{(1-u)^\beta}, \quad (30)$$

which gives the constant K and result (10). Figure 5 (top) shows that the exact MSD obtained numerically becomes a linear function of $\ln t$ at large times, in agreement with theory. We have also displayed in Fig. 5 (bottom) the derivative of the

MSD with respect to $\ln t$: This quantity tends to a constant in very good agreement with the calculated K .

B. Higher-order moments

In a similar way, we now consider the asymptotic behavior of the moments $M_{2p}(t)$. We start from the exact relation which follows from the master equation (1),

$$\begin{aligned} M_{2p}(t+1) &= 1 - q + (1-q)M_{2p}(t) \\ &+ (1-q) \sum_{k=1}^{p-1} C_{2p}^{2k} M_{2k}(t) \\ &+ \frac{q}{C(t)} \sum_{t'=0}^t F(t-t') M_{2p}(t'). \end{aligned} \quad (31)$$

Assuming the scaling hypothesis (13), and given that $M_2(t)$ always diverges as $t \rightarrow \infty$, we can neglect the term $1 - q$ as well as the terms proportional to $M_{2k}(t)$ for all $k < p - 1$ in the right-hand side of (31). Substituting $M_{2p}(t+1) - M_{2p}(t)$ by dM_{2p}/dt in the long-time limit, Eq. (31) becomes the following:

$$\begin{aligned} a_p \frac{dM_2(t)^p}{dt} &\simeq (1-q)p(2p-1)a_{p-1}M_2(t)^{p-1} \\ &+ \frac{qa_p}{C(t)} \sum_{t'=0}^t F(t-t') [M_2(t')^p - M_2(t)^p]. \end{aligned} \quad (32)$$

If $1 < \beta < 2$, we have $M_2(t) \simeq Kt^{\beta-1}$ and Eq. (32) reads as follows:

$$\begin{aligned} K^p a_p p(\beta-1)t^{(\beta-1)p-1} \\ \simeq (1-q)p(2p-1)a_{p-1}K^{p-1}t^{(\beta-1)(p-1)} \\ + \frac{qa_p K^p}{C(t)} \sum_{t'=0}^t \frac{t'^{(\beta-1)p} - t^{(\beta-1)p}}{(t-t'+1)^\beta}. \end{aligned} \quad (33)$$

Using formula (24), we obtain the following:

$$\begin{aligned} \sum_{t'=0}^t \frac{t'^{(\beta-1)p}}{(t-t'+1)^\beta} \\ \simeq t^{(\beta-1)p} \left\{ \zeta(\beta) - t^{-\beta+1} \left[\int_0^1 du \frac{1-u^{(\beta-1)p}}{(1-u)^\beta} + \frac{1}{\beta-1} \right] \right\}, \end{aligned} \quad (34)$$

whereas

$$\sum_{t'=0}^t \frac{t'^{(\beta-1)p}}{(t-t'+1)^\beta} \simeq t^{(\beta-1)p} \left[\zeta(\beta) - \frac{t^{-\beta+1}}{\beta-1} \right]. \quad (35)$$

Since $1 < \beta < 2$, then $t^{(\beta-1)p-1} \ll t^{(\beta-1)(p-1)}$ and the left-hand side of Eq. (33) can be neglected. Substituting (34) and (35) in (33), and then taking the limit $C(t) \rightarrow \zeta(\beta)$ at large t and using the expression (9) for K , Eq. (15) is obtained.

For the case $\beta < 1$, or $M_2(t) \simeq K \ln t$, the derivative in the left-hand side of (32) is $O(\ln^{p-1}(t)/t) \rightarrow 0$ and can still be neglected compared with terms growing with time in this

equation. Thus Eq. (32) becomes the following:

$$(1-q)p(2p-1)a_{p-1}K^{p-1}\ln^{p-1}t \\ = \frac{qa_pK^p}{C(t)} \int_0^1 dt' \frac{\ln^p t - \ln^p t'}{(t-t'+1)^\beta}, \quad (36)$$

where in this case the diverging sums have been replaced by integrals. We next write

$$\int_0^t dt' \frac{\ln^p t'}{(t-t'+1)^\beta} \\ \simeq t^{1-\beta} \int_0^1 du \frac{(\ln u + \ln t)^p}{(1-u)^\beta} \\ = t^{1-\beta} \sum_{k=0}^p C_p^k \ln^{p-k} t \int_0^1 du \frac{\ln^k u}{(1-u)^\beta} \\ \simeq t^{1-\beta} \left[\frac{\ln^p t}{1-\beta} + p \ln^{p-1} t \int_0^1 du \frac{\ln u}{(1-u)^\beta} \right], \quad (37)$$

where in the last step we have retained only the first two dominant terms of the binomial expansion, i.e., those of indices $k=0$ and $k=1$. We also have the following:

$$\int_0^t dt' \frac{\ln^p t'}{(t-t'+1)^\beta} \simeq t^{1-\beta} \frac{\ln^p t}{1-\beta}. \quad (38)$$

Substituting (37) and (38) into (36), and then using the expansion $C(t) \simeq t^{1-\beta}/(1-\beta)$ as well as the expression (10) for K , one obtains the following simple result:

$$a_p/a_{p-1} = 2p - 1, \quad (39)$$

which corresponds to the Gaussian distribution.

Note that the equation that leads to this Gaussian result [Eq. (36)] can be understood as a balance between diffusion due to random increments (the combinatorial left-hand side) and confining effects due to recurrent memory (the integral right-hand side). This balance arises because the term $a_p dM_2(t)^p/dt$ in Eq. (32) is negligible. In the simple random walk, on the contrary, $a_p dM_2(t)^p/dt$ is dominant, since $M_2(t) \propto t$, and equals the combinatorial term (in the absence of a memory term). Therefore, the mechanism leading to Gaussianity in the memory walk here differs substantially from the central limit theorem case.

C. Numerical tests of the higher-order moments

Figure 6 shows several analytical curves a_p/a_{p-1} as a function of β , summarizing the three regimes discussed above. To check these results, we performed Monte Carlo simulations of the model and computed the quantity

$$Q_p(t) \equiv \frac{M_{2p}(t)}{M_{2p-2}(t)M_2(t)}, \quad (40)$$

after obtaining the corresponding moments by averaging X_t^{2p} , X_t^{2p-2} , and X_t^2 over many independent walks. If the scaling hypothesis holds, $Q_p(t) \rightarrow a_p/a_{p-1}$. In each case, we checked that $Q_p(t)$ tended to a constant at large times, which is indicated by a dot in Fig. 6.

The agreement between theory and simulations is very good. Note, however, that when β becomes small ($\beta < 1.3$),

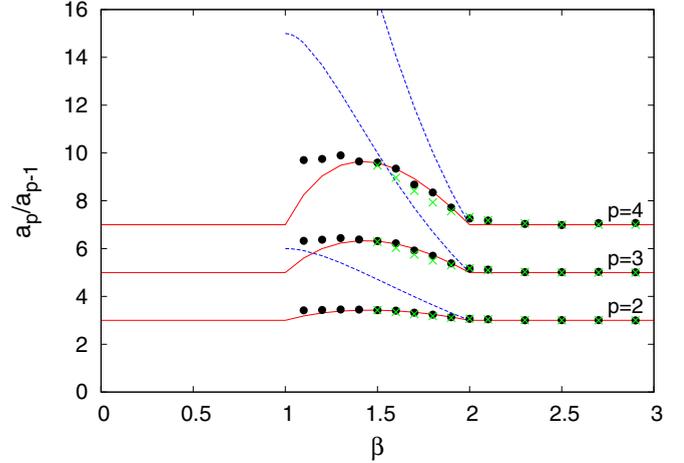


FIG. 6. (Color online) Moment ratios a_p/a_{p-1} (for $p = 2, 3,$ and 4) as a function of β . The solid red lines correspond to the analytical results. The two Gaussian regimes ($a_p/a_{p-1} = cst = 2p - 1$) are separated by non-Gaussian scalings in the interval $1 < \beta < 2$, given by Eq. (15). The dots correspond to the quantity $Q_p(t)$ computed from simulations ($q = 0.1, t = 10^6, 5 \times 10^4$ runs). The green crosses correspond to $Q_p(t)$ computed from exact numerical resolution of Eq. (1), with $q = 0.1$ and $t = 10^4$. The ratios a_p/a_{p-1} corresponding to the CTRW model (dotted blue lines) are shown for comparison for the same values of p . In that case, β is the waiting-time exponent.

the computed $Q_p(t)$ tends to be larger than the theoretical value. This is because diffusion is very slow in these cases and therefore the scaling regime becomes very difficult to reach in finite-time simulations. Recall that the scaling regime settles only when $M_2(t) \gg 1$. In Ref. [26], the approach toward the scaling regime was studied in detail for the case $\beta = 0$. It was shown that the first correction to scaling slowly decayed as $1/\ln t$ and thus could be neglected only at extremely long times, which cannot be reached in practice in simulations.

As an additional test, we solved Eq. (1) numerically, calculated the moments from $P(n, t)$, and then calculated $Q_p(t)$. The results are indicated by crosses in Fig. 6. The necessary computer memory scales as t^2 in this method, and times beyond $t = 15\,000$ could not be reached in practice (this is a low value, compared to $t = 10^6$ used in simulations). The scaling regime is thus less easy to observe, specially in the subdiffusive case, and the discrepancy with the theoretical a_p/a_{p-1} is a bit larger than in simulations.

V. COMPARISON OF THE CASE $1 < \beta < 2$ WITH THE CTRW

A. Qualitative analogy

The subdiffusive law (9) is the consequence of a diverging average backward time $\langle \tau \rangle_t$ as $t \rightarrow \infty$. Interestingly, this situation is analogous to that of the continuous time random walk (CTRW), where subdiffusion arises due to diverging waiting times [50,51]. In this well-known renewal Markovian process, an unbiased random walker remains at its current location during a random time τ , independent and identically distributed from a distribution $\psi(\tau)$, before performing the next step to a nearest-neighbor site [43]. If $\psi(\tau) \sim \tau^{-\beta}$

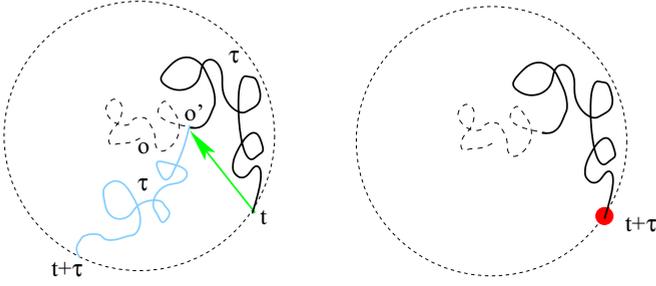


FIG. 7. (Color online) Qualitative comparison between the current model (left) and a CTRW (right); see text.

with $1 < \beta < 2$ (or setting $\beta = 1 + \mu$, if $0 < \mu < 1$), then $\langle \tau \rangle = \infty$ and a diffusion constant cannot be properly defined. In this case, motion is subdiffusive with $M_2(t) \sim t^{\beta-1}$, the same scaling as in Eq. (9).

The following qualitative argument can be made to explain the similar behaviors of the mean-square displacements in these two models. Let us consider a walker in our model starting from the origin O and taking a relocation step at a time t (Fig. 7, left). This step, represented by the arrow, brings the walker back to some position O' that was occupied earlier, τ time units before time t , where τ is a random variable $\leq t$ drawn from the probability density function $F(\tau) \propto \tau^{-\beta}$. Just before taking this step, the walker was at some distance from the origin, represented by the large circle. In order to reach this circle again from O' , the walker will typically need to diffuse during another time interval τ again (assuming stationarity in the process). Hence, only at about time $t + \tau$ may the walker's displacement grow beyond the radius of the circle. This situation can be seen as analogous to keeping the walker immobile at the position reached at t during a time τ , which is the basic rule of the CTRW (Fig. 7, right).

However, one may intuitively guess that the two models differ quantitatively, namely that their probability densities $P(n, t)$ are given by different scaling functions $g(x)$. This is actually the case, since $g(x)$ is exponential when $\beta \rightarrow 1$ in the one-dimensional CTRW, i.e., $g^{\text{CTRW}}(x) \rightarrow \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|x|)$ [11], instead of the Gaussian form in the memory model. More generally, for $1 < \beta < 2$ the expression of $g^{\text{CTRW}}(x)$ is exactly known in terms of the fully asymmetric Lévy law of index $(\beta - 1)/2$. Its moments can be calculated exactly, giving [11]

$$\left. \frac{a_p}{a_{p-1}} \right|_{\text{CTRW}} = (2p-1)(p-1) \frac{\Gamma[\mu(p-1)]\Gamma(\mu+1)}{\Gamma(\mu p)}. \quad (41)$$

The above expression indeed differs from (15) or (18). The limit distribution $g(x)$ in the present model does not reduce, to our knowledge, to a standard distribution. Figure 6 displays a_p/a_{p-1} as a function of β for $p = 2, 3$, and 4 for the CTRW and the present model, in which deviations from Gaussianity are weaker.

B. Large x behavior of $g(x)$

Further insights on $g(x)$ in the memory model can be gained by studying its behavior at large x for β in the range $(1, 2)$. This asymptotic behavior can be derived by analyzing the moments

in the limit $p \gg 1$. Knowing the properties of $g(x)$ is also useful to understand how the distribution tends to a Gaussian instead of an exponential as $\beta \rightarrow 1^+$.

The large p behavior of (18) for any $0 < \mu < 1$ can be obtained by using the Stirling formula $\Gamma(az + b) \simeq \sqrt{2\pi} e^{-az} (az)^{az+b-1/2}$ for large z :

$$\frac{a_p}{a_{p-1}} \simeq \chi_\mu p^{2-\mu} \quad \text{with} \quad \chi_\mu = 2 \frac{\frac{\pi\mu}{\sin(\pi\mu)} - 1}{\Gamma(1-\mu)\mu^\mu}. \quad (42)$$

Clearly, the asymptotic result (42) is valid as long as χ_μ is nonzero. This prefactor is strictly positive if $0 < \mu \leq 1$ but vanishes at $\mu = 0$ and this case will receive special attention below.

If $0 < \mu < 1$, Eq. (42) implies that the scaling function follows, at large x , the following leading behavior:

$$g(x) \propto e^{-b_\mu |x|^{1/\mu}}, \quad |x| \gg 1, \quad (43)$$

with

$$\delta_\mu = \frac{1}{1-\mu/2} \quad \text{and} \quad b_\mu = 2\chi_\mu^{-\frac{\delta}{2}}/\delta_\mu. \quad (44)$$

Let us notice that the large- x behavior above is similar to that of the CTRW with a waiting-time distribution $\psi(\tau) \sim \tau^{-(1+\mu)}$, namely the leading part of $g^{\text{CTRW}}(x)$ is also of the form (43) and with the same exponent δ_μ as given by (44). This property stems from the fact that the moment relation for CTRW, Eq. (41), also becomes proportional to $p^{2-\mu}$ at large p , like in (42). However, the scaling functions do differ in the two models since the prefactors χ_μ (and therefore b_μ in the exponential) differ.

This difference increases drastically as $\mu \rightarrow 0$: It is easy to see from (42) that

$$\chi_\mu \rightarrow 0 \quad \text{as} \quad \mu \rightarrow 0, \quad (45)$$

whereas $\chi_{\mu=0}^{\text{CTRW}}$ is > 0 . The fact that the prefactor vanishes is a crucial property of the memory model. It indicates that the leading term of the scaling function at large x is no longer given by (43) and (44). A higher-order calculation shows that, for any p ,

$$\frac{a_p}{a_{p-1}} \rightarrow 2p-1 \quad \text{as} \quad \mu \rightarrow 0^+, \quad (46)$$

which confirms our previous result that $g(x)$ becomes Gaussian again at $\mu = 0$. An expansion of the moment relation (18) at large p and small μ generalizes Eq. (42),

$$\frac{a_p}{a_{p-1}} \simeq \frac{\pi^2 \mu^2}{3} p^{2-\mu} + 2p-1, \quad (p \gg 1, \mu \ll 1). \quad (47)$$

Therefore the exponential behavior of $g(x)$ at large x (i.e., $a_p/a_{p-1} \propto p^2$ for large p) is avoided at $\mu = 0$ due to the vanishing prefactor and the distribution adopts a Gaussian profile instead.

VI. LÉVY-LIKE DISTRIBUTIONS OF RELOCATION LENGTHS

Every time the walker uses memory, it performs a relocation step, of length, say, ℓ , corresponding to the distance between its current location and the newly chosen site. These lengths ℓ

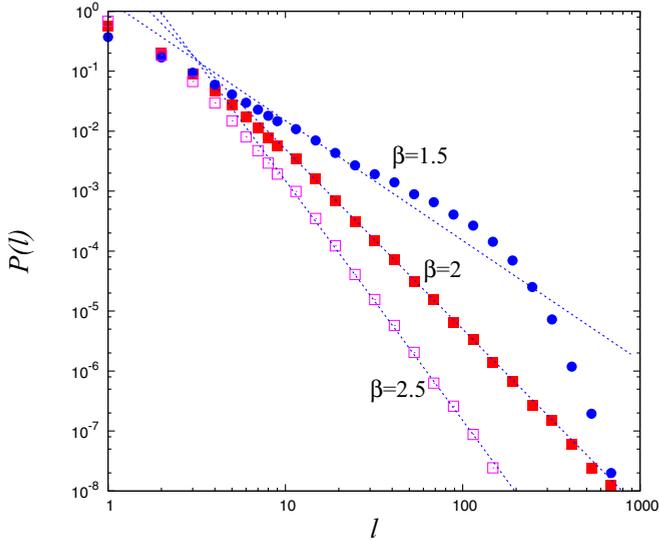


FIG. 8. (Color online) Probability distribution function of the relocation length ℓ in simulations for several values of β ($q = 0.01$ in all cases). The solid lines have slopes $-(2\beta - 1)$, i.e., -4 , -3 , and -2 , from bottom to top. The variance of the distribution thus becomes infinite at the subdiffusion threshold $\beta = 2$.

are represented by the thick (black) vertical lines in the one-dimensional simulations of Fig. 2. They can be quite large in the subdiffusive regime. We present below a scaling argument to obtain an expression for the probability distribution of ℓ .

Since the site occupied after a relocation is the site that was occupied τ time units ago, one may expect that, typically,

$$\ell \sim [M_2(\tau)]^{1/2}. \quad (48)$$

In the case with $\beta > 2$, we know that $M_2(t) \propto t$, thus the random variables ℓ and τ should be related through the scaling relation $\ell \sim \tau^{1/2}$. Denoting by $P(\ell)$ the distribution function of ℓ , probability conservation implies $P(\ell)d\ell = \psi(\tau)d\tau$, where $\psi(\tau) \sim \tau^{-\beta}$. Hence

$$P(\ell) \sim \ell (\ell^2)^{-\beta} \sim \ell^{-(2\beta-1)}. \quad (49)$$

Thus, when $\beta = 2$, the distribution $P(\ell)$ obeys a ‘‘Lévy-like’’ power-law distribution with marginal exponent $2\beta - 1 = 3$, which is characterized by an infinite variance $\langle \ell^2 \rangle$. If $\beta > 2$, Eq. (49) indicates that the step-length distribution exponent is > 3 , outside the Lévy range, and $\langle \ell^2 \rangle$ is therefore finite. Note that, unlike in a genuine Lévy process, the steps in the present model are long-range correlated and thus not independent. Consequently, our model exhibits a paradoxical behavior: The divergence of the variance of the relocation length ℓ happens at the onset of subdiffusion, and not superdiffusion, as in usual Lévy flights. This important feature can be understood by the fact that when the relocation lengths become large, they also become strongly anticorrelated and thus limit the walker’s diffusion overall. The scaling law predicted in (49) agrees very well with simulation results, as shown in Fig. 8.

The case $\beta \in (1, 2)$, corresponding to subdiffusive motion, is more complicated and the above argument to obtain $P(\ell)$ is not applicable strictly speaking. Equation (48) may not hold due to nonergodic effects, which are well known to be strong in subdiffusive processes such as the CTRW [52,53]. If we

naively assume that (48) is correct (which makes the implicit assumption that the mean-square displacement of the walker during an interval of time τ is independent of the time taken as the origin along the trajectory), we have $\ell \sim \tau^{(\beta-1)/2}$ from (48), which gives $P(\ell) \sim \ell^{\frac{2}{\beta-1}-1} [\ell^{\frac{2}{\beta-1}}]^{-\beta} \sim \ell^{-3}$. Hence, the relocation length distribution would remain at the border of the Lévy range for any $1 < \beta < 2$.

But this prediction is not confirmed by the numerical results of Fig. 8, which show that for intermediate values of ℓ the distribution $P(\ell)$ decays more slowly than the law ℓ^{-3} . If we assume that the average displacement ℓ taken along the trajectory looks normally diffusive, like in the ordinary CTRW process [52,54], then we should use the scaling $\ell \sim \tau^{1/2}$ instead of $\tau^{(\beta-1)/2}$. This relation leads to the same form (49) obtained for $\beta > 2$. Thus the step-length exponent would be $2\beta - 1$, spanning the whole Lévy range (1,3) as β is varied is between 1 and 2. The numerical simulations support this qualitative prediction partly. As displayed by Fig. 8, $P(\ell)$ exhibits an intermediate power-law regime with an exponent quite close to $2\beta - 1$ (case $\beta = 1.5$). Nevertheless, the actual distribution seems more complicated than an inverse power law and it is truncated at large ℓ .

VII. CASE $\beta < 1$ AND THE SCALED BROWNIAN MOTION

Scaled Brownian motion (SBM) is a well-known anomalous diffusion process with Gaussian property. Typically, it is simply generated by rescaling the time variable of an ordinary Brownian motion as $t \rightarrow t^\alpha$, with $0 < \alpha < 2$ a constant (see, e.g., Ref. [44]). Thus, the diffusion front $P(n, t)$ of a SBM is Gaussian, but with variance $2D_0 t^\alpha$, instead of $2D_0 t$ in the original process. Consequently, it obeys a Fokker-Planck equation (in continuous space) of the form

$$\frac{\partial P}{\partial t} = D(t) \frac{\partial^2 P}{\partial n^2} \quad (50)$$

with $D(t)$ a time-dependent diffusion coefficient, given by $D(t) = \alpha D_0 t^{\alpha-1}$.

When $\beta < 1$ in the present model, and in the very long time limit, $P(n, t)$ becomes a Gaussian, too, but with variance $M_2(t) \simeq K_\beta \ln t$, where the generalized mobility K_β is given by Eq. (10). Therefore, setting $T = \ln t$ we see that $P(n, T)$ obeys the simple effective diffusion equation $\partial_T P \simeq \frac{K_\beta}{2} \partial_{n^2} P$ in the limit $T \gg 1$. Coming back to the time variable, the equation becomes

$$\frac{\partial P}{\partial t} \simeq \frac{K_\beta}{2t} \frac{\partial^2 P}{\partial n^2}, \quad (51)$$

which is a SBM in the limit $\alpha \rightarrow 0$, i.e., with $D(t) \propto 1/t$ at large times. This result highlights the nonstationary nature of the process due to the very long-ranged memory, which gradually slows down the walker. This mapping is also interesting because it provides a concrete example of a random walk with constant parameters and constant time increment that can be described as a SBM asymptotically [at least, as far as the probability function $P(n, t)$ is concerned]. Here the SBM time rescaling is not of the power-law form usually considered but is given by

$$t \rightarrow \ln t. \quad (52)$$

Conversely, this type of SBM can be considered as a mean-field model for the path-dependent random walks studied here (individual trajectories clearly differ in the two processes). As noted earlier, we emphasize that this correspondence is valid only at very large times. At moderate $\ln t$ (e.g., for $t \sim 10^{10}$ or any numerical simulation time in practice), the memory model does not obey simple scaling *a priori*.

The SBM analogy could be used to infer the first-passage properties of the model (although they would be limited to the regime of extremely long times). In addition, the nonergodic properties of SBM systems have been recently studied in detail in Refs. [55,56] and could be used here.

VIII. CONCLUSIONS

Non-Markovian, path-dependent processes are in general difficult to analyze for a single particle, since a rigorous description usually requires the introduction of multiple-time distribution functions that are related to each other via a hierarchy of master equations [16].

Here we have introduced a path-dependent random-walk model that does not exhibit such complications and can be described by the single-time distribution function. This property stems from the linear nature of the model, in which the walker can perform nonlocal steps backwards in time. With this form of self-attraction, the process can be exactly described by one master equation. Additionally, the equation is solved exactly in the long-time limit thanks to the scaling hypothesis. Therefore, a rather precise picture of the dynamics is obtained, which is not limited to the derivation of the MSD.

The model exhibit three regimes, depending on how fast memory decays. In the weakly non-Markovian regime, diffusion is asymptotically normal but with a reduced diffusion coefficient. A transition to a second regime occurs if the mean memory time diverges, i.e., when the memory kernel decays as $\tau^{-\beta}$ with $\beta = 2$. In this case, recurrence to previously visited sites increases sharply and motion becomes subdiffusive. When memory decays very slowly, slower than τ^{-1} , a third, “ultra-slow,” diffusion regime with logarithmic behavior settles.

The diffusion front is a function which is practically unknown for reinforced walk models and almost never studied for other non-Markovian models which are solvable for the MSD [13–15] (but see Ref. [57] for a discussion on the so-called elephant walk). The knowledge of the scaling function here has allowed us to establish useful connections between a reinforced walk and well-known Markovian models of anomalous diffusion: namely the CTRW ($1 < \beta < 2$) and the SBM ($\beta < 1$). The present model exhibits distinctive features, such as a new scaling function defined through Eq. (18), and anticorrelated Lévy steps with exponent -3 at the subdiffusive transition. The visitation and first-passage statistics are also likely to be peculiar in this model, and their study certainly deserves future work.

Our results also unveil a novel, memory-driven mechanism for the emergence of Lévy flights. Whereas Lévy flights are a paradigmatic model of superdiffusive behavior in physical [11] and biological [35] systems, the model exposed here illustrates that (truncated) Lévy flights can be also closely linked to subdiffusion. This aspect is particularly relevant to the

mobility of many living organisms like humans [19,23,24,42] or nonhuman primates [58], to name a few, who often combine two apparently contradictory patterns: a power-law step length distribution and a very slow diffusion or limited space use.

ACKNOWLEDGMENTS

We thank L. Lacasa, H. Larralde, F. Leyvraz, I. Pérez-Castillo, A. Robledo, and S. Thurner for fruitful discussions and D. Aguilar for technical support. J.C.R.C. acknowledges supports from CONACYT, PAEP (Posgrado en Ciencias Físicas, UNAM), and the Universidad Autónoma de la Ciudad de México. This work was also supported by PAPIIT Grant No. IN101712 of the Universidad Nacional Autónoma de México.

APPENDIX A: DERIVATION OF THE MASTER EQUATION (1)

Let X_t be the position of the walker at time t and let $P(n, t + 1 | i_t, i_{t-1}, \dots, i_0)$ define the probability that $X_{t+1} = n$, given that $\{X_0, X_1, \dots, X_t\} = \{i_0, i_1, \dots, i_t\}$, a prescribed set of integers. The model is defined by the following:

$$P(n, t + 1 | i_t, i_{t-1}, \dots, i_0) = \frac{1-q}{2} \delta_{i_t, n-1} + \frac{1-q}{2} \delta_{i_t, n+1} + q \sum_{t'=0}^t p_t(t') \delta_{i_{t'}, n}, \quad (\text{A1})$$

where the last non-Markovian term indicates that n can be revisited if it has been visited at any previous time t' . In this memory movement mode, which is used with probability q , site n does not need to be a nearest-neighbor of the current position i_t , and the probability that n is chosen depends on all its previous visits weighted with the memory kernel $p_t(t')$.

Let $p(i_1, \dots, i_t | i_0)$ define the probability of a particular trajectory generated with the model rules and starting at i_0 . We have the general relation

$$\sum_{i_1} \dots \sum_{i_t} P(n, t + 1 | i_t, \dots, i_0) p(i_1, \dots, i_t | i_0) = P(n, t + 1 | i_0), \quad (\text{A2})$$

where $P(n, t + 1 | i_0)$ is the single-time occupation probability, evaluated at time $t + 1$. In addition, for any integer t' in $[0, t]$,

$$\sum_{i_1} \dots \sum_{\substack{i_k \\ k \neq t'}} \dots \sum_{i_t} p(i_1, \dots, i_t | i_0) = P(i_{t'}, t' | i_0). \quad (\text{A3})$$

Multiplying Eq. (A1) by $p(i_1, \dots, i_t | i_0)$ and summing over all possible values of i_1, \dots, i_t , we obtain, using Eqs. (A2) and (A3),

$$P(n, t + 1 | i_0) = \frac{1-q}{2} \sum_{i_t} \delta_{i_t, n-1} P(i_t, t | i_0) + \frac{1-q}{2} \sum_{i_t} \delta_{i_t, n+1} P(i_t, t | i_0) + q \sum_{t'=0}^t p_t(t') \sum_{i_{t'}} \delta_{i_{t'}, n} P(i_{t'}, t' | i_0), \quad (\text{A4})$$

or

$$P(n, t+1|i_0) = \frac{1-q}{2} P(n-1, t|i_0) + \frac{1-q}{2} P(n+1, t|i_0) + q \sum_{t'=0}^t p_t(t') P(n, t'|i_0), \quad (\text{A5})$$

This is the master equation (1) for the single particle. In this paper, $i_0 = 0$ and we rewrite $P(n, t|i_0) \rightarrow P(n, t)$ for convenience.

APPENDIX B: MEMORY TERM IN THE SUBDIFFUSIVE CASE

We choose a fixed $\epsilon \ll 1$ and t large enough so $\epsilon t \gg 1$. Hence, the memory term of Eq. (20) for the evolution of $M_2(t)$ reads as follows:

$$\begin{aligned} \mathcal{F}\{Kt^\nu\} &= K \sum_{\tau=0}^t \frac{(t-\tau)^\nu}{(1+\tau)^\beta} = Kt^\nu \sum_{\tau=0}^t \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta} \\ &= Kt^\nu \left[\sum_{\tau=0}^{\epsilon t} \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta} + \sum_{\tau=\epsilon t+1}^t \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta} \right] \\ &\simeq Kt^\nu \left[\sum_{\tau=0}^{\epsilon t} \frac{1}{(1+\tau)^\beta} - \frac{\nu}{t} \sum_{\tau=0}^{\epsilon t} \frac{\tau}{(1+\tau)^\beta} \right. \\ &\quad \left. + \sum_{\tau=\epsilon t+1}^t \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta} \right]. \quad (\text{B1}) \end{aligned}$$

The first sum in the brackets of (B1) can be written as

$$\sum_{\tau=0}^{\epsilon t} \frac{1}{(1+\tau)^\beta} = \zeta(\beta) - \sum_{\tau=\epsilon t+1}^{\infty} \frac{1}{(1+\tau)^\beta} \quad (\text{B2})$$

$$\simeq \zeta(\beta) - \int_{\epsilon t}^{\infty} d\tau (1+\tau)^{-\beta} \quad (\text{B3})$$

$$\simeq \zeta(\beta) - \frac{(\epsilon t)^{-\beta+1}}{\beta-1}, \quad (\text{B4})$$

due to the Euler-Maclaurin expansion of the discrete sum. Similarly, the third sum in the brackets of (B1) reads as follows:

$$\sum_{\tau=\epsilon t+1}^t \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta} \simeq \int_{\epsilon t}^t d\tau \frac{(1-\tau/t)^\nu}{(1+\tau)^\beta}, \quad (\text{B5})$$

$$\simeq t^{-\beta+1} \int_{\epsilon}^1 du \frac{(1-u)^\nu}{u^\beta}, \quad (\text{B6})$$

$$= O((\epsilon t)^{-\beta+1}). \quad (\text{B7})$$

The second term in the brackets of (B1) is

$$\frac{\nu}{t} \sum_{\tau=0}^{\epsilon t} \frac{\tau}{(1+\tau)^\beta} = O\left(t^{-1} \int_{\epsilon t}^{\epsilon t} d\tau \tau^{1-\beta}\right) = O((\epsilon t)^{-\beta+1}), \quad (\text{B8})$$

and thus negligible compared to (B4) and (B6). We then substitute (B4) and (B6) into (B1), and Eq. (24) is obtained after taking the limit $\epsilon \rightarrow 0$.

-
- [1] E. Bolthausen and U. Schmock, *Ann. Probab.* **25**, 531 (1997).
[2] H. G. Othmer and A. Stevens, *SIAM J. Appl. Math.* **57**, 1044 (1997).
[3] R. Pemantle, *Probab. Surv.* **4**, 1 (2007).
[4] A. O. Gautestad and I. Mysterud, *Am. Nat.* **165**, 44 (2005).
[5] A. O. Gautestad and I. Mysterud, *Ecol. Complex.* **3**, 44 (2006).
[6] B. Davis, *Probab. Theory Relat. Fields* **84**, 203 (1990).
[7] A. Ordemann, E. Tomer, G. Berkolaiko, S. Havlin, and A. Bunde, *Phys. Rev. E* **64**, 046117 (2001).
[8] J. W. Lee, *J. Phys. A: Math. Gen.* **31**, 3929 (1998).
[9] J. G. Foster, P. Grassberger, and M. Paczuski, *New J. Phys.* **11**, 023009 (2009).
[10] J. Choi, J. I. Sohn, K. I. Goh, and I. M. Kim, *Europhys. Lett.* **99**, 50001 (2012).
[11] J.-P. Bouchaud and A. Georges, *Phys. Rep.* **195**, 127 (1990).
[12] R. Metzler and J. Klafter, *Phys. Rep.* **339**, 1 (2000).
[13] G. M. Schütz and S. Trimper, *Phys. Rev. E* **70**, 045101(R) (2004).
[14] J. C. Cressoni, Marco Antonio Alves da Silva, and G. M. Viswanathan, *Phys. Rev. Lett.* **98**, 070603 (2007).
[15] M. Serva, *Phys. Rev. E* **88**, 052141 (2013).
[16] L. Peliti, *J. Phys.* **46**, 1469 (1985).
[17] L. Peliti and L. Pietronero, *Riv. Nuovo Cimento* **10**, 1 (1987).
[18] V. B. Sapozhkov, *J. Phys. A: Math. Gen.* **27**, L151 (1994).
[19] C. Song, T. Koren, P. Wang, and A.-L. Barabási, *Nat. Phys.* **6**, 818 (2010).
[20] R. Nathan *et al.*, *Proc. Natl. Acad. Sci. USA* **105**, 19052 (2008).
[21] D. W. Sims *et al.*, *Nature* **451**, 1098 (2008).
[22] N. E. Humphries, H. Weimerskirch, N. Queiroz, E. J. Southall, and D. W. Sims, *Proc. Natl. Acad. Sci. USA* **109**, 7169 (2012).
[23] D. Brockmann, L. Hufnagel, and T. Geisel, *Nature* **439**, 462 (2006).
[24] M. C. González, C. A. Hidalgo, and A.-L. Barabási, *Nature* **453**, 779 (2008).
[25] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, *Science* **327**, 1018 (2010).
[26] D. Boyer and C. Solis-Salas, *Phys. Rev. Lett.* **112**, 240601 (2014).
[27] J. A. Merkle, D. Fortin, and J. M. Morales, *Ecol. Lett.* **17**, 924 (2014).
[28] P. R. Moorcroft and M. A. Lewis, *Mechanistic Home Range Analysis* (Princeton University Press, Princeton, 2006).
[29] L. Börger, B. D. Dalziel, and J. M. Fryxell, *Ecol. Lett.* **11**, 637 (2008).
[30] B. van Moorter *et al.*, *Oikos* **118**, 641 (2009).
[31] D. Boyer and P. D. Walsh, *Phil. Trans. R. Soc. A* **368**, 5645 (2010).
[32] J. Nabe-Nielsen, J. Tougaard, J. Teilmann, K. Lucke, and M. C. Forchhammer, *Oikos* **122**, 1307 (2013).
[33] P. Turchin, *Quantitative Analysis of Movement* (Sinauer, Sunderland, MA, 1998).

- [34] E. A. Codling, M. J. Plank, and S. Benhamou, *J. R. Soc. Interface* **5**, 813 (2008).
- [35] G. M. Viswanathan, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, *The Physics of Foraging* (Cambridge University Press, Cambridge, 2011).
- [36] O. Bénichou, M. Coppey, M. Moreau, P.-H. Suet, and R. Voituriez, *Phys. Rev. Lett.* **94**, 198101 (2005).
- [37] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, *Rev. Mod. Phys.* **83**, 81 (2011).
- [38] F. Bartumeus, *Oikos* **118**, 488 (2009).
- [39] J. M. Morales, D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell, *Ecology* **85**, 2436 (2004).
- [40] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, *Phys. Life Rev.* **5**, 133 (2008).
- [41] F. Bartumeus, M. G. E. da Luz, G. M. Viswanathan, and J. Catalan, *Ecology* **86**, 3078 (2005).
- [42] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong, *IEEE/ACM Trans. Netw.* **19**, 630 (2011).
- [43] E. W. Montroll and G. H. Weiss, *J. Math. Phys.* **6**, 167 (1965).
- [44] S. C. Lim and S. V. Muniandy, *Phys. Rev. E* **66**, 021114 (2002).
- [45] G. U. Yule, *Phil. Trans. R. Soc. (London) B* **213**, 21 (1925).
- [46] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [47] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).
- [48] M. R. Evans and S. N. Majumdar, *Phys. Rev. Lett.* **106**, 160601 (2011).
- [49] Namely $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$, $x\Gamma(x) = \Gamma(x + 1)$, and $\Gamma(x)\Gamma(1 - x) = \pi/\sin(\pi x)$.
- [50] M. F. Shlesinger, *J. Stat. Phys.* **10**, 421 (1974).
- [51] M. Kotulski, *J. Stat. Phys.* **81**, 777 (1995).
- [52] Y. He, S. Burov, R. Metzler, and E. Barkai, *Phys. Rev. Lett.* **101**, 058101 (2008).
- [53] A. Rebenshtok and E. Barkai, *Phys. Rev. Lett.* **99**, 210601 (2007).
- [54] A. Lubelski, I. M. Sokolov, and J. Klafter, *Phys. Rev. Lett.* **100**, 250602 (2008).
- [55] F. Thiel and I. M. Sokolov, *Phys. Rev. E* **89**, 012115 (2014).
- [56] J.-H. Jeon, A. V. Chekin, and R. Metzler, [arXiv:1405.2193](https://arxiv.org/abs/1405.2193) [cond-mat.stat-mech].
- [57] M. A. A. da Silva, J. C. Cressoni, G. M. Schütz, G. M. Viswanathan, and S. Trimper, *Phys. Rev. E* **88**, 022115 (2013).
- [58] D. Boyer, M. C. Crofoot, and P. D. Walsh, *J. R. Soc. Interface* **9**, 842 (2012).