

**Statistics of occupation times and connection to local properties of nonhomogeneous random walks**Mattia Radice<sup>✉,\*</sup>, Manuele Onofri, Roberto Artuso<sup>✉</sup>, and Gaia Pozzoli*Dipartimento di Scienza e Alta Tecnologia and Center for Nonlinear and Complex Systems, Università degli studi dell'Insubria, Via Valleggio 11, I-22100 Como, Italy  
and I.N.F.N. Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*

(Received 30 January 2020; accepted 16 March 2020; published 3 April 2020)

We consider the statistics of occupation times, the number of visits at the origin, and the survival probability for a wide class of stochastic processes, which can be classified as renewal processes. We show that the distribution of these observables can be characterized by a single exponent, that is connected to a local property of the probability density function of the process, viz., the probability of occupying the origin at time  $t$ ,  $P(t)$ . We test our results for two different models of lattice random walks with spatially inhomogeneous transition probabilities, one of which of non-Markovian nature, and find good agreement with theory. We also show that the distributions depend only on the occupation probability of the origin by comparing them for the two systems: When  $P(t)$  shows the same long-time behavior, each observable follows indeed the same distribution.

DOI: [10.1103/PhysRevE.101.042103](https://doi.org/10.1103/PhysRevE.101.042103)**I. INTRODUCTION**

Survival and persistence problems for stochastic processes are often considered in the study of critical phenomena in equilibrium and nonequilibrium systems, for instance, spin systems in one or higher dimensions [1], phase-ordering kinetics [2–4], and twisted nematic liquid crystals exhibiting planar Ising model dynamics [5]. Such problems are characterized by the *persistence exponent* [6], which gives the scaling of the probability that the order parameter  $x(t)$  (for example, the magnetization of a ferromagnet) of a system quenched from the disordered phase to its critical point has not changed sign in a time interval  $t$  following the quench [7,8]. In this context the time evolution of the order parameter is treated as a stochastic process and other questions regarding the statistics of  $x(t)$  naturally arise: For instance, one can ask what is the fraction  $t^+$  of time in which the process has assumed positive values [9–11], which is associated, e.g., with the mean magnetization. For many physical systems [12–14] the distribution of the mean magnetization displays a U-shaped curve, reflecting the fact that, contrary to intuition, the order parameter is more likely to preserve its sign during the observation time. Interestingly, it is found that the exponent of the singularities at the outer values is closely related to the persistent exponent. This connection can be proved [9] by considering  $x(t)$  as generated by a renewal process: starting from the initial state  $x(t_0)$ , during the time evolution the process resets itself to the initial condition at random times  $t_i$ ,  $i = 1, 2, \dots$ , such that the intervals  $\Delta t_i = t_i - t_{i-1}$  are independent and identically distributed random variables. In this setting it is also worth asking what is the number of renewals observed up to time  $t$ , which is found to follow a Mittag-Leffler distribution of an adequate parameter. The value of the parameter and therefore the shape of the distribution depend

on the scaling exponent of the probability density function of waiting times between renewals: For a distribution which decays asymptotically as  $F(\Delta t) \sim \Delta t^{-1-\theta}$ , with  $0 \leq \theta < 1$ , one obtains a Mittag-Leffler of parameter  $\theta$  [15–18].

The advantage of renewal theory is that it applies to a broad range of stochastic processes, including, for example, random walks with spatially inhomogeneous transition probabilities and correlations between steps. Note that non-homogeneous diffusion has attracted recently attention in a variety of different contexts; see, for example, Refs. [19–22]. The major difficulty in applying renewal theory to general diffusion processes is that one has to determine the waiting-time distribution, which is often a difficult task to perform, especially when translational symmetry is broken or for walks of non-Markovian nature. In the language of random walks, for example, this corresponds to computing the probabilities of first return to the starting position, which can be analytically done only in few cases. In this paper we will show that it is possible to obtain the fraction of time spent in the positive axis, the number of renewals, and the persistence exponent just by considering the probability  $P(t)$  that at time  $t$  the process is returned to the initial state:

$$P(t) = \Pr \{x(t) = x(t_0)\}. \quad (1)$$

The paper is organized as follows: In the next section we introduce the class of processes to which our results apply; then we present the known results regarding the occupation time of the positive axis (Sec. III), the number of returns at the origin (Sec. IV) and the survival probability (Sec. V), and discuss how to establish a connection between the three, starting from the probability of occupying the initial state; in Sec. VI we describe the stochastic processes we have considered in our simulations and show the numerical results; finally in Sec. VII we draw our conclusions.

\*m.radice1@uninsubria.it

**II. THE CLASS OF PROCESSES CONSIDERED IN THE PAPER**

The class of stochastic processes for which the results of this paper apply is similar to that considered in a classic paper by Lamperti [23]. This class consists of processes (not necessarily Markovian) whose time evolution is described by a discrete parameter  $n$ , with the property that the states are divided into two sets, say  $A$  and  $B$ , which communicate through the occurrence of a recurrent state  $x_0$ , assumed as the initial state. More precisely, denoting with  $x_n$  the state at time  $n$  of the stochastic process, starting from  $x_0$ , we consider processes such that if  $x_{n-1} \in A$  and  $x_{n+1} \in B$  or vice versa, then  $x_n = x_0$ ; moreover, the occupation of  $x_0$  is a *persistent recurrent event* [24], by which we mean that calling  $F_n$  the probability that the process returns to  $x_0$  for the first time after  $n$  steps, having started from  $x_0$ , then

$$\sum_{n=1}^{\infty} F_n = 1, \tag{2}$$

i.e., the return to  $x_0$  is certain. We will furthermore assume that the return to  $x_0$  defines a renewal event, so that  $x_n$  can be treated as a renewal process.

In practice, in this paper we will consider one-dimensional random walks on the integer lattice, starting from  $x_0 = 0$ , with nearest-neighbor jumps (without specifying the rules followed by the jumps). We will call  $A(B)$  the set of positive(negative) integers, and will assume that the occupation of state  $x_0$ , i.e., the return to the origin, is a persistent recurrent event.

**III. OCCUPATION TIME OF THE POSITIVE AXIS**

For the class of processes we are considering in this paper, the result stated in [23] provides the distribution, as the number of steps tends to infinity, of the fraction of time spent in the positive axis, which we call the *Lamperti distribution*  $\mathcal{G}_{\eta,\rho}(\xi)$ . Such a distribution is defined through two parameters. The first parameter is

$$\eta = \lim_{n \rightarrow \infty} \mathbb{E} \left( \frac{k_n}{n} \right), \tag{3}$$

where  $k_n$  denotes the occupation time of set  $A$  up to step  $n$ , using the convention that the occupation of the origin is counted or not according to whether the last other state occupied was in  $A$  [25]. Clearly  $\eta$  is equal to  $1/2$  if the process is symmetric with respect to  $A$  and  $B$ . The second parameter is defined as the limit,

$$\rho = \lim_{z \rightarrow 1} \frac{(1-z)F'(z)}{1-F(z)}, \tag{4}$$

where  $F(z)$  denotes the generating function of the recurrence times of  $x_0$ , i.e., the first return probabilities  $F_n$ :

$$F(z) = \sum_{n=1}^{\infty} F_n z^n. \tag{5}$$

We have the following [23].

*Theorem 1.* Let  $x_n$  be the process described above. Then

$$\lim_{n \rightarrow \infty} \Pr\{k_n/n \leq u\} \equiv \mathcal{G}_{\eta,\rho}(u) \tag{6}$$

exists if and only if both limits  $0 \leq \eta \leq 1$ , Eq. (3), and  $0 \leq \rho \leq 1$ , Eq. (4), exist. In this case  $\mathcal{G}_{\eta,\rho}(u)$  is the distribution on  $[0, 1]$  which, provided both  $\eta$  and  $\rho \neq 0, 1$ , has the density,

$$\mathcal{G}'_{\eta,\rho}(u) = \mathcal{N} \frac{u^\rho(1-u)^{\rho-1} + u^{\rho-1}(1-u)^\rho}{a^2 u^{2\rho} + 2au^\rho(1-u)^\rho \cos(\pi\rho) + (1-u)^{2\rho}}, \tag{7}$$

where

$$a = \frac{1-\eta}{\eta}, \tag{8}$$

$$\mathcal{N} = \frac{a \sin(\pi\rho)}{\pi}. \tag{9}$$

For  $\eta = 0, 1$ , and  $0 < u < 1$ , the distribution is

$$\mathcal{G}_{\eta,\rho}(u) = \begin{cases} 1 & \text{for } \eta = 0 \\ 0 & \text{for } \eta = 1; \end{cases} \tag{10}$$

for  $\rho = 1$  we have

$$\mathcal{G}_{\eta,1}(u) = \begin{cases} 0 & \text{for } u < \eta \\ 1 & \text{for } u \geq \eta, \end{cases} \tag{11}$$

while for  $\rho = 0, 0 \leq u < 1$ ,

$$\mathcal{G}_{\eta,0}(u) = 1 - \eta. \tag{12}$$

An important observation is that the existence of the limit (4) is equivalent to a condition on the form the generating function  $F(z)$  must assume, namely (see ref. [23]),

$$F(z) = 1 - (1-z)^\rho L\left(\frac{1}{1-z}\right), \tag{13}$$

where  $L(x)$  is a *slowly varying* function, by which we mean a continuous function, positive for large enough  $x$ , that for any  $y > 0$  satisfies

$$\lim_{x \rightarrow \infty} \frac{L(yx)}{L(x)} = 1. \tag{14}$$

Equation (13) suggests that the distribution of the occupation time can be determined by evaluating the analytical expression of  $F(z)$ . As we have already observed, however, in general the computation of the first return probabilities is hard to perform. Nevertheless, since we are taking as  $x_0$  the site  $j = 0$ , one can use a well-known formula, valid for any renewal process, relating  $F(z)$  to the generating function  $P(z)$  of the probabilities of occupying the origin at time  $n$ ,  $P_n$ , which reads [26]

$$F(z) = 1 - \frac{1}{P(z)}, \tag{15}$$

to recast condition (13) as

$$P(z) = \frac{1}{(1-z)^\rho} H\left(\frac{1}{1-z}\right), \tag{16}$$

where  $H(x) = 1/L(x)$  is a slowly varying function. In particular, Eq. (16) shows that the parameter  $\rho$  of the Lamperti distribution appears as an exponent in the generating function  $P(z)$ .

A first consequence is that the parameter  $\rho$  can be computed by evaluating  $P_n$ . In order to show this, we make use of the following Tauberian theorem [27].

*Theorem 2.* Let  $g_n \geq 0$  and suppose that

$$\sum_{n=0}^{\infty} g_n z^n = G(z) \tag{17}$$

converges for  $0 \leq z < 1$ . Then

$$G(z) \sim \frac{1}{(1-z)^\gamma} \mathcal{H}\left(\frac{1}{1-z}\right), \quad z \rightarrow 1^- \iff$$

$$g_0 + \dots + g_n \sim \frac{1}{\Gamma(\gamma+1)} n^\gamma \mathcal{H}(n), \quad n \rightarrow \infty, \tag{18}$$

where  $\mathcal{H}(x)$  is a slowly varying function and  $\gamma \geq 0$ .

Furthermore, if the sequence  $\{g_n\}$  is ultimately monotonic and  $\gamma > 0$ , it also holds

$$G(z) \sim \frac{1}{(1-z)^\gamma} \mathcal{H}\left(\frac{1}{1-z}\right), \quad z \rightarrow 1^- \iff$$

$$g_n \sim \frac{1}{\Gamma(\gamma)} n^{\gamma-1} \mathcal{H}(n), \quad n \rightarrow \infty. \tag{19}$$

By using Eq. (16) and applying the theorem, one has that, for  $0 < \rho \leq 1$ ,  $P_n$  decays as

$$P_n \sim \frac{1}{\Gamma(\rho)} \frac{H(n)}{n^{1-\rho}}, \tag{20}$$

meaning that  $\rho$  is related to the exponent appearing in the long-time limit of the occupation probability of the origin.

We remark that this result connects the behavior of the process regarding the occupation time of the sets  $A$  and  $B$ , which is a nonlocal property, to a local property. For instance, for a simple symmetric random walk it is known that  $P_n$  decays with the power law  $P_n \sim n^{-1/2}$ , which corresponds to  $\rho = \frac{1}{2}$ . In this case the distribution of the occupation time follows the *first arcsin law* [28], which is recovered by Theorem 1 in the case  $\rho = \eta = \frac{1}{2}$ . For  $0 < \rho < 1$  the probability of being at the origin has the asymptotic decay  $P_n \sim n^{-(1-\rho)}$ , up to a factor given by the slowly varying function, and the distribution of the occupation time is represented by U-shaped curves. From formula (7) we see that the divergence of these curves at  $u = 0$  and  $u = 1$  is given exactly by the exponent  $1 - \rho$ . The situation is different for  $\rho = 1$ : We have  $P_n \sim H(n)$ , hence  $P_n$  does not decay as a power law. Instead it must behave for large  $n$  as a (ultimately) decreasing slowly varying function, converging to a constant. In this case the occupation time is split among the two sets, in such a way that the process spends a fraction  $\eta$  of time in  $A$  and the remaining in  $B$ . The distribution of the fraction of time in  $A$  is therefore a Dirac delta function centered around  $u = \eta$  and we will refer to this as the *ergodic case* [23]. In the opposite case,  $\rho = 0$ , regarding  $P_n$  we can only conclude that

$$\sum_{m=0}^n P_m \sim H(n), \tag{21}$$

where this time  $H(n)$  must be (ultimately) increasing. Since by using Eqs. (2) and (15), one can show that a necessary and sufficient condition for recurrence is the divergence of  $P(z)$  [26], we can say that  $H(n)$  must diverge, but we expect the divergence to be slow. In this sense, the case  $\rho = 0$  corresponds to a crossover for the occupation of  $x_0$  between

being or not a persistent recurrent event. This can be better understood by observing that the distribution of the occupation time has masses  $m_A = \eta$  on  $u = 1$  and  $m_B = 1 - \eta$  on  $u = 0$ , meaning that the process spends all the time either in  $A$ , with probability  $\eta$ , or in  $B$ , with probability  $1 - \eta$ .

#### IV. NUMBER OF VISITS AT THE ORIGIN

The number of renewals for a random walk is closely related to the number of visits at the origin. Indeed, if the return defines a renewal event,  $N$  visits at the starting site for a walker correspond to  $N - 1$  renewals. The distribution of the occupation time of the origin can be obtained from a classic result by Darling and Kac [29], where they showed that the limiting distribution of the occupation time of a set of finite measure for a Markov process is the Mittag-Leffler distribution:

$$\mathcal{M}_\nu(\xi) = \frac{1}{\nu \xi^{1+\frac{1}{\nu}}} L_\nu\left(\frac{1}{\xi^{\frac{1}{\nu}}}\right), \tag{22}$$

where  $L_\nu(x)$  denotes the Lévy one-sided density of parameter  $\nu$ , defined through the inverse Laplace transform from  $p$  to  $x$ :  $L_\nu(x) = \mathcal{L}^{-1}[\exp(-p^\nu)]$ ; the parameter  $\nu$  depends on the process itself. For the sake of clarity, here we briefly state the result, limiting ourselves to the case of random walks on a lattice—we point out, however, that the result holds in a more general setting.

Let  $x_n$  be a random walk on the integer lattice. Consider the generating function of the probabilities  $P_n(j|j_0)$  of arriving at site  $j$  in  $n$  steps, having started from  $j_0$ :

$$P_z(j|j_0) = \sum_{n=0}^{\infty} P_n(j|j_0) z^n. \tag{23}$$

Let  $V(j)$  be an integrable, non-negative function and suppose there exists a function  $\pi(z)$ ,  $\pi(z) \rightarrow \infty$  as  $z \rightarrow 1^-$ , and a positive constant  $c$ , such that

$$\lim_{z \rightarrow 1^-} \frac{1}{\pi(z)} \sum_j P_z(j|j_0) V(j) = c, \tag{24}$$

the convergence being uniform in  $E = \{j_0 | V(j_0) > 0\}$ . Then the following result holds [29].

*Theorem 3.* For some normalizing sequence  $\{u_n\}$  the limiting distribution of

$$\frac{1}{u_n} \sum_{m=0}^n V(x_m) \tag{25}$$

exists and it is nonsingular if and only if, for some  $0 \leq \nu < 1$ ,

$$\pi(z) = \frac{1}{(1-z)^\nu} \mathcal{H}\left(\frac{1}{1-z}\right), \tag{26}$$

where  $\mathcal{H}(x)$  is a slowly varying function. Moreover, if (26) is satisfied,  $u_n$  can be taken to be  $c\pi(1 - \frac{1}{n})$  and the limiting distribution is the Mittag-Leffler distribution  $\mathcal{M}_\nu(\xi)$ .

We will use this result to find the distribution of the occupation time of the origin. In order to do so, we take  $V(j) = \delta_{j,0}$  so that

$$\sum_j P_z(j|j_0) V(j) = \sum_j P_z(j|j_0) \delta_{j,0} = P_z(0|j_0). \tag{27}$$

Now, since  $\delta_{j,0} > 0$  only for  $j = 0$ , we have to prove the existence of  $\pi(z)$  of the desired form, Eq. (26), such that

$$\lim_{z \rightarrow 1^-} \frac{P_z(0|0)}{\pi(z)} = c. \tag{28}$$

Now, by definition,

$$P_z(0|0) \equiv P(z), \tag{29}$$

and we know from the discussion made in Sec. III [see Eq. (16)] that

$$P(z) = \frac{1}{(1-z)^\rho} H\left(\frac{1}{1-z}\right), \tag{30}$$

where  $H(x)$  is slowly varying. Then, for any positive constant  $c$ , we can take

$$\pi(z) = \frac{1}{c(1-z)^\rho} H\left(\frac{1}{1-z}\right), \tag{31}$$

and have

$$\lim_{z \rightarrow 1^-} \frac{P(z)}{\pi(z)} = c. \tag{32}$$

Hence, for the theorem stated above, the limiting distribution of the random variable,

$$T_n \equiv \frac{1}{H(n)n^\rho} \sum_{m=0}^n \delta_{x_m,0}, \tag{33}$$

describing the occupation time of the origin, is the Mittag-Leffler distribution of parameter  $\rho$ , for  $0 \leq \rho < 1$ .

We point out that when  $\rho = 0$  the Mittag-Leffler distribution becomes the exponential distribution, while for  $\rho = \frac{1}{2}$  we have a half-Gaussian:  $\mathcal{M}_{1/2}(\xi) = \pi^{-1/2} \exp(-\xi^2/4)$ , which is the limiting distribution for the simple symmetric random walk [28]. For  $\rho = 1$  one has a degenerate case, with the convergence,

$$\frac{1}{H(n)n} \sum_{m=0}^n \delta_{x_m,0} \rightarrow 1, \tag{34}$$

in probability, which is a kind of weak ergodic theorem [29], as the long-time limit of  $H(n)$  gives in this case the probability of occupying the origin, which decays to a constant (see discussion in Sec. III). This means that the process possesses a stationary distribution and the value of such a distribution at  $j = 0$  corresponds to the ensemble average of  $V(j)$ . Therefore we have the convergence of the time average of  $V(j)$  over a single trajectory to its ensemble average, so that the density of  $T_n$  converges to a Dirac delta function centered around  $\xi_0 = 1$ .

In Appendix A we show that in the long-time limit  $T_n$  is proportional to the number of visits at the origin up to step  $n$ , which we denote as  $M_n$ , rescaled for its mean value,

$$T_n \sim \frac{1}{\Gamma(1+\rho)} \frac{M_n}{\langle M_n \rangle}. \tag{35}$$

We may therefore conclude that the result states that the random variable,

$$\xi = \lim_{n \rightarrow \infty} \frac{1}{\Gamma(1+\rho)} \frac{M_n}{\langle M_n \rangle}, \tag{36}$$

follows a Mittag-Leffler distribution of parameter  $\rho$ , for  $0 \leq \rho < 1$ , and a degenerate Mittag-Leffler distribution for  $\rho = 1$ , whose density is

$$P(\xi) = \delta(\xi - 1). \tag{37}$$

We remark that the result also holds if  $x_n$  is not a Markov process, provided that the return to  $x_0$  defines a renewal event, so that the transition can be characterized by a waiting-time distribution between renewals, i.e., by the distribution of the first return times. This happens, for example, if  $x_n$  is symmetric with respect to the starting point. Indeed, we could obtain the same from renewal theory (see refs. [15–18] and references therein). It is worth observing that the parameter characterizing the distribution of the occupation time of the origin must be the same parameter of the Lamperti distribution, which in our setting describes the occupation time of the positive(negative) axis for a symmetric process. We also point out that a similar connection was proved in [9] by using scaling arguments, and also in the context of infinite ergodic theory for deterministic systems [30].

### V. DECAY OF THE SURVIVAL PROBABILITY

As we have seen, the Lamperti and the Mittag-Leffler distribution are closely related, all due to the particular form that the generating functions  $P(z)$  and  $F(z)$  must assume. We will show that this is also related to the asymptotic decay of the survival probability in the set  $A(B)$ . We define the survival probability in a set for a random walk on the integers with nearest-neighbor jumps as the probability  $Q_n$  of never leaving the setup to step  $n$ . If  $A$  is the set of positive integers, then, following the convention in [23] on how to count the occupation time, we have

$$Q_n = \Pr\{x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 | x_0 = 0\}, \tag{38}$$

with  $Q_0 = 1$ . Such a quantity can be computed exactly for random walks with i.i.d. jumps drawn from a continuous distribution, by using a well-known combinatorial identity known as the Sparre-Andersen theorem [31]:

$$Q(z) = \sum_{n=0}^{\infty} Q_n z^n = \exp \left[ \sum_{n=1}^{\infty} \frac{z^n}{n} \Pr\{x_n \geq 0\} \right]. \tag{39}$$

For any symmetric jump distribution, one obtains the behavior  $Q_n \sim n^{-1/2}$  for large  $n$ , independently of the distribution itself. It can be shown that such a decay also holds for walks with nearest-neighbor jumps, i.e., a particular case of noncontinuous jump distribution, provided that the jumps are symmetric, independent, and identically distributed [32]. Therefore, in the paradigmatic case of the simple symmetric random walk on the integers, one finds that the value  $\frac{1}{2}$  describes the power-law decay of the survival probability and gives the correct parameter describing both the Lamperti and the Mittag-Leffler distributions. However, no results for the survival probability are available if jumps are correlated or not identically distributed.

In our setting, we can obtain a relation between the survival probability  $Q_n$  and the *persistence* probability [9], namely the probability  $U_n$  of not observing any return up to time  $n$ , which

can be computed as

$$U_n = 1 - \sum_{m=0}^n F_m. \tag{40}$$

In Appendix B we show that the generating functions  $Q(z)$  and  $U(z)$  satisfy the relation,

$$Q(z) = \frac{1 + U(z)}{1 + (1 - z)U(z)}, \tag{41}$$

and that this implies  $Q(z) \sim U(z)$  as  $z \rightarrow 1$ , for  $0 \leq \rho < 1$ . This means that the survival and the persistence probabilities have the same behavior for large  $n$ .

By using Eq. (40) and the condition  $U_0 = 1$ , we can compute the generating function,

$$U(z) = \frac{1 - F(z)}{1 - z}, \tag{42}$$

and hence, by using Eq. (13), we find that  $U(z)$  must be of the form,

$$U(z) = \frac{1}{(1 - z)^{1-\rho}} L\left(\frac{1}{1 - z}\right), \tag{43}$$

where  $L(x) = 1/H(x)$  is a slowly varying function, and  $H(x)$  is the same appearing in Eq. (16). Since, as we already stated, for  $z \rightarrow 1$  we have  $Q(z) \sim U(z)$ , the use of the Tauberian theorem implies that the survival probability  $Q_n$  decays as

$$Q_n \sim \frac{1}{\Gamma(1 - \rho)} \frac{n^{-\rho}}{H(n)}. \tag{44}$$

Once again, the quantity of interest is characterized by the Lamperti parameter. We remark that this result holds for the class of processes we are considering, therefore not only for walks with i.i.d. jumps. For  $\rho = 1$  the Tauberian theorem only assures that as  $n \rightarrow \infty$ ,

$$\sum_{m=0}^n Q_m \sim \frac{1}{H(n)}, \tag{45}$$

where  $H(n)$  is a (ultimately) decreasing slowly varying function, hence the asymptotic relation (44) is not valid in this regime. We recall that in this case  $P_n$  does not decay as a power law (see Sec. III).

### VI. NUMERICAL RESULTS

In this section we present numerical results for two different classes of walks. The first class is the *Gillis random walk* [33], which is a random walk on the integer lattice, starting from  $x_0 = 0$ , with nontrivial jump probabilities; depending on the position of the walker, the probabilities of jumping from site  $j$  to site  $j'$  are given by the following rules:

$$p(j', j) = \frac{1}{2} \left(1 - \frac{\epsilon}{j}\right) \delta_{j', j+1} + \frac{1}{2} \left(1 + \frac{\epsilon}{j}\right) \delta_{j', j-1}, \tag{46}$$

for  $j \neq 0$ , and

$$p(j', 0) = \frac{1}{2} \delta_{j', 1} + \frac{1}{2} \delta_{j', -1} \tag{47}$$

for  $j = 0$ , where  $-1 < \epsilon < 1$  is a real parameter. For  $\epsilon > 0$  there is a bias towards the origin, while for  $\epsilon < 0$  there is a bias away from it, which in both cases decreases with the

distance from the origin. It can be shown that the random walk is recurrent only for  $\epsilon \geq -\frac{1}{2}$  [26,33], so we will consider this range only. We point out that this is one of the few examples of nonhomogeneous random walks for which exact analytical results are given. Indeed, Gillis proved that the generating function  $P(z)$  reads

$$P(z) = \frac{{}_2F_1\left(\frac{1}{2}\epsilon + 1, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z^2\right)}{{}_2F_1\left(\frac{1}{2}\epsilon, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z^2\right)}, \tag{48}$$

where  ${}_2F_1(a, b, c; z)$  is the Gaussian hypergeometric function [34].

The second class of random walks we want to consider is the *averaged Lévy-Lorentz gas* (ALL), which was presented in [35] and [36] in two different versions. This model is closely related to the well-known Lévy-Lorentz gas, quite extensively studied in the literature [37–42]. The ALL consists in a generalization of the correlated random walk [43]: A particle starts from  $x_0 = 0$  choosing with equal probability the initial direction of motion and performing only nearest-neighbor jumps. After each jump the walker can reverse the direction of motion with probability  $\tau$ , which, depending on the value of a real parameter  $\alpha$ , can assume a nontrivial position-dependent behavior: In the first version  $\tau$  decays as a power law with the distance from the origin:

$$\tau(j) \sim |j|^{-(1-\alpha)}, \quad 0 < \alpha < 1. \tag{49}$$

In the second version, instead, it is the transmission probability  $t = 1 - \tau$ , i.e, the probability of preserving the direction of motion, that decays with the same power law. In both cases at  $j = 0$  the reflection probability is  $\tau(0) = \frac{1}{2}$ . It is possible to derive the long-time properties of the model by using appropriate continuum limits, which lead to the following diffusion equation for the evolution of the PDF of the process:

$$\frac{\partial P(x, t)}{\partial t} = \frac{\partial}{\partial x} \left[ D_\alpha(x) \frac{\partial P(x, t)}{\partial x} \right]. \tag{50}$$

Here  $D_\alpha(x)$  is a position-dependent diffusion coefficient, whose behavior is related to the form of the reflection probability. Equation (50) corresponds to a Langevin equation interpreted following Hänggi-Klimontovich (isothermal interpretation). For a discussion on how different interpretations (Itô, Stratonovich, or Hänggi-Klimontovich) affect the statistical properties of the system (see ref. [18]). Here we only point out that one can compute the solution [18,44] to get the asymptotic growth of the mean square displacement for both versions of the model: In the first case,  $\langle x_t^2 \rangle \sim t^{\frac{2}{1+\alpha}}$ , hence transport is superdiffusive and we will refer to this as the *superdiffusive* version; in the second case,  $\langle x_t^2 \rangle \sim t^{\frac{2}{3-\alpha}}$ , transport is subdiffusive, and we will name this the *subdiffusive* version. We can obtain in both cases the asymptotic behavior of the probability density function at  $x = 0$ , which reads

$$P(0, t) \sim \begin{cases} t^{-1/(1+\alpha)} & \text{superdiffusive} \\ t^{-1/(3-\alpha)} & \text{subdiffusive.} \end{cases} \tag{51}$$

From the results in Eqs. (48) and (51) we are able to compute the parameter  $\rho$  for both models, yielding the shape of the Lamperti distribution, the form of the Mittag-Leffler distribution, and the asymptotic decay of the survival probability.

**A. Evaluation of the Lamperti parameter**

In order to obtain the analytical predictions to compare with the simulation results, all we need to do is to evaluate the Lamperti parameter  $\rho$ .

For the Gillis random walk we use the analytical result regarding the generating function [Eq. (48)]. By using the properties of the hypergeometric function [34], we can rewrite  $P(z)$  in the form given in Eq. (16), obtaining (see Appendix C)

$$\rho = \begin{cases} 0 & \text{for } \epsilon = -\frac{1}{2} \\ \frac{1}{2} + \epsilon & \text{for } -\frac{1}{2} < \epsilon < \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} \leq \epsilon < 1. \end{cases} \quad (52)$$

For the averaged Lévy-Lorentz gas we evaluate  $\rho$  by using the long time asymptotics of the probability of occupying the origin, which can be computed performing a continuum limit. As we reported in Eq. (51), it can be shown that the probability  $P_n$  decays as [35,36]

$$P_n \sim \begin{cases} n^{-1/(1+\alpha)} & \text{superdiffusive} \\ n^{-1/(3-\alpha)} & \text{subdiffusive,} \end{cases} \quad (53)$$

where  $0 < \alpha < 1$ . Since the exponent is connected to  $\rho$  [see Eq. (20)], we immediately get

$$\rho = \begin{cases} 1 - \frac{1}{1+\alpha} & \text{superdiffusive} \\ 1 - \frac{1}{3-\alpha} & \text{subdiffusive.} \end{cases} \quad (54)$$

**B. Occupation time of the positive axis**

Here we provide the results of simulations regarding the occupation time of the set  $A$  for the Gillis random walk and both versions of the averaged Lévy-Lorentz gas.

For the Gillis random walk we can recognize two different behaviors. When  $\epsilon < 0$  there is a bias away from the origin, and the distribution of the occupation time is represented by a U-shaped curve, meaning that the particle most likely spends all the time in one of the two sets. When  $\epsilon > 0$  there is a bias towards the origin, so we expect a higher contribution from walks which spend an equal amount of time in the two sets. This is confirmed by the plots in Fig. 1, where we consider the cases  $\epsilon = -0.2$  and  $\epsilon = 0.2$ .

When the bias towards the origin becomes sufficiently strong, i.e., for  $\epsilon \geq \frac{1}{2}$ , the outer values of the distribution cease to be the most probable. The process enters in an ergodic regime where the fraction of time spent in a set converges to its expected value, which in our case is  $\eta = \frac{1}{2}$ . In other words, the distribution is a Dirac delta function centered around  $\eta$ . Figure 2 shows the behavior of the distribution as the number of steps grows, for  $\epsilon = 0.8$ , confirming the convergence to a Dirac delta distribution.

For the averaged Lévy-Lorentz gas, the behavior of the distribution depends on which version of the model we are considering. In the superdiffusive case the reflection probability decays as a power law with the distance from the origin,  $\tau(j) \sim |j|^{-(1-\alpha)}$ , therefore a particle tends to preserve its direction of motion as the distance from the starting point increases. As  $\alpha$  varies in  $(0, 1)$ , the Lamperti parameter varies in  $(0, \frac{1}{2})$  [see Eq. (53)]. In the subdiffusive case instead the reflection probability converges to 1 as the distance from the origin increases, with the transmission coefficient decaying

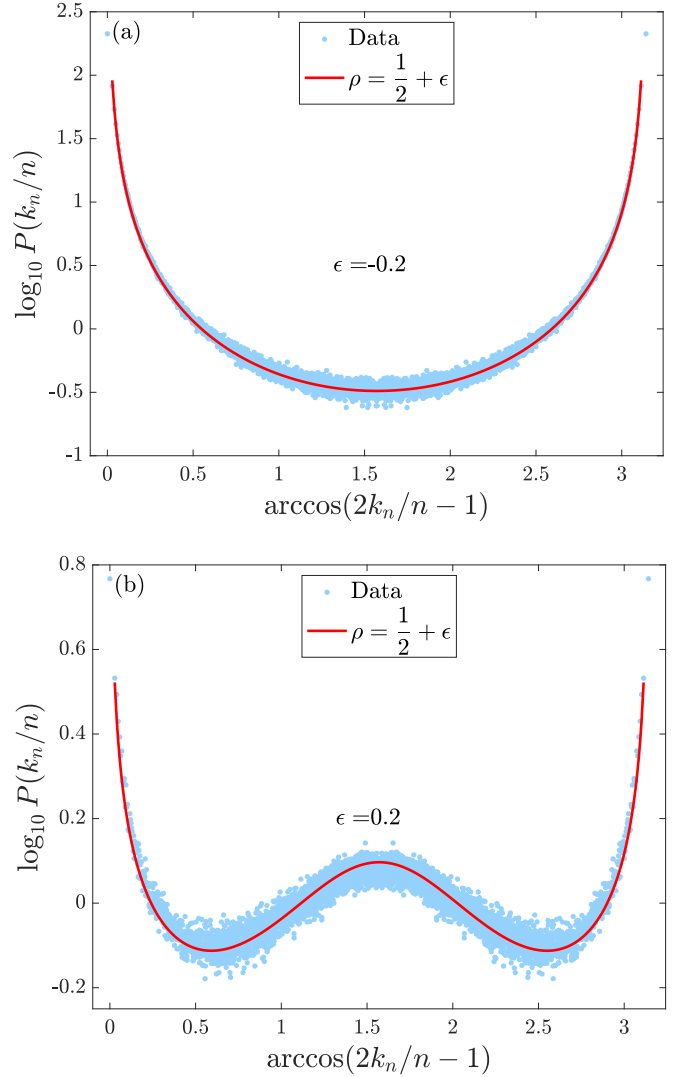


FIG. 1. Distribution of the fraction of time spent in the positive axis for the Gillis random walk, in semilogarithmic scale. To enhance readability of the outer values, it has been performed the transformation  $x \rightarrow \arccos(2x - 1)$  on the  $x$  axis. The (light blue) dots represent the simulation results, the (red) line the theoretical prediction. (a) The case  $\epsilon = -0.2$ ; (b)  $\epsilon = 0.2$ . In both cases the results are obtained simulating  $10^6$  walks of  $10^4$  steps.

as a power law. The Lamperti parameter is in the range  $\frac{1}{2} < \rho < \frac{2}{3}$ . The behavior of both models is presented in Fig. 3, for  $\alpha = 0.7$  in the superdiffusive case and  $\alpha = 0.3$  in the subdiffusive one. We point out that for both versions of the model we never enter the ergodic regime, as  $\rho \neq 1$  for any value of  $\alpha$ .

**C. Occupation time of the origin**

As we have already shown, the distribution of the occupation time of the origin follows a Mittag-Leffler distribution of the same parameter characterizing the Lamperti distribution. We consider the random variable,

$$\xi = \lim_{n \rightarrow \infty} \frac{1}{\Gamma(1 + \rho)} \frac{M_n}{\langle M_n \rangle}. \quad (55)$$

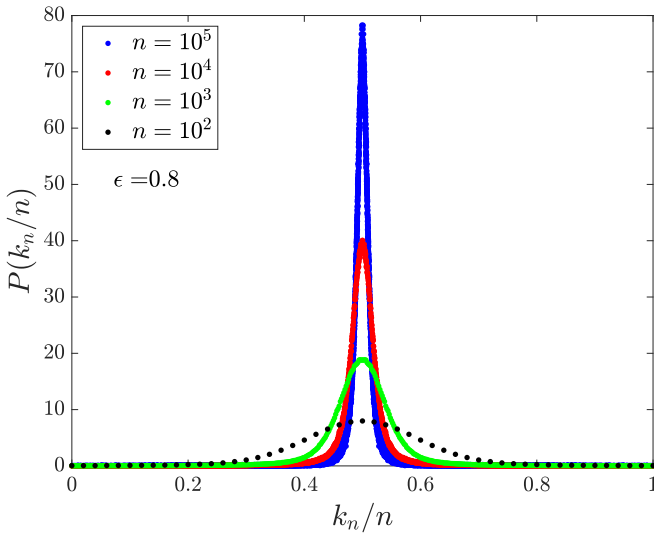


FIG. 2. Distribution of the fraction of time spent in the positive axis for the Gillis random walk, ergodic case, with  $\epsilon = 0.8$ . Data are obtained simulating  $10^6$  walks of different numbers of steps. As the maximum number of steps grows, the distribution converges to a Dirac delta, centered around  $k_n/n = 1/2$ .

Once again the Gillis random walk is the model displaying the richest behavior. As shown in Fig. 4, for  $\epsilon < 0$  the distribution of  $\xi$  is monotonically decreasing, reflecting the fact that the particle is biased away from the origin. Indeed, the walks that do not return to the starting point have the highest probability. For  $\epsilon > 0$ , instead, the bias is towards the origin, therefore the probability of returning increases. The shape of the distribution is quite different, and we have a pronounced peak close to  $\xi_0 = 1$ . For values  $\epsilon \geq \frac{1}{2}$  we enter in the ergodic regime and the distribution converges to a Dirac delta function centered around  $\xi_0 = 1$  (Fig. 5).

We can recognize a similar behavior for the averaged Lévy-Lorentz gas, as shown in Fig. 6. Here the shape of the distribution depends on which version of the model is considered: For the superdiffusive case we have a monotonically decreasing curve, while for the subdiffusive one the distribution presents a peak close to  $\xi_0 = 1$ . The values of  $\alpha$  chosen for the two systems are  $\alpha = 0.7$  for the superdiffusive version and  $\alpha = 0.3$  for the subdiffusive one.

**D. Decay of the survival and persistence probabilities**

For both models we also provide simulations regarding the asymptotic decay of the survival and persistence probabilities. As we have seen, we expect both quantities to decay as  $n^{-\rho}$ , where  $\rho$  depends on a parameter characterizing the model ( $\epsilon$  for Gillis,  $\alpha$  for Lévy-Lorentz). We confirm our prediction by plotting the exponent of the asymptotic decay of  $Q_n$  and  $U_n$ , obtained from simulations, versus the characteristic parameter.

For the Gillis random walk we have good agreement between the two computed exponents and the theoretical values (Fig. 7). We point out that  $\epsilon$  is taken in the range  $(-\frac{1}{2}, \frac{1}{2})$ , so that  $0 < \rho < 1$ . We observe that the agreement gets worse when  $\epsilon$  gets closer to the boundaries of the considered

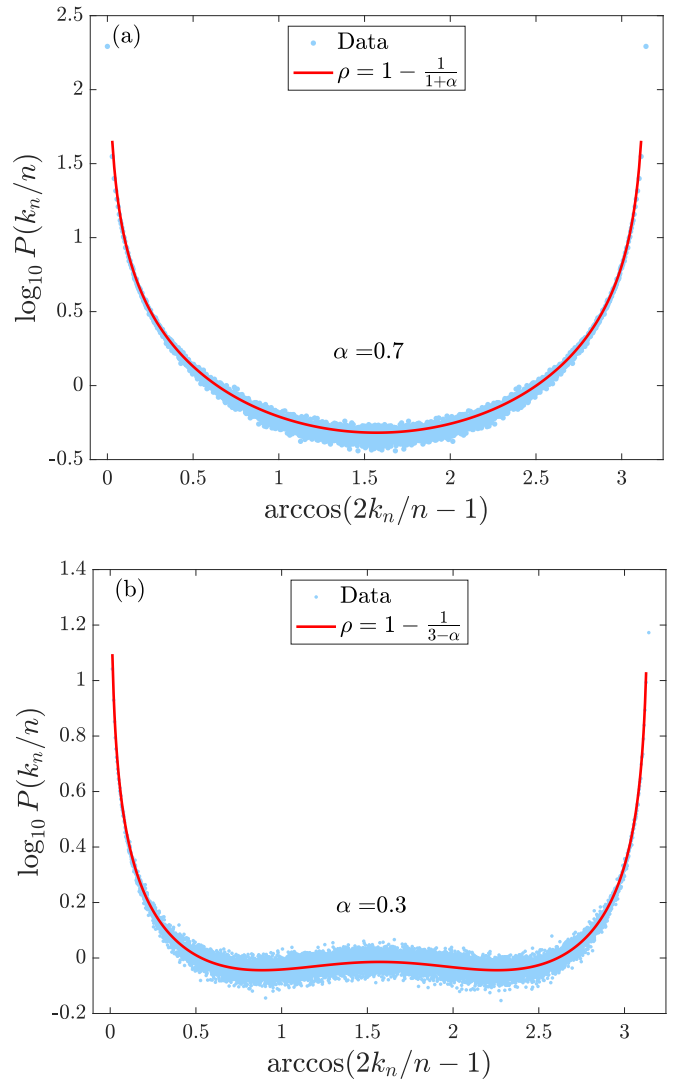


FIG. 3. Distribution of the fraction of time spent in the positive axis for the averaged Lévy-Lorentz gas, in semilogarithmic scale. To enhance readability of the outer values, it has been performed the transformation  $x \rightarrow \arccos(2x - 1)$  on the  $x$  axis. The (light blue) dots represent the simulation results, the (red) line the theoretical prediction. (a) The superdiffusive case, with  $\alpha = 0.7$ . Data are obtained simulating  $10^6$  walks of  $10^4$  steps. (b) The subdiffusive version, with  $\alpha = 0.3$ . In this case data are obtained simulating  $10^7$  walks of  $10^5$  steps.

interval: We can explain this fact by considering that as  $\epsilon \rightarrow -\frac{1}{2}$  convergence to the theoretical values becomes slower, while in the opposite case,  $\epsilon \rightarrow \frac{1}{2}$ , the system is getting closer to the regime  $\rho = 1$ , where  $Q_n$  and  $U_n$  are not guaranteed to decay in the same way.

For the superdiffusive averaged Lévy-Lorentz gas we have good agreement when  $\alpha \geq 0.4$ , while for lower values of the parameter we observe a non-negligible difference between the two computed exponents. However, we point out that this is due to the fact that the continuum limit used to describe the long time properties of the system becomes effective after a preasymptotic regime, which depends on  $\alpha$ , and the diffusive asymptotic regime is not yet captured at the number of steps

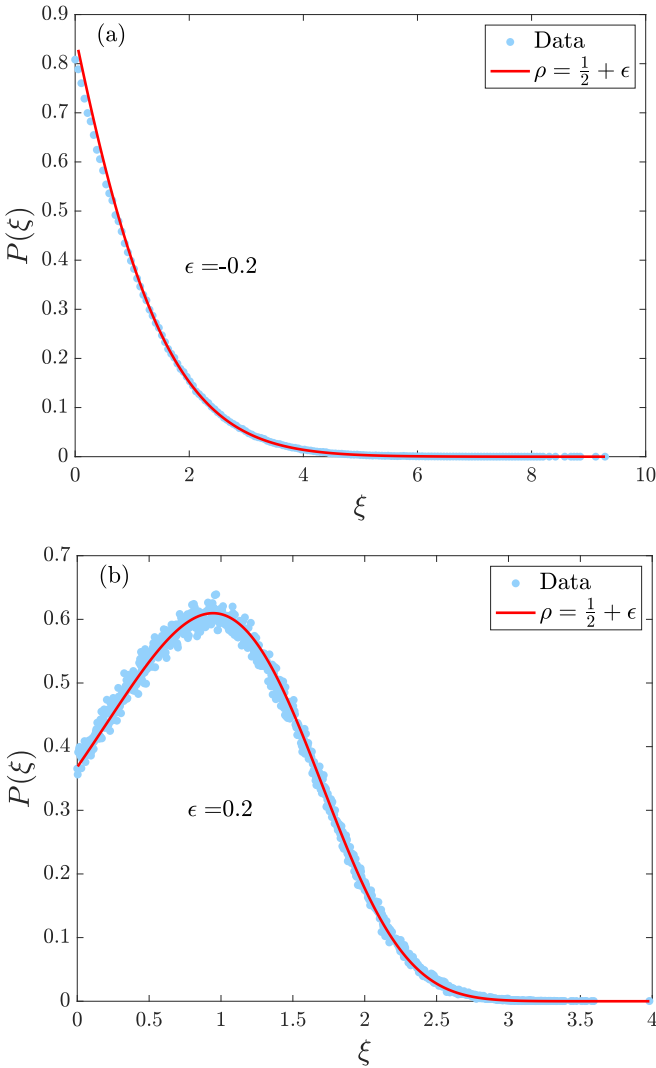


FIG. 4. Distribution of the random variable  $\xi$  representing the rescaled number of steps in which the process occupies the origin. The (light blue) dots represent the simulation results, the (red) line the theoretical prediction. (a) The case  $\epsilon = -0.2$ ; (b)  $\epsilon = 0.2$ . In both cases the results are obtained simulating  $10^6$  walks of  $10^4$  steps.

of our simulations. Indeed, we observed in [35] the same discrepancies, in the same range of  $\alpha$ , in the evaluation of the moments. For the subdiffusive version instead the difficulties to capture the asymptotic regime may be traced back to the fact that in order to observe cleanly the decay of the quantities of interest we need a larger number of steps with respect to the superdiffusive version. However, for both versions of the model we have in general a good agreement with the theoretical predictions (Fig. 8).

**E. Comparison of different systems with the same Lamperti parameter**

From the discussion made so far it should be clear that the Lamperti parameter  $\rho$ , characterizing the distributions of the observables we have considered in this paper, only depends on a local property of the PDF of the process, namely the probability  $P_n$  of occupying the origin at time  $n$ . It can

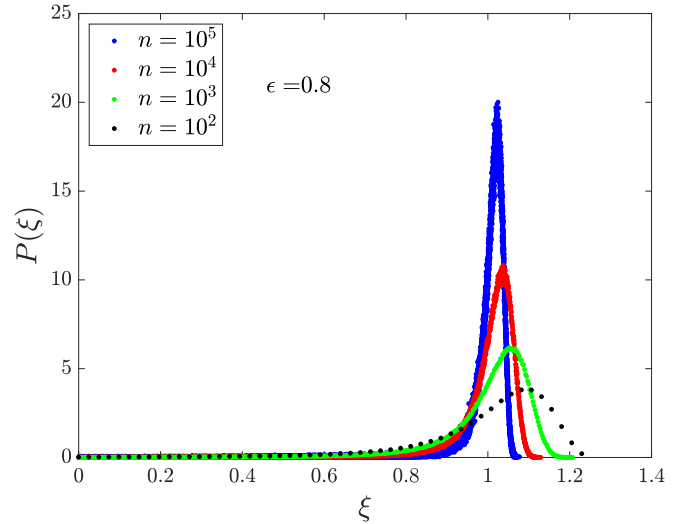


FIG. 5. Distribution of the random variable  $\xi$  representing the rescaled number of steps in which the process occupies the origin, ergodic case, with  $\epsilon = 0.8$ . Data are obtained simulating  $10^6$  walks of different numbers of steps. As the maximum number of steps grows, the distribution converges to a Dirac delta, centered around  $\xi_0 = 1$ .

happen that two stochastic processes are described by two different sets of evolution laws, but share the same asymptotic power-law decay for the distribution of the occupation time of the origin, i.e., the  $P_n$  decay with the same exponent. As a consequence, the distributions of the occupation time of the positive axis and the number of returns to the origin will be the same.

In order to show this, we compare the two distributions for the Gillis random walk and both versions of the averaged Lévy-Lorentz gas. For the latter system we consider the values of  $\alpha$  already chosen in the previous sections, viz.  $\alpha = 0.7$  for the superdiffusive version and  $\alpha = 0.3$  for the subdiffusive one. The two corresponding values of the Lamperti parameter are  $\rho = \frac{7}{17}$  (superdiffusive) and  $\rho = \frac{17}{27}$  (subdiffusive), which are obtained in the case of the Gillis random walk for  $\epsilon = -0.0882$  and  $\epsilon = 0.1296$ , respectively. The results are presented in Figs. 9 and 10. In both cases the simulations agree with the theoretical predictions.

**VII. CONCLUSIONS AND DISCUSSION**

We have shown that for a general class of stochastic processes there is a deep connection between the statistics of the occupation times, the number of visits at the origin, and the survival probability. The distributions of these observables can be characterized by the same exponent, which is related to the asymptotic power-law decay of the probability of occupying the origin. However, we remark that this connection holds true for infinite systems, and may be no longer valid when the dynamics is confined on a finite domain, e.g., by the presence of reflecting walls or a force field. In such cases we expect the process to reach an equilibrium distribution, so that the probability of occupying any point in the domain is time independent. Nevertheless, studies on the statistics of occupation times have shown that the exponent characterizing the Lamperti law is not related to the steady state. Examples



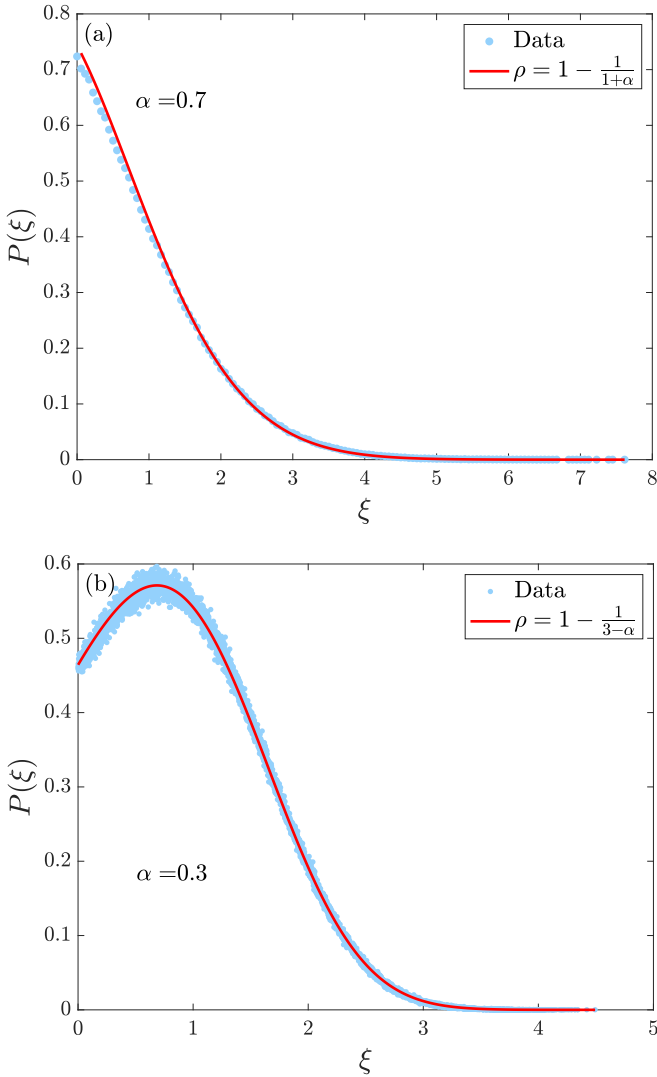


FIG. 6. Distribution of the random variable  $\xi$  representing the rescaled number of steps in which the process occupies the origin. The (light blue) dots represent the simulation results, the (red) line the theoretical prediction. (a) The superdiffusive case, with  $\alpha = 0.7$ . The results are obtained simulating  $10^6$  walks of  $10^4$  steps. (b) Corresponds to the subdiffusive version, with  $\alpha = 0.3$ . Data are obtained simulating  $10^7$  walks of  $10^5$  steps.

can be found in the contexts of continuous-time random walks [45], fractional diffusion [46], and in the quenched trap model [47], to cite a few. All these cases show that the correct characterization of the Lamperti distribution is possible by evaluating the first-passage exponent: In this sense, the relation between the first-passage time distribution and the Lamperti and Mittag-Leffler distributions is valid in a more general setting.

For infinite systems, we point out that the results of this paper are also associated with infinite ergodic theory. In particular, let us consider the Darling-Kac theorem, that we used in Sec. IV to obtain the statistics of the occupation time of the origin, in its continuous-time version [29]. The theorem first requires that, for a given non-negative and integrable

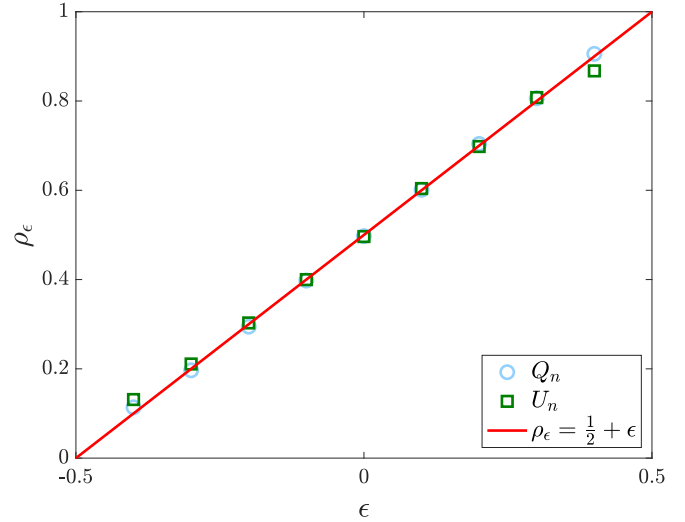


FIG. 7. Exponents of the asymptotic power-law decay of the persistence and survival probabilities for the Gillis random walk. Data are obtained simulating  $10^7$  walks of  $10^5$  steps. The (green) squares represent the persistence probability, while the (light blue) circles refer to the survival probability.

function  $V(x)$ , one has

$$\lim_{s \rightarrow 0} \frac{1}{\pi(s)} \int P_s(x|x_0)V(x)dx = c, \quad (56)$$

where  $c$  is a positive constant,  $P_s(x|x_0)$  is the Laplace transform from  $t$  to  $s$  of the probability of arriving at  $x$  starting from  $x_0$  in time  $t$ , and  $\pi(s)$  is a function such that  $\pi(s) \rightarrow \infty$  as  $s \rightarrow 0$ . Now suppose that we have

$$\lim_{s \rightarrow 0} \frac{P_s(x|x_0)}{\pi(s)} = \mathcal{I}_\infty(x), \quad (57)$$

where  $\mathcal{I}_\infty(x)$  is called, in the language of infinite ergodic theory, the infinite density [18], since it is not normalizable. Note that in this case, if  $V(x)$  is measurable with respect to the infinite density, the condition given in Eq. (56) is satisfied. Therefore, if  $\pi(s) = s^{-\rho}H(1/s)$ , with  $H(u)$  slowly varying, the Darling-Kac theorem states that the random variable,

$$\xi = \lim_{t \rightarrow \infty} \frac{1}{c\pi(1/t)} \int_0^t V(x(\tau))d\tau \quad (58)$$

$$= \lim_{t \rightarrow \infty} \frac{1}{ct^\rho H(t)} \int_0^t V(x(\tau))d\tau, \quad (59)$$

follows a Mittag-Leffler distribution of order  $\rho$ . Now we observe that using Eq. (57) we can say

$$P_s(x|x_0) \sim \pi(s)\mathcal{I}_\infty(x), \quad (60)$$

and therefore for the ensemble average of  $V(x)$  we have

$$\langle V_s \rangle = \int P_s(x|x_0)V(x)dx \sim c\pi(s). \quad (61)$$

In the case  $\pi(s) = s^{-\rho}H(1/s)$ , by using the Tauberian theorem [28] we find that the ensemble average in the long-time limit behaves as

$$\langle V_t \rangle \sim \frac{c}{\Gamma(\rho)} t^{\rho-1} H(t), \quad (62)$$

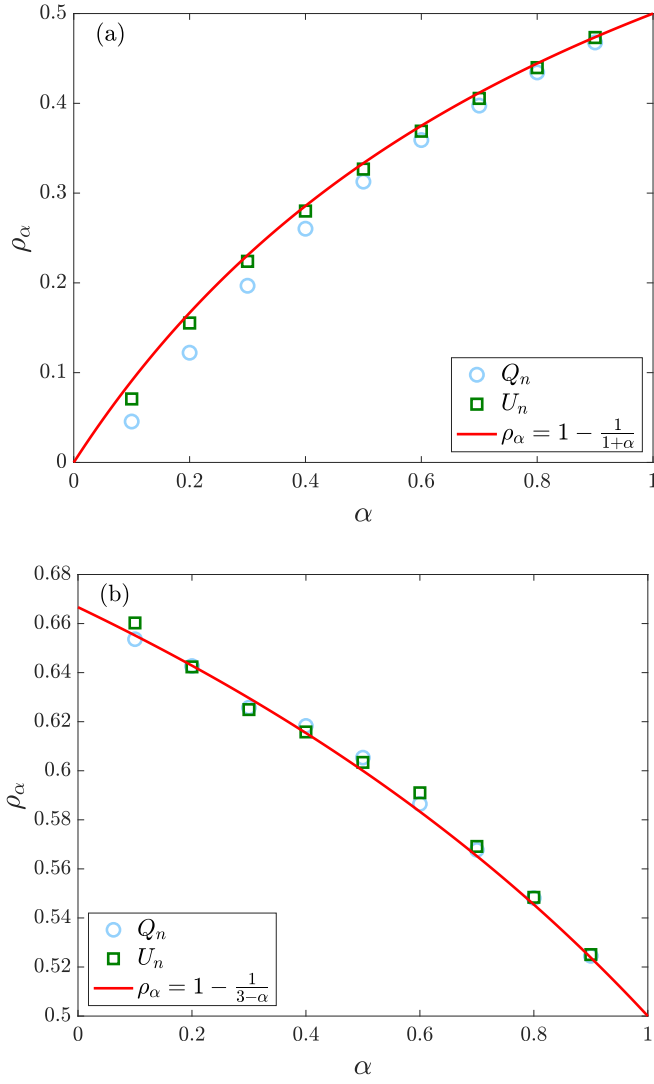


FIG. 8. Exponents of the asymptotic power-law decay of the persistence and survival probabilities for the averaged Lévy-Lorentz gas. (a) The superdiffusive version; (b) the subdiffusive one. In both cases data are obtained simulating  $10^7$  walks of  $10^5$  steps. The (green) squares represent the persistence probability, while the (light blue) circles refer to the survival probability.

and therefore

$$\xi = \lim_{t \rightarrow \infty} \frac{1}{\Gamma(\rho)} \frac{\overline{V}_t}{\langle V_t \rangle}, \quad (63)$$

where  $\overline{V}_t$  indicates the time average of  $V(x)$  over a single realization. Such a ratio is a random variable distributed according to a Mittag-Leffler of order  $\rho$ . This is the main difference with standard ergodic theory, where instead time averages converge to ensemble averages, and hence  $\xi$  is expected to be distributed according to a Dirac delta function centered around  $\xi_0 = 1$ . Now the important point is that the scaling function  $\pi(s)$ , which determines the distribution of  $\xi$ , i.e., the value of  $\rho$ , is a property of the propagator  $P_s(x|x_0)$ . In other words, for any function which is measurable with respect to the infinite density, the distribution of  $\xi$  only depends on the long-time properties of the propagator. Therefore, it is

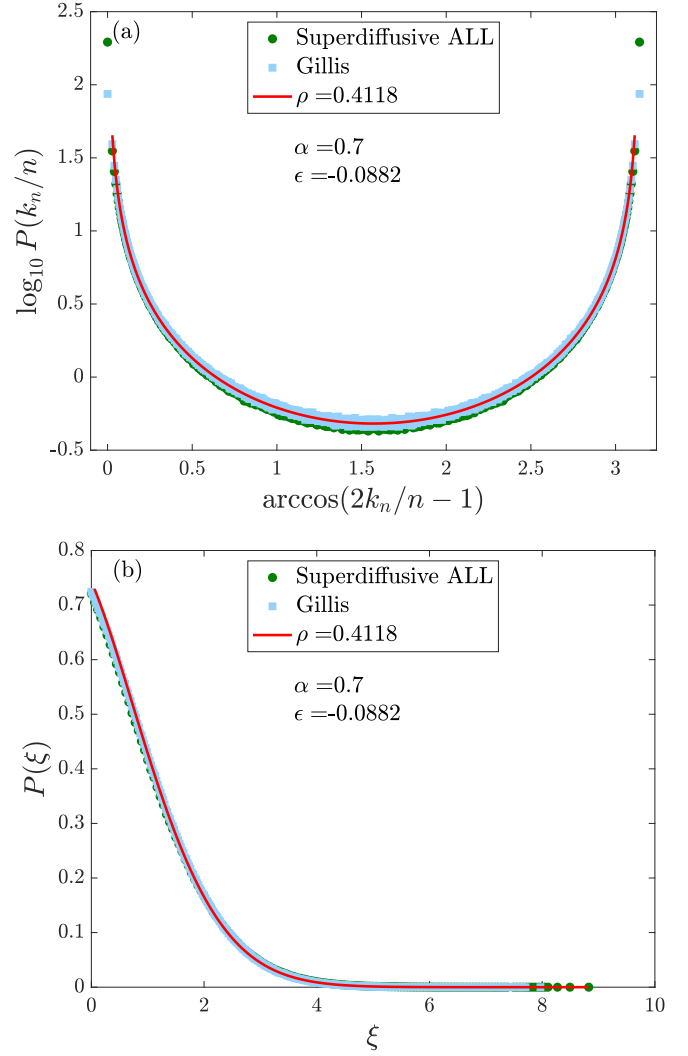


FIG. 9. Distributions of the occupation time (a) and the number of returns (b) for the averaged Lévy-Lorentz gas, superdiffusive version (green points), and the Gillis random walk (light blue squares), compared to the theoretic result. The values of the corresponding parameters are  $\alpha = 0.7$  and  $\epsilon = -0.0882$ , which in both cases yield  $\rho = 7/17$ . For both systems we considered  $10^7$  walks evolved for  $10^4$  steps.

possible to determine the distribution by just evaluating the long-time behavior of  $P(x, t)$  in a given set, as we have done in the paper by considering the probability of occupying the origin. However, we point out that if we focus on a point off the origin, the evaluation of  $\rho$  is not sufficient to determine the Lamperti distribution, since it is also necessary to determine the asymmetry parameter  $\eta$  [Eq. (3)].

### ACKNOWLEDGMENTS

The authors are grateful to the referees for their comments which helped improve the paper and acknowledge partial support from PRIN Research Project No. 2017S35EHN “Regular and stochastic behavior in dynamical systems” of the Italian Ministry of Education, University and Research (MIUR).

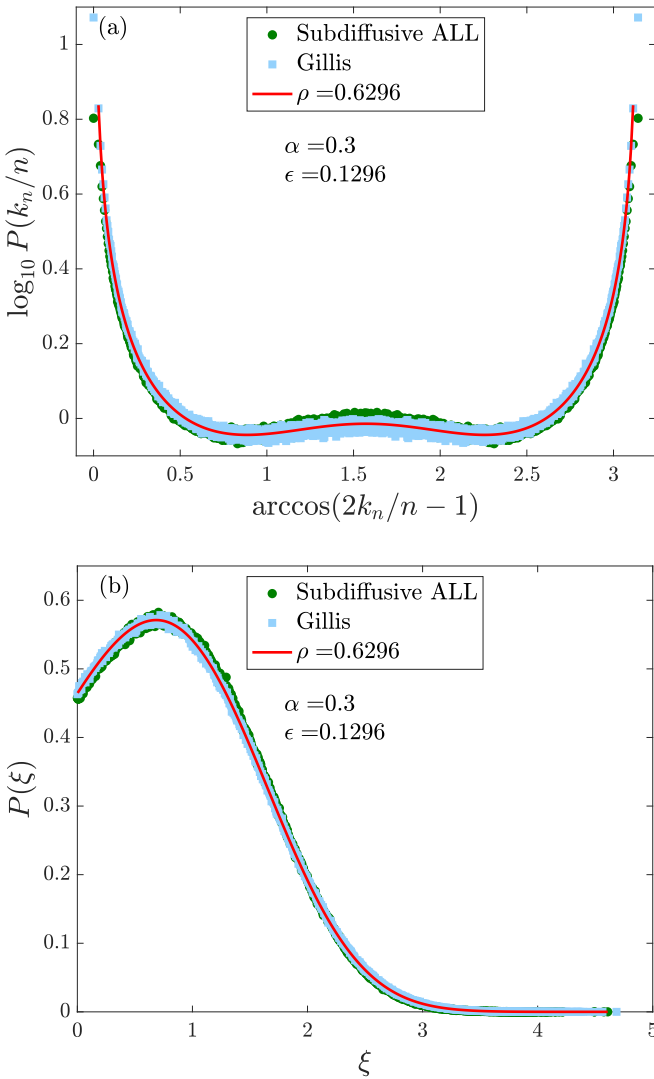


FIG. 10. Distributions of the occupation time (a) and the number of returns (b) for the averaged Lévy-Lorentz gas, subdiffusive version (green points), and the Gillis random walk (light blue squares), compared to the theoretic result. The values of the corresponding parameters are  $\alpha = 0.3$  and  $\epsilon = 0.1296$ , which in both cases yield  $\rho = 17/27$ . For both systems we considered  $10^7$  walks evolved for  $10^4$  steps.

#### APPENDIX A: MEANING OF THE RANDOM VARIABLE $T_n$

We consider the random variable,

$$T_n \equiv \frac{1}{H(n)n^\rho} \sum_{m=0}^n \delta_{x_m,0}. \quad (\text{A1})$$

The sum,

$$M_n = \sum_{m=0}^n \delta_{x_m,0}, \quad (\text{A2})$$

clearly represents the number of times the random walk has visited the origin up to time  $n$ , while it is possible to show that the denominator  $n^\rho H(n)$  is connected to the asymptotics of the mean occupation time. Indeed, for  $M \geq 1$ , let us call  $\psi_n(M)$  the probability that the  $M$ th visit occurs at step  $n$ , and

$U_n$  the probability of observing no returns to the origin up to step  $n$ , with the initial conditions  $\psi_0(M) = \delta_{M,1}$  and  $U_0 = 1$ . We have

$$U_n = 1 - \sum_{m=0}^n F_m, \quad (\text{A3})$$

$$\psi_n(1) = \delta_{n,0}, \quad (\text{A4})$$

while for  $M \geq 2$  we can write the recurrence relation,

$$\psi_n(M) = \sum_{m=0}^n F_m \psi_{n-m}(M-1). \quad (\text{A5})$$

From Eqs. (A3), (A4), and (A5) we can compute the generating functions,

$$U(z) = \frac{1 - F(z)}{1 - z}, \quad (\text{A6})$$

$$\psi_z(M) = [F(z)]^{M-1}. \quad (\text{A7})$$

Now, the probability  $\phi_n(M)$  of  $M$  visits in  $n$  steps is equal to the probability that the  $M$ th visit has occurred at step  $k \leq n$ , and then no other visit occurs up to time  $n$ :

$$\phi_n(M) = \sum_{m=0}^n \psi_k(M) U_{n-m}, \quad (\text{A8})$$

hence its generating function reads

$$\phi_z(M) = F^{M-1}(z) \frac{1 - F(z)}{1 - z}. \quad (\text{A9})$$

The generating function of the mean number of visits is

$$\langle M(z) \rangle = \sum_{M=1}^{\infty} M \phi_z(M) = \frac{1}{1 - z} \frac{1}{1 - F(z)}, \quad (\text{A10})$$

and since we know the relation between  $F(z)$  and  $P(z)$ , Eq. (15), and the form that  $P(z)$  must assume, Eq. (16), we have

$$\langle M(z) \rangle = \frac{1}{(1 - z)^{1+\rho}} H\left(\frac{1}{1 - z}\right), \quad (\text{A11})$$

and the Tauberian theorem implies

$$\langle M_n \rangle \sim \frac{1}{\Gamma(1 + \rho)} n^\rho H(n), \quad (\text{A12})$$

which is valid for  $0 \leq \rho \leq 1$ . We conclude that the random variable  $T_n$  represents, up to a constant factor, the asymptotic value of the occupation time of the origin rescaled for its mean value:

$$T_n \sim \frac{1}{\Gamma(1 + \rho)} \frac{M_n}{\langle M_n \rangle}. \quad (\text{A13})$$

#### APPENDIX B: THE RELATION BETWEEN THE SURVIVAL AND PERSISTENCE PROBABILITIES AND THEIR ASYMPTOTIC BEHAVIOR

We consider the survival probability in the set  $A$ . Define

$$F_n = \Pr\{x_1 \neq 0, x_2 \neq 0, \dots, x_n = 0 | x_0 = 0\}, \quad (\text{B1})$$

$$Q_n = \Pr\{x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 | x_0 = 0\}, \quad (\text{B2})$$

$$U_n = \Pr\{x_1 \neq 0, x_2 \neq 0, \dots, x_n \neq 0 | x_0 = 0\}, \quad (\text{B3})$$

with the initial conditions  $F_0 = 0$ ,  $Q_0 = 1$ , and  $U_0 = 1$ , and the generating functions,

$$F(z) = \sum_{n=1}^{\infty} F_n z^n, \tag{B4}$$

$$Q(z) = \sum_{n=0}^{\infty} Q_n z^n, \tag{B5}$$

$$U(z) = \sum_{n=0}^{\infty} U_n z^n. \tag{B6}$$

It is easy to see that if the process is symmetric with respect to the two sets, the following relation holds:

$$2Q_n = \delta_{n,0} + U_n + \sum_{m=1}^n F_m Q_{n-m}. \tag{B7}$$

By passing to the generating function we get

$$2Q(z) = 1 + U(z) + F(z)Q(z), \tag{B8}$$

and by using Eq. (A6) in Appendix A, after some algebra we obtain

$$Q(z) = \frac{1 + U(z)}{1 + (1 - z)U(z)}. \tag{B9}$$

To show that  $Q(z)$  and  $U(z)$  have the same  $z \rightarrow 1$  behavior, we use a result by Karamata [48]: If  $L(x)$  is a slowly varying function, then for any  $\gamma > 0$ ,

$$\lim_{x \rightarrow \infty} x^{-\gamma} L(x) = 0, \tag{B10}$$

$$\lim_{x \rightarrow \infty} x^\gamma L(x) = \infty. \tag{B11}$$

We showed in the main text, Eq. (43), that  $U(z)$  is of the form,

$$U(z) = \frac{1}{(1 - z)^{1-\rho}} L\left(\frac{1}{1 - z}\right), \tag{B12}$$

therefore, as  $z \rightarrow 1$ ,  $U(z)$  diverges and  $(1 - z)U(z)$  converges to 0. For  $\rho = 0$  we still have the divergence of  $U(z)$ , but we cannot use the previous result by Karamata for  $(1 - z)U(z)$ , because

$$(1 - z)U(z) = L\left(\frac{1}{1 - z}\right). \tag{B13}$$

However, since in this case,

$$F(z) = 1 - L\left(\frac{1}{1 - z}\right), \tag{B14}$$

and recurrence implies  $F(z) \rightarrow 1$ , we still have  $(1 - z)U(z) \rightarrow 0$ . Hence, it follows from Eq. (B9) that  $Q(z) \sim U(z)$  for any  $0 \leq \rho < 1$ .

**APPENDIX C: EVALUATION OF THE LAMPERTI PARAMETER FOR THE GILLIS RANDOM WALK**

The strategy is to put the generating function  $P(z)$ , Eq. (48) in the main text, in the form,

$$P(z) = \frac{1}{(1 - z)^v} H\left(\frac{1}{1 - z}\right), \tag{C1}$$

where  $H(x)$  is a slowly varying function. We make use of the transformation formulas [34]:

$$\begin{aligned} {}_2F_1(a, b; c; z) &= \frac{\Gamma(c)\Gamma(c - a - b)}{\Gamma(c - a)\Gamma(c - b)} {}_2F_1(a, b; a + b - c + 1; 1 - z) \\ &+ (1 - z)^{c-a-b} \frac{\Gamma(c)\Gamma(a + b - c)}{\Gamma(a)\Gamma(b)} {}_2F_1(c - a, c - b; c - a - b + 1; 1 - z), \end{aligned} \tag{C2}$$

valid for  $c - a - b$  noninteger, while for the integer case we use

$$\begin{aligned} {}_2F_1(a, b; a + b + m; z) &= \frac{\Gamma(m)\Gamma(a + b + m)}{\Gamma(a + m)\Gamma(b + m)} \sum_{n=0}^{m-1} \frac{(a)_n (b)_n}{n!(1 - m)_n} (1 - z)^n - (z - 1)^m \frac{\Gamma(a + b + m)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{(a + m)_n (b + m)_n}{n!(n + m)!} (1 - z)^n \\ &\times [\log(1 - z) - \psi(n + 1) - \psi(n + m + 1) + \psi(a + n + m) + \psi(b + n + m)], \end{aligned} \tag{C3}$$

and

$$\begin{aligned} {}_2F_1(a, b; a + b - m; z) &= (1 - z)^{-m} \frac{\Gamma(m)\Gamma(a + b - m)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{m-1} \frac{(a - m)_n (b - m)_n}{n!(1 - m)_n} (1 - z)^n - (-1)^m \frac{\Gamma(a + b - m)}{\Gamma(a - m)\Gamma(b - m)} \\ &\times \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n!(n + m)!} (1 - z)^n [\log(1 - z) - \psi(n + 1) - \psi(n + m + 1) + \psi(a + n) + \psi(b + n)], \end{aligned} \tag{C4}$$

for  $m = 1, 2, \dots$ , or

$${}_2F_1(a, b; a + b, z) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(n!)^2} (1 - z)^n [2\psi(n + 1) - \psi(a + n) - \psi(b + n) - \log(1 - z)], \tag{C5}$$

where  $\psi(z) \equiv \frac{d}{dz} \log \Gamma(z)$  is the digamma function, and  $(z)_n \equiv \Gamma(z + n)/\Gamma(z)$  denotes the Pochhammer's symbol [34].

Now, since the generating function  $P(z)$ ,

$$P(z) = \frac{{}_2F_1(\frac{1}{2}\epsilon + 1, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z^2)}{{}_2F_1(\frac{1}{2}\epsilon, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z^2)}, \tag{C6}$$

is a function of  $z^2$ , for the sake of simplicity we consider

$$P(\sqrt{z}) \equiv \Pi(z) = \sum_{n=0}^{\infty} \varpi_n z^n = \frac{{}_2F_1(\frac{1}{2}\epsilon + 1, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z)}{{}_2F_1(\frac{1}{2}\epsilon, \frac{1}{2}\epsilon + \frac{1}{2}; 1; z)}, \tag{C7}$$

so that the  $n$ th coefficient  $\varpi_n$  corresponds to  $P_{2n}$ . It is easy to show that if  $P(z)$  is of the form,

$$P(z) = \frac{1}{(1 - z)^\rho} H\left(\frac{1}{1 - z}\right), \tag{C8}$$

then also  $\Pi(z)$  can be written as

$$\Pi(z) = \frac{1}{(1 - z)^\rho} G\left(\frac{1}{1 - z}\right), \tag{C9}$$

where  $G(x)$  is slowly varying and related to  $H(x)$  by

$$G(x) = \frac{1}{x^\rho (1 - \sqrt{1 - \frac{1}{x}})^\rho} H\left(\frac{1}{1 - \sqrt{1 - \frac{1}{x}}}\right). \tag{C10}$$

This means that the transformation does not change the exponent  $\rho$ . By using Eq. (C7) we obtain the following results.

(1) In the case  $\epsilon = -\frac{1}{2}$  we get

$$\Pi(z) = G\left(\frac{1}{1 - z}\right), \tag{C11}$$

where the slowly varying function is

$$G(x) = \frac{\sum_{n=0}^{\infty} \frac{(3/4)_n(1/4)_n}{(n!)^2} (x)^{-n} [2\psi(n + 1) - \psi(\frac{3}{4} + n) - \psi(\frac{1}{4} + n) + \log(x)]}{4 + \frac{1}{4} \sum_{n=0}^{\infty} \frac{(3/4)_n(5/4)_n}{n!(n+1)!} (x)^{-n-1} [\log(x) + \psi(n + 1) + \psi(n + 2) - \psi(\frac{3}{4} + n) - \psi(\frac{5}{4} + n)]}. \tag{C12}$$

(2) In the range  $\epsilon \in (-\frac{1}{2}, \frac{1}{2})$  the generating function has the form,

$$\Pi(z) = \frac{1}{(1 - z)^{1/2+\epsilon}} G\left(\frac{1}{1 - z}\right), \tag{C13}$$

with

$$G(x) = a_1 \frac{{}_2F_1(-\frac{1}{2}\epsilon, \frac{1}{2} - \frac{1}{2}\epsilon; \frac{1}{2} - \epsilon; \frac{1}{x}) + a_2 x^{-1/2-\epsilon} {}_2F_1(\frac{1}{2}\epsilon + 1, \frac{1}{2}\epsilon + \frac{1}{2}; \frac{3}{2} + \epsilon; \frac{1}{x})}{{}_2F_1(\frac{1}{2}\epsilon, \frac{1}{2}\epsilon + \frac{1}{2}; \frac{1}{2} + \epsilon; \frac{1}{x}) + a_3 x^{-1/2+\epsilon} {}_2F_1(1 - \frac{1}{2}\epsilon, \frac{1}{2} - \frac{1}{2}\epsilon; \frac{3}{2} - \epsilon; \frac{1}{x})}, \tag{C14}$$

where  $a_1, a_2$ , and  $a_3$  are numerical coefficients (depending on  $\epsilon$ ) which can be determined from formula (C2).

(3) For  $\epsilon = \frac{1}{2}$  we have

$$\Pi(z) = \frac{1}{1 - z} G\left(\frac{1}{1 - z}\right), \tag{C15}$$

where  $G(x)$  has the expression,

$$G(x) = \frac{4 - \frac{1}{4} \sum_{n=0}^{\infty} \frac{(5/4)_n(3/4)_n}{n!(n+1)!} (x)^{-n-1} [\log(x) + \psi(n + 1) + \psi(n + 2) - \psi(\frac{5}{4} + n) - \psi(\frac{3}{4} + n)]}{\sum_{n=0}^{\infty} \frac{(1/4)_n(3/4)_n}{(n!)^2} (x)^{-n} [2\psi(n + 1) - \psi(\frac{1}{4} + n) - \psi(\frac{3}{4} + n) + \log(x)]}. \tag{C16}$$

(4) Finally when  $\epsilon \in (\frac{1}{2}, 1)$  the generating function has the same form as the previous case,

$$\Pi(z) = \frac{1}{1 - z} G\left(\frac{1}{1 - z}\right), \tag{C17}$$

but with

$$G(x) = b_1 \frac{{}_2F_1\left(-\frac{1}{2}\epsilon, \frac{1}{2} - \frac{1}{2}\epsilon; \frac{1}{2} - \epsilon; \frac{1}{x}\right) + b_2 x^{-1/2-\epsilon} {}_2F_1\left(\frac{1}{2}\epsilon + 1, \frac{1}{2}\epsilon + \frac{1}{2}; \frac{3}{2} + \epsilon; \frac{1}{x}\right)}{{}_2F_1\left(1 - \frac{1}{2}\epsilon, \frac{1}{2} - \frac{1}{2}\epsilon; \frac{3}{2} - \epsilon; \frac{1}{x}\right) + b_3 x^{1/2-\epsilon} {}_2F_1\left(\frac{1}{2}\epsilon, \frac{1}{2}\epsilon + \frac{1}{2}; \frac{1}{2} + \epsilon; \frac{1}{x}\right)}, \quad (\text{C18})$$

where once again  $b_1$ ,  $b_2$ , and  $b_3$  can be determined from Eq. (C2).

We remark that if one wishes to go back to  $P(z)$ , it is now sufficient to recover the expression of  $H(x)$  by using Eq. (C10).

- 
- [1] D. Stauffer, *J. Phys. A: Math. Gen.* **27**, 5029 (1994).  
 [2] A. J. Bray, B. Derrida, and C. Godrèche, *Europhys. Lett.* **27**, 175 (1994).  
 [3] P. L. Krapivsky, E. Ben-Naim, and S. Redner, *Phys. Rev. E* **50**, 2474 (1994).  
 [4] S. N. Majumdar and C. Sire, *Phys. Rev. Lett.* **77**, 1420 (1996).  
 [5] B. Yurke, A. N. Pargellis, S. N. Majumdar, and C. Sire, *Phys. Rev. E* **56**, R40 (1997).  
 [6] A. J. Bray, *Phys. Rev. E* **62**, 103 (2000).  
 [7] S. N. Majumdar, A. J. Bray, S. J. Cornell, and C. Sire, *Phys. Rev. Lett.* **77**, 3704 (1996).  
 [8] K. Oerding, S. J. Cornell, and A. J. Bray, *Phys. Rev. E* **56**, R25 (1997).  
 [9] C. Godrèche and J. M. Luck, *J. Stat. Phys.* **104**, 489 (2001).  
 [10] G. Bel and E. Barkai, *J. Phys.: Condens. Matter* **17**, S4287 (2005).  
 [11] N. Korabel and E. Barkai, *J. Stat. Mech.* (2011) P05022.  
 [12] I. Dornic and C. Godrèche, *J. Phys. A: Math. Gen.* **31**, 5413 (1998).  
 [13] J. M. Drouffe and C. Godrèche, *J. Phys. A: Math. Gen.* **31**, 9801 (1998).  
 [14] Z. Toroczkai, T. J. Newman, and S. Das Sarma, *Phys. Rev. E* **60**, R1115(R) (1999).  
 [15] Y. He, S. Burov, R. Metzler, and E. Barkai, *Phys. Rev. Lett.* **101**, 058101 (2008).  
 [16] J. H. P. Schulz, E. Barkai, and R. Metzler, *Phys. Rev. X* **4**, 011028 (2014).  
 [17] R. Metzler, J. H. Jeon, A. G. Cherstvy, and E. Barkai, *Phys. Chem. Chem. Phys.* **16**, 24128 (2014).  
 [18] N. Leibovich and E. Barkai, *Phys. Rev. E* **99**, 042138 (2019).  
 [19] M. Dentz, A. Cortis, H. Scher, and B. Berkowitz, *Adv. Water Resources* **27**, 155 (2004).  
 [20] T. Kühn, T. O. Ihalainen, J. Hyväluoma, N. Dross, S. F. Willman, J. Langowski, M. Vihinen-Ranta, and J. Timonen, *PLoS ONE* **6**, e22962 (2011).  
 [21] A. Nissan and B. Berkowitz, *Phys. Rev. Lett.* **120**, 054504 (2018).  
 [22] G. Altan-Bonnet, T. Mora, and A. M. Walczak, [arXiv:1907.3891](https://arxiv.org/abs/1907.3891).  
 [23] J. Lamperti, *Trans. Am. Math. Soc.* **88**, 380 (1958).  
 [24] W. Feller, *Trans. Am. Math. Soc.* **67**, 98 (1949).  
 [25] This is the same convention used in [23].  
 [26] B. D. Hughes, *Physica A* **134**, 443 (1986).  
 [27] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2 (Wiley, New York, 1971).  
 [28] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1 (Wiley, New York, 1968).  
 [29] D. A. Darling and M. Kac, *Trans. Am. Math. Soc.* **84**, 444 (1957).  
 [30] M. Thaler and R. Zweimüller, *Probab. Theory Relat. Fields* **135**, 15 (2006).  
 [31] E. Sparre-Andersen, *Math. Scand.* **2**, 195 (1954).  
 [32] C. Godrèche, S. N. Majumdar, and G. Schehr, *J. Phys. A: Math. Theor.* **50**, 333001 (2017).  
 [33] J. Gillis, *Quart. J. Math.* **7**, 144 (1956).  
 [34] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1974).  
 [35] R. Artuso, G. Cristadoro, M. Onofri, and M. Radice, *J. Stat. Mech.* (2018) 083209.  
 [36] M. Radice, M. Onofri, R. Artuso, and G. Cristadoro, *J. Phys. A: Math. Theor.* **53**, 025701 (2019).  
 [37] E. Barkai, V. Fleurov, and J. Klafter, *Phys. Rev. E* **61**, 1164 (2000).  
 [38] R. Burioni, L. Caniparoli, and A. Vezzani, *Phys. Rev. E* **81**, 060101(R) (2010).  
 [39] A. Bianchi, G. Cristadoro, M. Lenci, and M. Ligabò, *J. Stat. Phys.* **163**, 22 (2016).  
 [40] A. Vezzani, E. Barkai, and R. Burioni, *Phys. Rev. E* **100**, 012108 (2019).  
 [41] A. Bianchi, M. Lenci, and F. Pène, *Stochastic Process. Appl.* **130**, 708 (2020).  
 [42] R. Burioni and A. Vezzani, *J. Stat. Mech.* (2020) 034005.  
 [43] E. Renshaw and R. Henderson, *J. Appl. Prob.* **18**, 403 (1981).  
 [44] S. Regev, N. Grønbech-Jensen, and O. Farago, *Phys. Rev. E* **94**, 012116 (2016).  
 [45] G. Bel and E. Barkai, *Phys. Rev. Lett.* **94**, 240602 (2005).  
 [46] E. Barkai, *J. Stat. Phys.* **123**, 883 (2006).  
 [47] S. Burov and E. Barkai, *Phys. Rev. Lett.* **98**, 250601 (2007).  
 [48] J. Karamata, *Some Theorems Concerning Slowly Varying Functions*, Technical Report No. 369 ( Mathematics Research Centre, University of Wisconsin, Madison, 1962).