

Large deviations in continuous-time random walksAdrian Pacheco-Pozo^{1,*} and Igor M. Sokolov²¹*Institut für Physik, Humboldt Universität zu Berlin, Newtonstraße 15, D-12489 Berlin, Germany*²*Institut für Physik and IRIS Adlershof, Humboldt Universität zu Berlin, Newtonstraße 15, D-12489 Berlin, Germany*

(Received 18 December 2020; accepted 19 March 2021; published 9 April 2021)

We discuss large deviation properties of continuous-time random walks (CTRWs) and present a general expression for the large deviation rate in CTRWs in terms of the corresponding rates for the distributions of steps' lengths and waiting times. In the case of Gaussian distribution of steps' lengths the general expression reduces to a sequence of two Legendre transformations applied to the cumulant generating function of waiting times. The discussion of several examples (Bernoulli and Gaussian random walks with exponentially distributed waiting times, Gaussian random walks with one-sided Lévy and Pareto-distributed waiting times) reveals interesting general properties of such large deviations.

DOI: [10.1103/PhysRevE.103.042116](https://doi.org/10.1103/PhysRevE.103.042116)**I. INTRODUCTION**

The continuous-time random walk (CTRW), introduced in physics by Montroll and Weiss [1], is a generalization of a simple random walk model, in which the steps follow inhomogeneously in time. However, the precursors of this model can be traced back as early as 1903 in actuarial mathematics [2] (see [3] for a discussion).

In the standard variant of the CTRW, times of steps follow a renewal process in which the waiting times for subsequent steps are independent and identically distributed (i.i.d.) random variables. The model is then fully defined by specifying the probability density function (PDF) of waiting times and of spacial displacements in single steps (also being i.i.d. random variables) [4]. In physics, CTRW is used to model systems showing anomalous diffusion, e.g., situations when the mean square displacement (MSD) $\langle x^2(t) \rangle$ does not grow linearly in time, as predicted by Fick's law, especially in the cases of subdiffusion, when $\langle x^2(t) \rangle \propto t^\alpha$ with $\alpha < 1$. Such anomalous subdiffusion often arises due to the lack of the first moments of the distribution of waiting times. The anomaly in MSD is accompanied by a non-Gaussian shape of PDF of displacements during a given time t . This line of modeling, following the pioneering work [5], is discussed in the review articles [6,7], with some newer developments discussed in chapters of collective monographs [8,9]. Many recent works have reported a weaker type of anomaly, in which the time-dependence of the MSD is linear, but the form of the PDF is pronouncedly non-Gaussian, especially at shorter times. The decay of the tails of the PDFs is often exponential; see, e.g., [10–12] as well as [13–15] and references therein. In Ref. [14], Barkai and Burov showed that such exponential decay is a universal behavior of far tails in CTRW, presenting the approach based on subordination of random processes and large deviation theory. In Ref. [16], Wang, Barkai, and Burov presented explicit calculations for the case of exponential and

gamma (Erlang) distributions of waiting times also using large deviation theory.

The large deviation theory (LDT) is an important mathematical tool, which first arose from the analysis of insurability and ruin problems, and found its application also for a large class of other problems. In physics, the interest in large deviations is partly due to its intimate connection with the thermodynamical formalism, which leads to additional insights in the thermodynamical limit, and to new results [17].

In application to the problems of diffusion or anomalous diffusion [14,16] the value of the large deviation theory lies in the fact that it adequately describes the behaviors of the PDF in its tails, and also shows where these tails essentially begin, i.e., what part of the distribution is indeed reproduced by the typical approaches aiming for the adequate description of the PDF close to its mode (like the fluid limits of simple or continuous-time random walks). While the very far tails, in the asymptotic domain of LDT, may be hardly accessible to experiments, since the corresponding events might be extremely rare, knowledge of the transition from the central domain to the tail one (i.e., essentially the experimentally accessible preasymptotical wing of the distribution) might be of primary importance, especially when the properties of the walks are important for the understanding of the outcomes prone to the influence of rare fluctuations, like chemical reactions are.

In the present work we build on this approach and present a general expression for the large deviations' rate function in CTRW, which we use for discussing several examples beyond the ones considered in [14,16], where we consider both the case when the mean waiting time density (WTD) is finite and the case when it diverges. For the case of Gaussian step length distribution, as in [16], the analysis is especially simple, and the general expression reduces to a sequence of two Legendre transformations of a cumulant generating function of waiting times. We note that considerable work was done in mathematics to investigate large deviation properties of random walks (see, e.g., [18,19] and references therein) and of CTRW (see,

*Corresponding author: adrian.pacheco@physik.hu-berlin.de

e.g., the discussion of the existence and general properties of large deviation rates in Refs. [20,21]), but we concentrate here on physical examples close to the ones discussed in [14,16].

The structure of the present work is as follows: In Sec. II we conduct a preliminary discussion of the model and introduce notation used throughout the work. In Sec. III we derive the main result, which we use in the discussion of several examples, and consider the first example of a Bernoulli random walk. In Sec. IV we consider the case with Gaussian step length distribution, for which the expressions are especially simple. Here different further examples with and without mean waiting times are considered, which are enough for understanding the physics behind the large deviations in CTRW. Sec. VI concludes the work. To make the text self-contained, we present in Appendix A sketches of the proofs of some standard properties of the rate functions used in the main text, and in Appendix B the demonstration, on a physical level of rigor, of the known relation between the large deviation rates of leading and directing processes of CTRW. Appendix C gives an additional discussion of the asymptotic behavior of Legendre transformations in the case of Pareto distributions supporting intuitive arguments used in the main text.

II. PRELIMINARIES AND NOTATION

A continuous-time random walk (CTRW) is a random process which ensues due to subordination of a simple random walk (RW) process $x(n)$ to another random process $n(t)$, which gives us the number of steps done by the random walk up to time t , and is called a *directing process*, or operational time of the CTRW scheme [4]. The process $x(n)$ is called with this respect a *parent process* of the CTRW (see Refs. [22,23] for a systematic discussion and terminology). The “inverse” process $t(n)$ corresponding to the sum of independent, identically distributed waiting times in CTRW is called a *leading process*.

Let the PDF of the parent process be $p_n(x)$, and the probability of making exactly n steps up to time t be $\chi_n(t)$. Then the PDF of displacements in CTRW is [4]

$$p(x, t) = \sum_{n=0}^{\infty} p_n(x) \chi_n(t).$$

For long t , the typical number of steps is very large, and the operational time $\tau = n$ can be taken to be continuous. In this limit

$$p(x, t) = \int_0^{\infty} f(x, \tau) T(\tau, t) d\tau, \quad (1)$$

where $f(x, \tau)$ is a continuous approximation to (the fluid limit of) $p_n(x)$ and $T(\tau, t)$ is the same kind of approximation to $\chi_n(t)$. Equation (1) is called the integral formula of subordination. In what follows we will only work in this limit.

The parent and the leading processes of the CTRW scheme are processes with independent and identically distributed (i.i.d.) increments, i.e., they correspond to sums of i.i.d. random variables. Exactly such sums constitute the elementary (and elemental) class of objects studied in the theory of large deviations. Let x_n be a sequence of independent random vari-

ables, $s_n = (x_1 + x_2 + \dots + x_n)/n$ be the empirical mean of the first n elements in the sequence, and $P_n(s)$ be the probability distribution of s_n . Then, under known conditions, there exists the limit

$$\mathcal{C}(s) = - \lim_{n \rightarrow \infty} \frac{1}{n} \ln P_n(s)$$

called the rate function (RF), or Cramér’s function, of large deviations for s_n [17,24,25]. The distribution of s_n for sufficiently large n is then given, up to a normalization constant, by

$$P_n(s) \sim \exp[-n\mathcal{C}(s)].$$

When making the inverse variable transformation from s_n to the sum $S_n = x_1 + x_2 + \dots + x_n = ns_n$ of the first n random variables, one sees that its PDF takes the form

$$p_n(S) \sim \exp\left[-n\mathcal{C}\left(\frac{S}{n}\right)\right], \quad (2)$$

up to an n -dependent normalization constant. Equation (2) will be called the *large deviation form* for the PDF of the corresponding sum. To avoid ambiguity, we will denote the rate functions by script letters.

At the beginning, only two PDFs are known: the one of the single step lengths, and the one of the waiting times. Therefore, in our discussion, we will start from the rate functions of sums of independent and identically distributed random variables (step times and step lengths), and derive the properties of all other rate functions from two rate functions: $\mathcal{I}(t)$ and $\mathcal{R}(x)$, the ones for the leading process and for the parent process (simple RW), which can be readily calculated. The large deviation forms for the PDFs of the corresponding sums are then

$$f(x, \tau) \sim \exp\left[-\tau\mathcal{R}\left(\frac{x}{\tau}\right)\right] \quad (3)$$

and

$$p(t, \tau) \sim \exp\left[-\tau\mathcal{I}\left(\frac{t}{\tau}\right)\right]. \quad (4)$$

To easily manipulate the variables we have to keep the arguments of all non-power-law functions dimensionless. To do so, we introduce the natural units of length and time. As a natural unit of length one may choose the mean squared displacement in a single step, and as a natural unit of time a characteristic time of a single step. If the mean waiting time for a step exists, we will choose this mean waiting time as the characteristic time. In the cases when the mean waiting time diverges, for the purpose of comparison of different situations, one should choose a characteristic time such that the mean number of steps $\langle n(t) \rangle$ made up to time t is the same for all cases compared. Note that this requirement is fulfilled when choosing the mean waiting time, provided it exists, as a characteristic time, as above. Thus, from here on, all our quantities are dimensionless. The units can be easily restored in the final expressions when necessary.

The two derived rate functions, the ones for the directing process and the one for the whole CTRW (a function we seek to know) will be denoted by $\mathcal{T}(\tau)$ and $\mathcal{X}(x)$, respectively. The

large deviation forms for the PDFs of τ and x are

$$T(\tau, t) \sim \exp \left[-t\mathcal{T} \left(\frac{\tau}{t} \right) \right] \tag{5}$$

and

$$p(x, t) \sim \exp \left[-t\mathcal{X} \left(\frac{x}{t} \right) \right]. \tag{6}$$

Before starting our derivation, let us go through some important properties of the rate functions [17,24,25]. The rate functions for the mean of i.i.d. random variables are convex [25]; see also Appendix A. Cramér’s theorem states that the rate function $\mathcal{C}(x)$ is the convex conjugate (a Legendre-Fenchel transformation) of the cumulant generation function $\mathcal{L}(q) = \ln \langle e^{qx} \rangle$ of the distribution of x (a simple explanation is given in Ref. [24]). For the sake of brevity we will call this transformation simply a Legendre transformation, as is done, say, when transforming a Lagrange to a Hamilton function in mechanics. For strictly convex rate functions (i.e., for all cases discussed in the present work) this transformation is invertible. The inverse of a Legendre transformation is given by a Legendre transformation again (an involution property). Two other properties of rate functions, which are important for what follows, are that the rate function is nonnegative and that it vanishes identically at the value of $x = \mu$ corresponding to the mean of the distribution of the single variable. The rate function can be defined on the whole real line (e.g., for a Gaussian distribution of x), on a half-line (like it is for waiting times), or on a finite interval. Outside of its domain of definition (i.e., outside of the domain available for the initial random variables) it is taken to be “literally infinite,” so that the probability to have such values (the ones outside the domain of definition) is strictly zero.

To keep equations concise we will introduce the following nonstandard notation for the Legendre transformation: Let $g(x)$ be a convex function defined on the corresponding interval, and let a function $f(z)$ of variable z be its Legendre transformation defined as

$$g(z) = \sup_x \{zx - f(x)\}.$$

The Legendre transformation is an operator which makes a mapping $f(x) \mapsto g(z)$. Our notation will keep track of the variable’s names in the corresponding functions

$$g(z) = \widehat{\mathcal{L}}_{\{z \leftarrow x\}} f(x).$$

The direction of the arrow is connected with the order of the symbols: the operator $\widehat{\mathcal{L}}$ acts on the function of x standing to the right from it, and transforms the function f of variable x into a function g of variable z (the names of the variables are changed from right to left). Since we choose to denote the variable’s names in the notation for the operator, we do not have to put them in the functions at all (but we will).

The main result of the present article is as follows:

$$\mathcal{X}(z) = - \sup_{\xi} \left[- \frac{\mathcal{R}(z\xi) + \mathcal{I}(\xi)}{\xi} \right], \tag{7}$$

with $\mathcal{I}(z) = \widehat{\mathcal{L}}_{\{z \leftarrow q\}} \mathcal{L}(q)$ and $\mathcal{R}(z) = \widehat{\mathcal{L}}_{\{z \leftarrow q\}} \mathcal{D}(q)$ where $\mathcal{L}(q) = \langle e^{qt} \rangle = \int_0^\infty e^{qt} \psi(t) dt$ with $\psi(t)$ being the PDF of waiting times (waiting time density, WTD), and $\mathcal{D}(q) =$

$\langle e^{qx} \rangle = \int_{-\infty}^\infty e^{qx} \lambda(x) dx$ with $\lambda(x)$ being the PDF of displacements in a single step (steps’ length density, SLD). For a Gaussian distribution of single steps’ lengths discussed in [14,16], the result can be put in a closed form involving only the Legendre transformations

$$\mathcal{X}(x) = - \widehat{\mathcal{L}}_{\{-x^2/2 \leftarrow \xi\}} [\xi^{-1} \widehat{\mathcal{L}}_{\{\xi \leftarrow q\}} \mathcal{L}(q)]. \tag{8}$$

III. THE LARGE DEVIATIONS RATE FOR CTRW

Let us now derive the forms (7) and (8). Substituting the large deviation forms for PDFs of the parent and directing processes [Eqs. (3) and (5), respectively] into the integral formula of subordination, Eq. (1), one gets

$$p(x, t) \sim \int_0^\infty \exp \left[-\tau \mathcal{R} \left(\frac{x}{\tau} \right) - t\mathcal{T} \left(\frac{\tau}{t} \right) \right] d\tau.$$

On the left-hand side (l.h.s.), we introduce the large deviation form of the whole CTRW [Eq.(6)],

$$\exp \left[-t\mathcal{X} \left(\frac{x}{t} \right) \right] \sim \int_0^\infty \exp \left[-\tau \mathcal{R} \left(\frac{x}{\tau} \right) - t\mathcal{T} \left(\frac{\tau}{t} \right) \right] d\tau,$$

and change to the new variables $z = x/t$ and $\xi = t/\tau$:

$$\exp[-t\mathcal{X}(z)] \sim \int_0^\infty \frac{t}{\xi^2} \exp \left\{ -t \left[\frac{\mathcal{R}(z\xi)}{\xi} + \mathcal{T} \left(\frac{1}{\xi} \right) \right] \right\} d\xi.$$

Now, we use the relation between the RF for the leading and directing processes, which is given by

$$\mathcal{T}(\tau) = \tau \mathcal{I} \left(\frac{1}{\tau} \right); \tag{9}$$

see Ref. [26]. The proof of this relation (and the discussion of conditions under which it holds) implies a longer chain of mathematical discussions. Although this relation is known, a simple explanation on the physical level of rigor is given in Appendix B.

Using this relation one gets

$$\exp[-t\mathcal{X}(z)] \sim \int_0^\infty \frac{t}{\xi^2} \exp \left\{ -t \left[\frac{\mathcal{R}(z\xi)}{\xi} + \frac{\mathcal{I}(\xi)}{\xi} \right] \right\} d\xi.$$

This integral can then be solved by Laplace’s method [27]. Thus, we can write

$$\exp[-t\mathcal{X}(z)] \sim \exp \left\{ -t \inf_{\xi} \left[\frac{\mathcal{R}(z\xi)}{\xi} + \frac{\mathcal{I}(\xi)}{\xi} \right] \right\},$$

where the preexponential term is disregarded, since it does not contribute to the large deviations when taking the limit $\lim_{t \rightarrow \infty} \ln[p(x, t)]/t$. Equating the arguments of the exponentials, we get

$$\mathcal{X}(z) = - \sup_{\xi} \left[- \frac{\mathcal{R}(z\xi) + \mathcal{I}(\xi)}{\xi} \right],$$

which is our main result, Eq. (7). Note that the infimum was changed to the supremum of the negative of the expression, which will be useful to later relate this quantity to a Legendre transformation in the case of a Gaussian SLD. Since the large deviation rates $\mathcal{R}(x)$ and $\mathcal{I}(t)$ are readily given by the Legendre transformations of the corresponding cumulant generating functions, the corresponding supremum can be easily calculated for the cases of interest. In what follows, we

consider in some detail the case brought to our attention by the works [14,16], namely, the case of Gaussian distribution of single step lengths. For this case, the RF $\mathcal{X}(z)$ follows from the cumulant generating function of waiting times by two Legendre transformations (see below). However, first, as an example, let us analyze the simplest random walk, i.e., the Bernoulli one.

CTRW with a fixed step length

As a first example, let us consider the Bernoulli random walk, with steps of fixed length ± 1 taken with probability $1/2$, respectively, and the simplest possible leading process, a Poisson process with rate 1, i.e., with $\psi(t) = e^{-t}$. The Bernoulli random walk is essentially the first example of Refs. [24,25]. For this case

$$p_N(S) = 2^{-N} \frac{N!}{\left(\frac{N}{2} - \frac{S}{2}\right)! \left(\frac{N}{2} + \frac{S}{2}\right)!}.$$

Applying Stirling's formula we get in the first order in N

$$\ln p_N(s) = -\frac{N}{2} [(1-s) \ln(1-s) + (1+s) \ln(1+s)]$$

with $s = S/N$, so that

$$\mathcal{R}(x) = \frac{1}{2} [(1-x) \ln(1-x) + (1+x) \ln(1+x)].$$

This function is parabolic around 0 and diverges for $x \rightarrow \pm 1$ [and is "literally infinite" outside of the interval $(-1, 1)$]. The function $\mathcal{I}(\xi) = \widehat{\mathcal{L}}_{\{\xi \leftarrow q\}} \mathcal{L}(q)$ corresponding to a Poisson process can be readily calculated (and essentially is long known), and reads

$$\mathcal{I}(\xi) = \xi - 1 - \ln \xi. \quad (10)$$

The supremum in Eq. (7) for fixed z is achieved at $\xi = 1/\sqrt{1+z^2}$, and $\mathcal{X}(z)$ is readily evaluated:

$$\begin{aligned} \mathcal{X}(z) = & 1 + \sqrt{1+z^2} \left\{ -1 + \ln \sqrt{1+z^2} \right. \\ & + \frac{1}{2} \left[\left(1 - \frac{z}{\sqrt{1+z^2}}\right) \ln \left(1 - \frac{z}{\sqrt{1+z^2}}\right) \right. \\ & \left. \left. + \left(1 + \frac{z}{\sqrt{1+z^2}}\right) \ln \left(1 + \frac{z}{\sqrt{1+z^2}}\right) \right] \right\}, \quad (11) \end{aligned}$$

which is the function with a quadratic behavior close to $z = 0$, $\mathcal{X}(z) \simeq z^2/2 + O(z^4)$, and with the large- z asymptotics being $\mathcal{X}(z) \simeq |z|(-1 + \ln 2 + \ln |z|)$ (note that the function is not "cut" at a finite value of z , different from the large deviation rate of the parent process).

IV. CTRW WITH GAUSSIAN DISTRIBUTION OF STEPS LENGTHS

For the Gaussian distribution of step lengths the large deviations rate for the parent process is given by

$$\mathcal{R}(x) = \frac{x^2}{2}.$$

Introducing this result into Eq. (7), one obtains

$$\mathcal{X}(x) = -\sup_{\xi} \left[-\frac{x^2 \xi}{2} - \frac{\mathcal{I}(\xi)}{\xi} \right], \quad (12)$$

which is given by a Legendre transformation of the function $\mathcal{I}(\xi)/\xi$ taken at the value of the Legendre variable $z = -x^2/2$. Using our notation for the Legendre transformations and the fact that $\mathcal{I}(\xi) = \widehat{\mathcal{L}}_{\{\xi \leftarrow q\}} \mathcal{L}(q)$ we arrive at the final result in a closed form,

$$\mathcal{X}(x) = -\widehat{\mathcal{L}}_{\{-x^2/2 \leftarrow \xi\}} [\xi^{-1} \widehat{\mathcal{L}}_{\{\xi \leftarrow q\}} \mathcal{L}(q)],$$

Eq. (8). Now we use this formula for discussing several important examples and then make some conclusions about the behavior of the large deviation rate functions of CTRW with Gaussian step lengths' distribution for very large deviations ($x \gg 1$).

A. Exponential waiting times

Let us assume that the waiting times follow a Poisson process, with $\psi(t) = e^{-t}$ and a rate function given by Eq. (10). Now, we define the function

$$g(\xi) = \frac{\mathcal{I}(\xi)}{\xi} = 1 - \frac{1}{\xi} - \frac{\ln \xi}{\xi},$$

and perform the Legendre transformation

$$z = \frac{d}{d\xi} g(\xi) = \frac{\ln \xi}{\xi^2}$$

so that

$$\xi = \exp \left[-\frac{1}{2} W_0(-2z) \right],$$

where $W_0(\cdot)$ is the Lambert function. Then, according to Eq. (8), we change the variable to $z = -x^2/2$ and invert the sign:

$$\begin{aligned} \mathcal{X}(x) = & 1 - \exp \left[\frac{1}{2} W_0(x^2) \right] + \frac{x^2}{2} \exp \left[-\frac{1}{2} W_0(x^2) \right] \\ & + \frac{1}{2} W_0(x^2) \exp \left[\frac{1}{2} W_0(x^2) \right]. \quad (13) \end{aligned}$$

Using the properties of the Lambert function, the two limits, of small and of large x , can be found. For $x \gg 1$, $W_0(x^2) \sim \ln x^2 + \ln \ln x^2$, and $\mathcal{X}(x) \sim |x| \sqrt{2 \ln |x|}$. On the other hand, for $x \ll 1$, $W_0(x^2) \sim x^2$, and, performing a Taylor expansion around zero, one has $\mathcal{X}(x) \sim x^2/2 + O(x^4)$, corresponding to a Gaussian. Figure 1 shows a comparison between the rate functions for the Bernoulli case with exponential WTD, Eq. (11), and for the Gaussian SLD with exponential WTD, Eq. (13). One can see that, as $x \rightarrow 0$, both curves coincide and show a parabolic behavior which is a consequence of the central limit theorem (CLT). For $x > 1$ the curves diverge, with the one for the Bernoulli RW growing faster. This shows that the large deviation rate of CTRWs is sensitive to the single step length distribution (and can be used as a probe for such). This discussion reproduces the results of Ref. [16].

B. One-sided Lévy law

Now we take $\psi(t)$ to follow a one-sided Lévy law with exponent α . Its Laplace characteristic function $\tilde{\psi}(s) = \langle e^{-st} \rangle$ reads $\tilde{\psi}(s) = \int_0^\infty \psi(t) e^{-st} dt = \exp(-\sigma s^\alpha)$ with $0 < \alpha < 1$ (for $s > 0$) and σ being the scale parameter. To allow for further comparison with the behavior for the Pareto-distributed

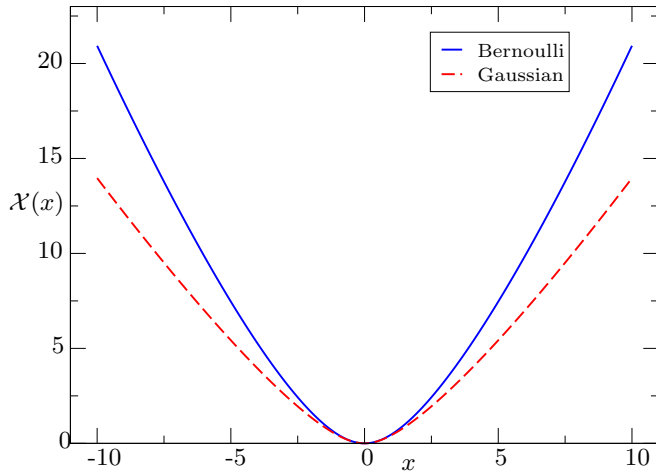


FIG. 1. A comparison of the rate functions $\mathcal{X}(x)$ for a Bernoulli CTRW and a CTRW with Gaussian distribution of step lengths, both with exponential WTD, Eq. (11). For $x \ll 1$, both curves have the same parabolic behavior, which is a consequence of the CLT; see the text for details.

waiting times, we fix this to be $\sigma = 1/\Gamma(\alpha + 1)$. This fixes the characteristic time of a step. The mean number of steps $\langle n(t) \rangle$ done during the time t is then

$$\langle n(t) \rangle = t^\alpha. \tag{14}$$

Let us denote $C_L = \Gamma(\alpha + 1)^{1/\alpha}$. Thus, $\tilde{\psi}(s) = \exp[-(s/C_L)^\alpha]$, so that $f(q) = \langle e^{qt} \rangle = \tilde{\psi}(-q)$ is given by

$$f(q) = \begin{cases} \exp[-(-q)^\alpha/C_L^\alpha], & q \leq 0, \\ +\infty, & q > 0, \end{cases}$$

and

$$\mathcal{L}(q) = \begin{cases} -(-q)^\alpha/C_L^\alpha, & q \leq 0, \\ +\infty, & q > 0. \end{cases}$$

Now, we perform the Legendre transform,

$$t = \frac{d}{dq} \mathcal{L}(q) = \frac{\alpha(-q)^{\alpha-1}}{C_L^\alpha} \rightarrow q = -\left(\frac{C_L^\alpha t}{\alpha}\right)^{\frac{1}{\alpha-1}},$$

which leads to

$$\mathcal{I}(t) = (1 - \alpha) \left(\frac{C_L t}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}.$$

Now

$$g(\xi) = \frac{\mathcal{I}(\xi)}{\xi} = (1 - \alpha) \left(\frac{\alpha}{C_L}\right)^{\frac{\alpha}{1-\alpha}} \xi^{-\frac{1}{1-\alpha}},$$

and now the second Legendre transformation can be performed:

$$z = \frac{d}{d\xi} g(\xi) = -\left(\frac{\alpha}{C_L}\right)^{\frac{\alpha}{1-\alpha}} \xi^{-\frac{2-\alpha}{1-\alpha}},$$

with

$$\xi = \left(\frac{C_L}{\alpha}\right)^{-\frac{\alpha}{2-\alpha}} (-z)^{-\frac{1-\alpha}{2-\alpha}}.$$

Finally, the rate function of displacements reads

$$\mathcal{X}(x) = (2 - \alpha) 2^{-\frac{1}{2-\alpha}} \left(\frac{\alpha}{C_L}\right)^{\frac{\alpha}{2-\alpha}} |x|^{\frac{2}{2-\alpha}}. \tag{15}$$

For $\alpha \rightarrow 1$ it tends to a parabolic RF of a Gaussian and for $\alpha \rightarrow 0$ to a wedge-like RF of ultraslow diffusion with a purely exponential PDF tails; see [4].

C. Pareto distributions

Now let us consider two other WTDs, namely the Pareto distributions of types I and II with parameter α (for the sake of brevity these will be sometimes referred to as Pareto I and Pareto II distributions). Both distributions have the same asymptotic behavior for long times, $\psi(t) \propto t^{-1-\alpha}$, but differ with respect to their short time behavior. The distributions with $0 < \alpha \leq 1$ are fat tailed and do not possess the mean; the distributions with $\alpha > 1$ do possess the mean waiting time.

To be able to compare them to each other, and to a Lévy distribution for $\alpha \in (0, 1)$, one has to fix the time scale. For $\alpha > 1$, both Pareto distributions possess a mean waiting time, so the scaling factor is this last quantity. For $0 < \alpha \leq 1$, we scale the distributions in such a way that the mean number of steps $\langle n(t) \rangle$ for a given t is the same for both distributions and exactly the same as for the Lévy case [Eq. (14)]. Thus, the Pareto type I WTD $\psi(t)$ is defined as

$$\psi(t) = \begin{cases} 0, & 0 \geq t > C_I, \\ \frac{\alpha C_I^\alpha}{t^{\alpha+1}}, & t \geq C_I, \end{cases} \tag{16}$$

where $C_I = [\Gamma(1 + \alpha)\Gamma(1 - \alpha)]^{-1/\alpha}$ if $\alpha \leq 1$, or $C_I = (\alpha - 1)/\alpha$ if $\alpha > 1$. Its Laplace transform reads

$$\tilde{\psi}(s) = \alpha C_I^\alpha s^\alpha \Gamma(-\alpha, C_I s), \tag{17}$$

with $\Gamma(x, y)$ being an upper incomplete Γ function. On the other hand, the Pareto type II WTD can be defined as

$$\psi(t) = \frac{\alpha C_{II}^\alpha}{(C_{II} + t)^{\alpha+1}}, \tag{18}$$

where $C_{II} = [\Gamma(1 + \alpha)\Gamma(1 - \alpha)]^{-1/\alpha}$ if $\alpha \leq 1$, and $C_{II} = (\alpha - 1)$ if $\alpha > 1$. Its Laplace transform reads

$$\tilde{\psi}(s) = \alpha C_{II}^\alpha s^\alpha e^{C_{II} s} \Gamma(-\alpha, C_{II} s). \tag{19}$$

Due to the presence of incomplete gamma functions, the Legendre transforms have to be performed numerically, by solving algebraic equations.

Figure 2 shows a comparison between the rate functions $\mathcal{X}(x)$ of displacements in the CTRW with Gaussian STD and the three following WTDs which do not possess mean waiting time: one-sided Lévy law with $\alpha = 0.5$, and Pareto I and II distributions, both with $\alpha = 0.5$, all with the same $\langle n(t) \rangle$ as given by Eq. (14). As in the case of the distributions with finite mean waiting times, the curves coincide in the central domain [although now they are not Gaussian but are given by Eq. (15)], but deviate for x large.

Fig. 3 shows the corresponding comparison between the Pareto type I and Pareto type II WTDs for the cases $\alpha = 1.5$ and $\alpha = 2.5$ when they do possess the mean waiting time. In this case the behavior for small x is universally parabolic, as it should be, but the RFs for Pareto type I WTDs universally

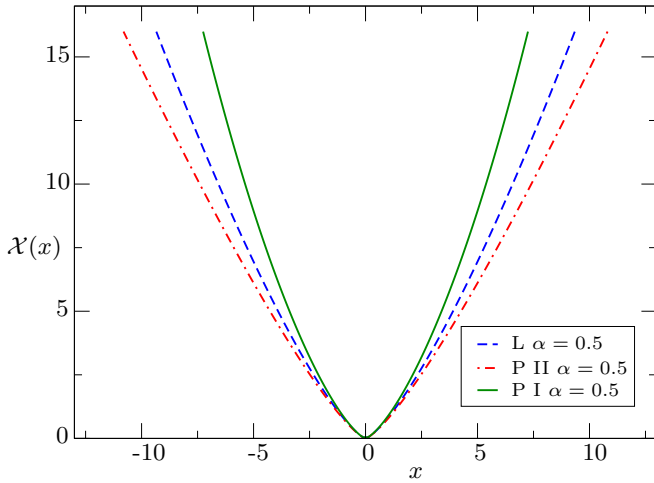


FIG. 2. A comparison of the rate functions $\mathcal{X}(x)$ for the CTRW with Gaussian SLD and the WTDs following a one-sided Lévy law, and Pareto type I and type II distributions, all with $\alpha = 0.5$ and the same $\langle n(t) \rangle$ as given by Eq. (14).

grow faster at large x than those for Pareto type II WTDs with the same α .

We now discuss in more detail these asymptotic growth properties, discussing the domain of very large deviations. The cases of the Pareto distributions give enough physical intuition to understand the general behavior of very large deviations.

V. VERY LARGE DEVIATIONS FOR PARETO WAITING TIME DENSITIES

Physically, the behavior of very large deviations ($x^2 \rightarrow \infty$) is dominated by realizations in which the number of steps

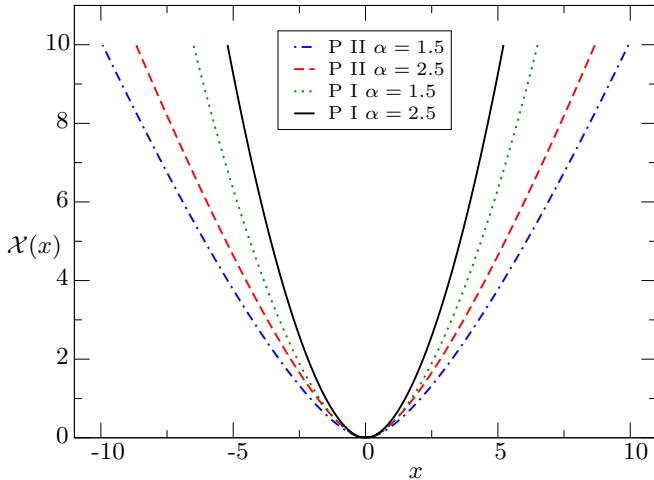


FIG. 3. A comparison of the rate function $\mathcal{X}(x)$ for the CTRW with Gaussian SLD and the following WTDs: Pareto I distributions with $\alpha = 1.5, 2.5$ and Pareto II with $\alpha = 1.5, 2.5$. Here, the variable x is the quotient between the position and the time. All these WTDs have a mean waiting time. Note the difference in the asymptotic behavior, which for the case of Pareto I is quadratic [Eq. (21)], whereas for the Pareto II is linear with a slowly varying correction [Eq. (23)].

is unusually large, and thus is governed by the behavior of $\psi(t)$ for very short t . Therefore the two Pareto cases serve as examples for the cases when $\psi(t)$ tends to a constant limit for $t \rightarrow 0$ (Pareto II), and when it vanishes identically for $t \rightarrow 0$ (Pareto I). There is no wonder that the Lévy case with $\psi(0) = 0$, but shooting up faster than any power of t for nonvanishing but small t , shows a behavior in between these two extrema.

Note that the expressions for both Pareto distributions [Eqs. (16) and (18)] are the same for all values of the parameter α , the only difference being the values of the scaling parameters C_I and C_{II} . Hence, all subsequent results hold independent of the value of α .

A. Pareto Type I WTD

First, let us consider the Pareto I WTD [Eq. (16)], and let us work with its Laplace transform as given by Eq. (17). The small time behavior ($t \rightarrow 0$) of the WTD is mirrored in the asymptotic behavior of its Laplace transform for $s \rightarrow \infty$; see Appendix C for additional discussions. For the Pareto I PDF this asymptotic behavior is given by

$$\tilde{\psi}(s) \sim \alpha C_I^{-1} s^{-1} e^{-C_I s}, \tag{20}$$

a form which can alternatively be obtained either by using the asymptotic expansion of the incomplete gamma function, or by evaluating the corresponding integral for the Laplace transform using the Laplace method. From this form it follows that

$$\mathcal{L}(q) = \ln \alpha - \ln C_I - \ln(-q) + C_I q.$$

Performing the Legendre transformation we get

$$t = \frac{d\mathcal{L}(q)}{dq} = C_I - \frac{1}{q} \rightarrow q = \frac{1}{C_I - t},$$

and finally obtain

$$\mathcal{I}(t) = -\ln(t - C_I)$$

in the leading order. Now,

$$g(\xi) = \frac{\mathcal{I}(\xi)}{\xi} = -\frac{\ln(\xi - C_I)}{\xi},$$

and the second Legendre transform can be performed:

$$-\frac{x^2}{2} = \frac{dg(\xi)}{d\xi} = -\frac{1}{\xi(\xi - C_I)} + \frac{\ln(\xi - C_I)}{\xi^2}.$$

Making the change of variable $u = \xi - C_I$, it can be rewritten as

$$-\frac{x^2}{2} = -\frac{1}{u(u + C_I)} + \frac{\ln u}{(u + C_I)^2}.$$

Very large and negative values of the l.h.s. correspond to $u \rightarrow 0$, so that

$$-\frac{x^2}{2} = -\frac{1}{uC_I} + \frac{\ln u}{C_I^2}.$$

To invert this expression, one can apply de Bruijn’s theorem for slowly varying functions; see [28]. Hence, one ends up

with

$$u = \frac{2}{C_I x^2} \left[1 + \frac{\ln(C_I x^2)}{C_I^2 x^2} \right],$$

and, going back to the variable ξ :

$$\xi = \frac{2}{C_I x^2} \left[1 + \frac{\ln(C_I x^2)}{C_I^2 x^2} \right] + C_I.$$

Finally, the asymptotic behavior ($|x| \rightarrow +\infty$) of the rate function for the CTRW has the form

$$\mathcal{X}(x) \sim \frac{C_I}{2} x^2 + \frac{2}{C_I} \ln |x|, \quad (21)$$

which is basically a quadratic behavior with a correction given by a slowly varying function. Note that, apart from the value of the constant C_I being a function of α , Eq. (21) does not depend on the parameter α .

B. Pareto Type II WTD

Let us now consider the case of a Pareto II WTD [Eq. (18)], which in the Laplace domain is given by Eq. (19). Following the same procedure as for the Pareto I WTD, let us consider the asymptotic behavior ($s \rightarrow \infty$) of Eq. (19) given by

$$\tilde{\psi}(s) \sim \alpha C_{II}^{-1} s^{-1}$$

[the difference with Eq. (20) is the absence of the exponential cutoff for very large s], which allows us to obtain the asymptotics of the cumulant generating function:

$$\mathcal{L}(q) = \ln \alpha - \ln C_{II} - \ln(-q). \quad (22)$$

Applying the Legendre transformation, we get

$$t = \frac{d\mathcal{L}(q)}{dq} = -\frac{1}{q} \rightarrow q = -\frac{1}{t},$$

so that the function $\mathcal{I}(t)$ in the leading order is given by

$$\mathcal{I}(t) = -\ln t.$$

Then we construct

$$\frac{\mathcal{I}(\xi)}{\xi} = -\frac{\ln \xi}{\xi},$$

and perform the second Legendre transformation

$$-\frac{x^2}{2} = \frac{dg(\xi)}{d\xi} = \frac{-1 + \ln(\xi)}{\xi^2}.$$

The values of $|x| \rightarrow \infty$ correspond to $\xi \rightarrow 0$, so that

$$\frac{x^2}{2} = -\frac{\ln \xi}{\xi^2}.$$

To invert this expression, one applies again de Bruijn's theorem. Hence, the inverse reads

$$\xi = |x|^{-1} \sqrt{2 \ln |x|}.$$

Finally, the asymptotic behavior ($|x| \rightarrow +\infty$) of the rate function for the CTRW is

$$\mathcal{X}(x) \sim |x| \sqrt{2 \ln |x|}, \quad (23)$$

which is an essentially linear behavior, with a correction given by a slowly varying function. This result does not depend on

the value of the parameter α at all (and not only up to the parameter values, like in the Pareto I case).

C. Erlang distributions

A similar analysis can be performed for gamma distributions (Erlang distributions) discussed in [14,16]

$$\psi(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!},$$

with $n \in \{1, 2, 3, \dots\}$, and $\lambda \in (0, \infty)$. The Laplace characteristic function of the Erlang distributions has the following form:

$$\tilde{\psi}(s) = \lambda^n (\lambda + s)^{-n}.$$

Then the asymptotic form of $\mathcal{L}(q)$ differs from Eq. (22) only by an additional proportionality factor in front of $\ln(-q)$:

$$\mathcal{L}(q) = -n \ln(-q),$$

in the limit $q \rightarrow -\infty$. Therefore, the essentially linear behavior with a slowly varying correction ensues. The difference between the Lévy and Pareto type I cases on one hand and Pareto type II and Erlang cases on the other hand is the fact that for the first class of distributions the WTD vanishes at zero together with all its derivatives, while in the second situation this is no longer the case. The full classification of possible behaviors will be discussed elsewhere. The lesson learned from these examples is that the essentially linear behavior of the large deviation rate function in CTRW is not universal but is pertinent to the specific classes of the waiting time distributions. The large deviation behavior of displacements probes the WTD for very short waiting times and is therefore a test for microscopic dynamics of the system.

VI. SUMMARY

In this paper, we presented a general procedure to compute the rate function for large deviations of displacements in CTRW in a general setting, i.e., for any step length distribution and any waiting time distribution. The situation with Gaussian step length distribution is especially simple. In this case the rate function for displacement is given by a sequence of Legendre transforms of the cumulant generating function for waiting times. The general discussion is accompanied by analyzing important particular examples like the one-sided Lévy and the Pareto-distributed waiting times. This discussion shows that the large deviation in displacement probes the waiting time density for very short times. The essentially linear behavior of the rate function for very large deviations is specific only for situations in which waiting time density does not vanish too fast when the waiting times approach zero.

ACKNOWLEDGMENTS

The work of A.P.P. was financially supported by ‘‘Doctoral Programmes in Germany’’ funded by the Deutscher Akademischer Austauschdienst (DAAD) (Programme ID 57440921).

APPENDIX A: PROPERTIES OF THE RATE FUNCTION

For completeness, we present here a sketch of the proof for some of the properties of the rate function, the ones mentioned in the main text of this work. For a more detailed discussion of the proof, see [29]. The rate function $\mathcal{C}(x)$ is defined as

$$\mathcal{C}(x) = \sup_q [qx - \ln f(q)] \quad (\text{A1})$$

where $f(q) = \langle e^{qx} \rangle$ is the moment generating function [note that we define $\mathcal{L}(q) = \ln f(q)$ in the main text].

With the definition (A1), the rate function is a convex nonnegative function which satisfies $\mathcal{C}(x = \mu) = 0$. In order to prove the convexity, we simply compute for $\lambda \in (0, 1)$

$$\begin{aligned} \mathcal{C}[\lambda x_1 + (1 - \lambda)x_2] &= \sup_q [q[\lambda x_1 + (1 - \lambda)x_2] - \ln f(q)] \\ &= \sup_q \{\lambda[qx_1 - \ln f(q)] + (1 - \lambda)[qx_2 - \ln f(q)]\} \\ &\leq \lambda \sup_q [qx_1 - \ln f(q)] + (1 - \lambda) \sup_q [qx_2 - \ln f(q)] \\ &= \lambda \mathcal{C}(x_1) + (1 - \lambda)\mathcal{C}(x_2). \end{aligned}$$

Now, to prove the nonnegativity of the rate function, let us use the fact that $f(0) = 1$, so $\ln f(0) = 0$. Now, since $\mathcal{C}(x) \geq qx - \ln f(q)$, taking $q = 0$ implies $\mathcal{C}(x) \geq 0x - \ln f(0) = 0$ and the nonnegativity is established. Now, let us use Jensen's inequality, which can be applied since $f(q)$ is a convex function; then

$$f(q) = \langle \exp(qx) \rangle \leq \exp(q\langle x \rangle) = \exp(q\mu).$$

Next, by applying the natural logarithm in both sides of the inequality, it can be expressed as

$$q\mu - \ln f(q) \leq 0.$$

Finally, because of the nonnegativity of the rate function, one concludes

$$\mathcal{C}(\mu) = 0 = \min_{x \in \mathbb{R}} \mathcal{C}(x).$$

$$\exp\left[-t\mathcal{T}\left(\frac{\tau}{t}\right)\right] \sim -\frac{d}{d\tau} \int_0^t \exp\left[-\tau\mathcal{I}\left(\frac{t'}{\tau}\right)\right] dt' = \int_0^t \left[\mathcal{I}\left(\frac{t'}{\tau}\right) - \frac{t'}{\tau}\mathcal{I}'\left(\frac{t'}{\tau}\right)\right] \exp\left[-\tau\mathcal{I}\left(\frac{t'}{\tau}\right)\right] dt' \quad (\text{B2})$$

or

$$\exp\left[-t\mathcal{T}\left(\frac{\tau}{t}\right)\right] \sim \frac{d}{d\tau} \int_t^\infty \exp\left[-\tau\mathcal{I}\left(\frac{t'}{\tau}\right)\right] dt' = -\int_t^\infty \left[\mathcal{I}\left(\frac{t'}{\tau}\right) - \frac{t'}{\tau}\mathcal{I}'\left(\frac{t'}{\tau}\right)\right] \exp\left[-\tau\mathcal{I}\left(\frac{t'}{\tau}\right)\right] dt'. \quad (\text{B3})$$

The expression in the nonexponential part of both Eqs. (B2) and (B3),

$$f(x) = \mathcal{I}(x) - x\mathcal{I}'(x)$$

(with $x = t'/\tau$) is nonnegative for $x < z_0$ and nonpositive for $x > z_0$. The first statement (for $x \geq 0$) follows immediately from the fact that $\mathcal{I}(x)$ is nonnegative and its derivative for $x < z_0$ nonpositive. The second statement is slightly finer and follows from the relation

$$g(z) \geq g(y) + g'(y)(z - y) \quad (\text{B4})$$

APPENDIX B: RELATION BETWEEN THE RATE FUNCTIONS FOR THE LEADING AND DIRECTING PROCESSES

The relation between the large deviation rates for the leading and the directing process of the CTRW scheme is based on the known relation between the cumulative distribution functions (CDFs) for the leading and directing processes [22,30]. Let $p(\tau|t) = T(\tau, t)$ be the probability density of having exactly τ steps up to time t (i.e., the probability density of τ conditioned on t), and $p(t|\tau)$ the corresponding density of the time of the last step conditioned on the number of steps. Then, from the fact that t is a monotonically nondecreasing function of τ , the integral relation between the two follows. In the continuous limit this relation takes the form

$$\int_\tau^\infty p(\tau'|t) d\tau' = \int_0^t p(t'|\tau) dt'. \quad (\text{B1})$$

We note that the rate function $\mathcal{I}(z)$ is already known. From the general properties of the rate functions it follows that $\mathcal{I}(z)$ is a convex function, monotonically nongrowing (in our case, essentially monotonically decaying) for $z < z_0$ with z_0 being equal to the mean waiting time, and monotonically nondecaying (in our case growing) for $z > z_0$. If the mean diverges, the function is always monotonically decaying. From Eq. (B1) it follows that

$$p(\tau|t) = -\frac{d}{d\tau} \int_0^t p(t'|\tau) dt',$$

or, equivalently,

$$p(\tau|t) = \frac{d}{d\tau} \left[1 - \int_0^t p(t'|\tau) dt' \right] = \frac{d}{d\tau} \int_t^\infty p(t'|\tau) dt'.$$

Substituting the large deviation forms we get

for convex differentiable functions, which we rewrite as

$$g(y) - yg'(y) \leq g(z) - g'(y)z.$$

Now one takes $g(y) = \mathcal{I}(y)$ and $z = z_0$, so that $g(z_0) = \mathcal{I}(z_0)$ vanishes, and notes that for $y > z_0$ the derivative $g'(y) = \mathcal{I}'(y)$ is nonnegative. Therefore the prefactors of the exponentials in both integrals on the r.h.s. of Eqs. (B2) and (B3) are nonnegative (essentially, positive for nondegenerated cases). For $z < z_0$ the function $\mathcal{I}(z)$ is monotonically decaying [$\mathcal{I}'(z) < 0$], and the argument of the exponential is therefore

monotonically growing towards the upper integration boundary. For $z > z_0$ the function $\mathcal{I}(z)$ is monotonically growing [$\mathcal{I}'(z) > 0$], and the absolute maximum of the integrand is achieved on the lower integration boundary.

To get the relation between \mathcal{I} and \mathcal{T} , the argument of the function on the l.h.s. has to be kept constant. Fixing $x = t/\tau$ and changing the variable of integration on the r.h.s. to $x' = t'/\tau$, one obtains

$$\exp\left[-\tau x \mathcal{T}\left(\frac{1}{x}\right)\right] \sim \tau \int_a^b |f(x')| \exp[-\tau \mathcal{I}(x')] dx',$$

where the limits of integration $\{a, b\}$ are $\{0, x\}$ in the case of Eq. (B2) and $\{x, \infty\}$ in the case of Eq. (B3). Assuming that, in the vicinity of x , $\mathcal{I}(x') \approx \mathcal{I}(x) + \mathcal{I}'(x)(x' - x) + o(x - x')$, both integrals can be estimated as

$$\frac{|f(x')|}{|\mathcal{I}'(x)|} \exp[-\tau \mathcal{I}(x)],$$

i.e., in the exponential order of magnitude

$$\exp\left[-\tau x \mathcal{T}\left(\frac{1}{x}\right)\right] \sim \exp[-\tau \mathcal{I}(x)],$$

from which Eq. (9) follows.

APPENDIX C: ASYMPTOTICS OF VERY LARGE DEVIATIONS FOR PARETO CASES

In the main text we relied on our physical intuition, and used the limiting form for $\tilde{\psi}(s)$ and therefore of $\mathcal{L}(q)$ for large absolute values of the arguments to calculate the Legendre transformation. Here we present some mathematical facts in support of our intuition.

The properties of $\mathcal{L}(q) = \ln \tilde{\psi}(-q)$ derive from those of $\tilde{\psi}(s)$. Some of these properties are general. Thus, $\tilde{\psi}(s) \leq 1$ and is monotonically nonincreasing ($\tilde{\psi}(s)$ is completely monotonic [31]). Therefore $\mathcal{L}(q) = \ln \tilde{\psi}(-q) < 0$ and is monotonically nondecreasing for all $-\infty < q \leq 0$. The asymptotics of $\mathcal{L}(q)$ follow from that of $\tilde{\psi}(s)$. For $q \rightarrow -\infty$ the function tends to $-\infty$; for $q \rightarrow 0_-$ it tends to zero.

The Pareto distributions belong to a class of infinitely divisible distributions [32]. For these distributions $\tilde{\psi}(s) = e^{-\phi(s)}$, where the function $\phi(s)$ is the one with completely monotonic derivative, and with $\phi(0) = 0$ [31]. Therefore $\mathcal{L}(q) = -\phi(-q)$ is infinitely differentiable on $(-\infty, 0_-)$ and convex.

For our discussions, however, we need only the properties of the two lower derivatives of $\mathcal{L}(q)$, which are common for all distributions with support on the right half-line, i.e., for all waiting time distributions, provided these are not the delta functions, like in simple random walks.

We show that both $Q(q) = \mathcal{L}'(q)$ and $Q'(q) = \mathcal{L}''(q)$ are strictly positive for all $-\infty < q < 0$ [the last statement means that $\mathcal{L}(q)$ is strictly convex]. Denoting $f(q) = \tilde{\psi}(-q) = \int_0^\infty e^{qt} \psi(t) dt$, as in the main text, we get $f'(q) = \int_0^\infty e^{qt} t \psi(t) dt$, $f''(q) = \int_0^\infty e^{qt} t^2 \psi(t) dt$. All three integrals are convergent for $q < 0$: the first one is always positive, and the last two are strictly positive for all $\psi(t) \neq \delta(t)$ (which holds in the Pareto case). The positivity of Q follows from

its representation

$$Q(q) = \frac{f'(q)}{f(q)},$$

where both the numerator and the denominator are positive and finite (for $q < 0$). Turning to $\mathcal{L}''(q)$ we write

$$\mathcal{L}''(q) = \frac{d}{dq} \frac{f'(q)}{f(q)} = \frac{f''(q)f(q) - [f'(q)]^2}{f^2(q)}.$$

The denominator is positive and finite. Now we write the expression in the numerator as a double integral,

$$\begin{aligned} & f''(q)f(q) - [f'(q)]^2 \\ &= \int_0^\infty \int_0^\infty (t''^2 - t''t') e^{q(t_1+t_2)} \psi(t'') \psi(t') dt'' dt', \end{aligned}$$

and symmetrize this expression,

$$\begin{aligned} & \int_0^\infty \int_0^\infty (t''^2 - t''t') e^{q(t_1+t_2)} \psi(t'') \psi(t') dt'' dt' \\ &= \frac{1}{2} \left[\int_0^\infty \int_0^\infty (t''^2 - t''t') e^{q(t_1+t_2)} \psi(t'') \psi(t') dt'' dt' \right. \\ & \quad \left. + \int_0^\infty \int_0^\infty (t'^2 - t''t') e^{q(t_1+t_2)} \psi(t'') \psi(t') dt'' dt' \right] \\ &= \int_0^\infty \int_0^\infty (t'' - t')^2 e^{q(t_1+t_2)} \psi(t'') \psi(t') dt'' dt', \end{aligned}$$

to see that it is greater than zero for all $\psi(t) \neq \delta(t - a)$. Therefore for our examples $\mathcal{L}''(q) > 0$ (for $-\infty < q < 0$), and thus $Q(q) = \mathcal{L}'(q)$ is a continuous, monotonically increasing function, which thus is invertible. Moreover, $Q(q) > 0$ for all $q < 0$.

The natural variable $\xi = Q(q) = \mathcal{L}'(q)$ of the Legendre-transformed function is therefore a monotonic function of q . The inverse function $q = Q^{-1}(\xi)$ (which defines the inverse transformation $q = d\mathcal{I}/d\xi$) exists, and is monotonic in ξ , namely a monotonically growing function thereof. Since $Q'(q) = \mathcal{L}''(q)$ is strictly positive, the inverse function $Q^{-1}(\xi)$ is differentiable as well. The continuity and differentiability of all involved functions within their definition intervals (with exclusion of the endpoints) allows us to exchange the limiting transitions, i.e., to consider the limits of the Legendre transformations as Legendre transformations of the limiting forms of the functions.

Now we return to our Pareto distributions. For $q \rightarrow -\infty$ the value of ξ tends to a positive constant C_I (for type I distribution) or to zero (for type II distribution). For $q \rightarrow 0_-$ it turns either to a positive constant μ (if the mean exists), which is larger than C_I or 0 for type I and type II distributions, respectively, or diverges going to $+\infty$ (if the mean is absent). The function $\mathcal{I}(\xi)$, being the Legendre transformation of $\mathcal{L}(q)$, is therefore defined either on a finite or on an infinite interval right from $\xi_{\min} \geq 0$ (being C_I or zero). It is important to note that in all cases the allowed values of ξ are positive: $\xi \geq 0$.

According to Eq. (12) the value of ξ which delivers the supremum necessary to calculate $\mathcal{X}(x)$ is given by the solution of the equation

$$\frac{x^2}{2} = -\frac{d}{d\xi} \left[\frac{\mathcal{I}(\xi)}{\xi} \right] = \frac{\mathcal{I}(\xi) - \mathcal{I}'(\xi)\xi}{\xi^2},$$

which we now consider for $x \rightarrow \infty$. The derivative of the numerator

$$\frac{d}{d\xi} [\mathcal{I}(\xi) - \mathcal{I}'(\xi)\xi] = -\xi \mathcal{I}''(\xi)$$

is negative for all $\xi > 0$, so that the whole expression (12) is a continuous and monotonically decreasing function of ξ . Therefore the solution to Eq. (12), provided it exists, is a decreasing function of x^2 . In this situation, going to the limit $x^2 \rightarrow \infty$ corresponds to $\xi \rightarrow \xi_{\min}$, which on its turn corresponds to $q \rightarrow -\infty$ and $s \rightarrow \infty$.

-
- [1] E. W. Montroll and G. H. Weiss, Random walks on lattices. II, *J. Math. Phys.* **6**, 167 (1965).
- [2] F. Lundberg, *Approximerad Framställning av Sannolikhetsfunktionerna: Återförsäkring av Kollektivrisker* (Almqvist & Wiksells, 1903).
- [3] E. Scalas, R. Gorenflo, H. Lückock, F. Mainardi, M. Mantelli, and M. Raberto, Anomalous waiting times in high-frequency financial data, *Quant. Finance* **4**, 695 (2004).
- [4] J. Klafter and I. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, 2011).
- [5] H. Scher and E. W. Montroll, Anomalous transit-time dispersion in amorphous solids, *Phys. Rev. B* **12**, 2455 (1975).
- [6] J. Haus and K. Kehr, Diffusion in regular and disordered lattices, *Phys. Rep.* **150**, 263 (1987).
- [7] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.* **339**, 1 (2000).
- [8] R. Klages, G. Radons, and I. Sokolov, *Anomalous Transport: Foundations and Applications* (Wiley, New York, 2008).
- [9] J. Klafter, S. C. Lim, and R. Metzler, *Fractional Dynamics* (World Scientific, Singapore, 2011).
- [10] P. Chaudhuri, L. Berthier, and W. Kob, Universal Nature of Particle Displacements Close to Glass and Jamming Transitions, *Phys. Rev. Lett.* **99**, 060604 (2007).
- [11] B. Wang, S. M. Anthony, S. C. Bae, and S. Granick, Anomalous yet brownian, *Proc. Natl. Acad. Sci. USA* **106**, 15160 (2009).
- [12] A. C. Silva, R. E. Prange, and V. M. Yakovenko, Exponential distribution of financial returns at mesoscopic time lags: a new stylized fact, *Physica A* **344**, 227 (2004).
- [13] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, Brownian Yet Non-Gaussian Diffusion: From Superstatistics to Subordination of Diffusing Diffusivities, *Phys. Rev. X* **7**, 021002 (2017).
- [14] E. Barkai and S. Burov, Packets of Diffusing Particles Exhibit Universal Exponential Tails, *Phys. Rev. Lett.* **124**, 060603 (2020).
- [15] E. B. Postnikov, A. Chechkin, and I. M. Sokolov, Brownian yet non-Gaussian diffusion in heterogeneous media: From superstatistics to homogenization, *New J. Phys.* **22**, 063046 (2020).
- [16] W. Wang, E. Barkai, and S. Burov, Large deviations for continuous time random walks, *Entropy* **22**, 697 (2020).
- [17] H. Touchette, The large deviation approach to statistical mechanics, *Phys. Rep.* **478**, 1 (2009).
- [18] B. Kamphorst and B. Zwart, Uniform asymptotics for compound poisson processes with regularly varying jumps and vanishing drift, *Stochastic Processes Appl.* **129**, 572 (2019).
- [19] D. Denisov, A. B. Dieker, and V. Shneer, Large deviations for random walks under subexponentiality: The big-jump domain, *Ann. Probab.* **36**, 1946 (2008).
- [20] A. A. Borovkov and A. A. Mogulskii, Large deviation principles for trajectories of compound renewal processes. I, *Theory Probab. Appl.* **60**, 207 (2016); Large deviation principles for trajectories of compound renewal processes. II, *ibid.* **60**, 349 (2016).
- [21] A. A. Borovkov, On large deviation principles for compound renewal processes, *Math. Notes* **106**, 864 (2019); Extension of the invariance principle for compound renewal processes to the zones of moderately large and small deviations, *Theory Probab. Appl.* **65**, 511 (2021).
- [22] R. Gorenflo, F. Mainardi, and A. Vivoli, Continuous-time random walk and parametric subordination in fractional diffusion, *Chaos Soliton Fractals* **34**, 87 (2007).
- [23] R. Gorenflo and F. Mainardi, Subordination pathways to fractional diffusion, *Eur. Phys. J. Spec. Top.* **193**, 119 (2011).
- [24] F. Cecconi, M. Cencini, A. Puglisi, D. Vergni, and A. Vulpiani, From the law of large numbers to large deviation theory in statistical physics: An introduction, in *Large Deviations in Physics: The Legacy of the Law of Large Numbers*, edited by A. Vulpiani, F. Cecconi, M. Cencini, A. Puglisi, and D. Vergni (Springer, Berlin, 2014), pp. 1–27.
- [25] S. R. S. Varadhan, Large deviations, in *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010)* (World Scientific, Singapore, 2011), pp. 622–639.
- [26] P. W. Glynn and W. Whitt, Large deviations behavior of counting processes and their inverses, *Queueing Syst.* **17**, 107 (1994).
- [27] R. W. Butler, *Saddlepoint Approximations with Applications*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, 2007).
- [28] R. Bojanic and E. Seneta, Slowly varying functions and asymptotic relations, *J. Math. Anal. Appl.* **34**, 302 (1971).
- [29] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability (Springer, Berlin, 2009).
- [30] A. Baule and R. Friedrich, Joint probability distributions for a class of non-Markovian processes, *Phys. Rev. E* **71**, 026101 (2005).
- [31] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. (John Wiley & Sons, New York, 1971), Vol. II.
- [32] O. Thorin, On the infinite divisibility of the Pareto distribution, *Scand. Actuar. J.* **1977**, 31 (1977).