Large Deviations in Continuous Time Random Walks

Adrian Pacheco-Pozo\textsuperscript{1} and Igor M. Sokolov\textsuperscript{1,2}

\textsuperscript{1}Institut für Physik, Humboldt-Universität zu Berlin, Newtonstraße 15, D-12489 Berlin, Germany
\textsuperscript{2}IRIS Adlershof, Humboldt-University of Berlin, Newtonstraße 15, 12489 Berlin, Germany

We discuss large deviation properties of continuous-time random walks (CTRW) and present a general expression for the large deviation rate in CTRW 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.

I. INTRODUCTION

Continuous time random walk (CTRW) introduced by Montroll and Weiss \cite{1} is a generalization of a simple random walk model, in which the steps follow inhomogeneously in time. 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 spatial displacements in single steps (also being i.i.d. random variables) \cite{2}. In physics, CTRW is used to model systems showing anomalous diffusion, e.g. situations when the mean square displacements (MSD) $\langle x^2(t) \rangle$ does not grow linearly in time, as predicted by the 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 modelling, following the pioneering work \cite{3}, is discussed in the review articles \cite{4} and \cite{5}, with some newer developments discussed in chapters of collective monographs \cite{6} and \cite{7}. Many recent works have reported 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. \cite{8,9,10} as well as \cite{11,12,13} and references therein.

In Ref. \cite{12}, 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. \cite{13}, Wang, Barkai, and Burov presented explicit calculations for the case of exponential and Gamma- (Erlang) distributions of waiting times. 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 \cite{12,13}, where we consider both the case when the mean WTD is finite and the case when it diverges. For the case of Gaussian step length distribution, as in \cite{14}, 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.

The structure of the present work is as follows: In Sec. \textsuperscript{II} we conduct a preliminary discussion of the model and introduce notation used throughout the work. In Sec. \textsuperscript{III} we derive the main result, which we use in the discussion of several examples; first of them already given in this section. In Sec. \textsuperscript{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. \textsuperscript{V} concludes the work. One “derivation” of a known mathematical result on a physical level of rigor, which we consider to be useful for understanding the approach, is placed in the Appendix.

II. PRELIMINARIES AND NOTATION

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 \cite{2}. The process $x(n)$ is called with this respect a parent process of the CTRW (see Refs. \cite{15,16} 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 to make exactly $n$ steps up to time $t$ be $\chi_n(t)$. Then the PDF of displacements in CTRW is \cite{2}

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

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
\[
p(x, t) = \int_0^\infty f(x, \tau)T(\tau, t)\,d\tau,
\] where $f(x, \tau)$ is a continuous approximation to (fluid limit of) $p_n(x)$ and $T(\tau, t)$ is the same kind of approximation to $\chi_n(t)$. Eq. (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. 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 variables, $s_n = (x_1 + x_2 + ... + x_n)/n$ 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
\[
C(s) = -\lim_{n \to \infty} \frac{1}{n} \ln P_n(s)
\]
called the rate function (RF), or Cramér’s function, of large deviations for $s_n$. The distribution of $s_n$ for sufficiently large $n$ is then given, up to a normalization constant, by
\[
P_n(s) \sim \exp\left[-nC(s)\right].
\]
When making the inverse variable transformation from $s_n$ to the sum $S_n = x_1 + x_2 + ... + x_n = ns_n$ of the first $n$ random variables, one sees that its PDF takes the form
\[
p_n(S) \sim \exp\left[-nC\left(\frac{S}{n}\right)\right],
\]
up to an $n$-dependent normalization constant. Eq. (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 scripted 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: $T(t)$ and $R(x)$, 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
\[
f(x, \tau) \sim \exp\left[-\tau R\left(\frac{x}{\tau}\right)\right],
\]
and
\[
p(t, \tau) \sim \exp\left[-\tau I\left(\frac{t}{\tau}\right)\right].
\]
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 direct- and the ones for the whole CTRW (a function we seek to know) will be denoted by $T(\tau)$ and $\chi(x)$, respectively. The large deviation forms for the PDFs of $\tau$ and $x$ are:
\[
T(\tau, t) \sim \exp\left[-tT\left(\frac{\tau}{t}\right)\right],
\]
and
\[
p(x, t) \sim \exp\left[-t\chi\left(\frac{x}{t}\right)\right].
\]
Before starting our derivation, let us go through some important properties of the rate functions [17–19]. The rate functions for the mean of i.i.d. random variables are convex [18]. Cramér’s theorem states that the rate function $C(x)$ is the convex conjugate (a Legendre-Fenchel transformation) of the cumulant generation function $L(q) = \ln\langle e^{qX} \rangle$ of the distribution of $x$. For the sake of brevity we will call this transformation simply a Legendre transformation, as it 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 non-negative, 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 “literaly 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 \( f(x) \rightarrow g(z) \). Our notation will keep track of the variable’s names in the corresponding functions
\[ g(z) = \mathcal{L}[(z-x)]f(x). \]

The direction of the arrow is connected with the order of the symbols: the operator \( \mathcal{L} \) acts on the function of \( x \) standing 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 the left). Since we choose to denote the variable’s names in the notation for the operator, we don’t have to put them in the functions at all (but we will).

The main result of the present article is as follows:
\[ \mathcal{X}(z) = -\sup_x \left[ -\mathcal{R}(z,x) + \mathcal{I}(x) \right], \quad (7) \]

with \( \mathcal{I}(z) = \mathcal{L}[(z-x)]\mathcal{L}(q) \) and \( \mathcal{R}(z) = \mathcal{L}[(z-x)]\mathcal{D}(q) \) where \( \mathcal{L}(q) = e^{qI} = \int_0^\infty e^{q\psi(t)}dt \) with \( \psi(t) \) being the PDF of waiting times (waiting time density, WTD), and \( \mathcal{D}(q) = e^{qI} = \int_0^\infty e^{q\lambda(x)}dx \) with \( \lambda(x) \) being the PDF of displacements in a single step (steps’ lengths density, SLD).

For a Gaussian distribution of single steps’ lengths discussed in [12, 14], the result can be put in a closed form involving only the Legendre transformations
\[ \mathcal{X}(x) = -\mathcal{L}[-x^2/2e^{-\xi}] \left[ \xi^{-1} \mathcal{L}(1)\mathcal{L}(q) \right]. \quad (8) \]

III. THE LARGE DEVIATIONS RATE FOR CTRW

Let us now derive the forms, Eq. (7) and (8). Substituting the large deviation forms for PDFs of the parent and directing processes (Eqs. 6 and 5, respectively) into the integral formula of subordination, Eq. 1, one gets
\[ p(x,t) \sim \int_0^\infty \exp \left[ -\tau \mathcal{R} \left( \frac{x}{\tau} \right) - tT \left( \frac{T}{t} \right) \right] d\tau. \]

On the l.h.s., we introduce the large deviation form of the whole CTRW (Eq. 6)
\[ \exp \left[ -t\mathcal{X} \left( \frac{x}{t} \right) \right] \sim \int_0^\infty \exp \left[ -\tau \mathcal{R} \left( \frac{x}{\tau} \right) - tT \left( \frac{T}{t} \right) \right] d\tau \]

and change to the new variables \( z = x/t \) and \( \xi = t/\tau \):
\[ \exp \left[ -t\mathcal{X} \left( \frac{z}{t} \right) \right] \sim \int_0^\infty \frac{t}{\xi^2} \exp \left\{ -t \left[ \frac{\mathcal{R}(z,\xi)}{\xi} + T \left( \frac{1}{\xi} \right) \right] \right\} d\xi. \]

Now, we use the relation between the RF for the leading and directing processes, which is given by
\[ T(\tau) = \tau I \left( \frac{1}{\tau} \right), \quad (9) \]

see Ref. [21]. 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, we sketch a simple explanation on the physical level of rigor is given in Appendix A.

Using this relation one gets
\[ \exp \left[ -t\mathcal{X} \left( \frac{z}{t} \right) \right] \sim \int_0^\infty \frac{t}{\xi^2} \exp \left\{ -t \left[ \frac{\mathcal{R}(z,\xi)}{\xi} + I \left( \frac{1}{\xi} \right) \right] \right\} d\xi. \]

This integral can then be solved by the Laplace’s method [21]. Thus, we can write
\[ \exp \left[ -t\mathcal{X} \left( \frac{z}{t} \right) \right] \sim \exp \left\{ -t \inf \left[ \frac{\mathcal{R}(z,\xi)}{\xi} + I \left( \frac{1}{\xi} \right) \right] \right\}, \]

where the pre-exponential term is disregarded, since it does not contribute to the large deviations when taking the limit \( \lim_{\xi \rightarrow \infty} \ln[p(x,t)]/t \). Equating the arguments of the exponentials, we get
\[ \mathcal{X}(z) = -\sup_x \left[ -\mathcal{R}(z,x) + \mathcal{I}(x) \right], \]

which is our main result, Eq. (7). Note that, the infimum was changed for 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 [12, 14], 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 analyse the simplest random walk, i.e., the Bernoulli one.

A. CTRW with a fixed step length

As a first example, let us consider the Bernoulli random walk, with steps of fixed length \( \pm 1 \) 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. [17] and [18]. For this case
\[ p_N(S) = 2^{-N-1} \frac{N!}{(N/2 - S/2)!} \frac{(N/2 + S)!}{(N/2)!}. \]
Applying Stirling formula we get in the first order in $N$

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

with $s = S/N$, so that

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

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 $z$, at difference to 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

$$\mathcal{X}(z) = 1 + \sqrt{1 + z^2} \left\{ -1 + \ln \sqrt{1 + z^2} + \frac{1}{2} \left[ \left(1 - \frac{z}{\sqrt{1 + z^2}}\right) \ln \left(1 - \frac{z}{\sqrt{1 + z^2}}\right) + \left(1 + \frac{z}{\sqrt{1 + z^2}}\right) \ln \left(1 + \frac{z}{\sqrt{1 + z^2}}\right) \right] \right\}.$$

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 $\mathcal{I}(\xi) = \mathcal{I}(\xi_{-q}) \mathcal{L}(q)$ corresponding to a Poisson process can be readily calculated (and essentially is known since long ago), and reads

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

The supremum in Eq. (9), for fixed $z$ is achieved at $\xi = 1/\sqrt{1 + z^2}$, and $\mathcal{X}(z)$ is readily evaluated:

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

Eq. (10). Now, we define the function

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

and perform the Legendre transformation

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

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:

$$\mathcal{X}(x) = 1 - e^{-\frac{1}{2} W_0(x^2)} + \frac{x^2}{2} e^{-\frac{1}{2} W_0(x^2)} + \frac{1}{2} W_0(x^2) e^{-\frac{1}{2} W_0(x^2)}.$$ (12)

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 $x << 1$, $W_0(x^2) \sim x^2$, and performing a Taylor expansion around zero, one has $\mathcal{X}(x) \sim x^2/2 + O(x^4)$ corresponding to a Gaussian. Fig. 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. (12). 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. [13].
Now we take \( \psi(t) \) to follow a one-sided Lévy law with exponent \( \alpha \). Its Laplace characteristic function \( f(s) = \langle e^{-st} \rangle \) reads: \( f(s) = \int_0^\infty \psi(t) e^{-st} dt = \exp(-\sigma s^\alpha) \) with \( 0 < \alpha < 1 \) (for \( s > 0 \)) and \( \sigma \) being the scale parameter. For \( \alpha = 1 \) it tends to a parabolic RF of a Gaussian, for \( \alpha \to 0 \) to a purely exponential PDF tails, see [2].

### B. One-sided Lévy law

Now we take \( \psi(t) \) to follow a one-sided Lévy law with exponent \( \alpha \). Its Laplace characteristic function \( f(s) = \langle e^{-st} \rangle \) reads: \( f(s) = \int_0^\infty \psi(t) e^{-st} dt = \exp(-\sigma s^\alpha) \) with \( 0 < \alpha < 1 \) (for \( s > 0 \)) and \( \sigma \) being the scale parameter. For \( \alpha = 1 \) it tends to a parabolic RF of a Gaussian, for \( \alpha \to 0 \) to a purely exponential PDF tails, see [2].

For \( \alpha \to 0 \) it tends to a parabolic RF of a Gaussian, for \( \alpha \to 0 \) to a purely exponential PDF tails, see [2].
Due to the presence of incomplete Gamma functions, the Legendre transforms have to be performed numerically, by solving algebraic equations.

Fig. 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. (13). 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. (14)), but deviate for $x$ large.

![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, both with $\alpha = 0.5$, and the same $\langle n(t) \rangle$ as given by Eq. (13). 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. (14)), but deviate for $x$ large.](image)

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

![Fig. 3. A comparison of the rate function $\mathcal{X}(x)$ for the CTRW with Gaussian SLD and the WTDs following a one-sided Lévy law, and Pareto type I and type II distributions, both 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. (20)), whereas for the Pareto II is linear with a slowly varying correction (Eq. (22)).](image)

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

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

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 ($x^2 \to \infty$) is dominated by realizations in which the number of steps is unusually large, and thus is governed by the behavior of $\psi(t)$ for very short $t$. Therefore the two Pareto

A. Pareto Type I WTD

First, let us consider the Pareto I WTD (Eq. (15)), and let us work with its Laplace transform as given by Eq. (16). The small time behavior ($t \to 0$) of the WTD is mirrored in the asymptotic behavior of its Laplace transform for $s \to \infty$. For the Pareto I PDF this asymptotic behavior is given by

$$f(s) \sim \alpha C_I^{-1}s^{-1}e^{-Cs},$$

which form 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

$$\mathcal{L}(q) = \ln \alpha - \ln C_I - \ln(-q) + C_Iq.$$
Performing the Legendre transformation we get
\[ t = \frac{d\mathcal{L}(q)}{dq} = C_I - \frac{1}{q} \quad \rightarrow \quad q = \frac{1}{C_I - t}, \]
and finally obtain
\[ \mathcal{I}(t) = -\ln(t - C_I) \]
in the leading order. Now,
\[ \frac{\mathcal{I}(\xi)}{\xi} = -\frac{\ln(\xi - C_I)}{\xi}, \]
and the second Legendre transform can be performed:
\[ -\frac{x^2}{2} = \frac{df(\xi)}{d\xi} = -\frac{1}{\xi(\xi - C_I)} + \frac{\ln(\xi - C_I)}{\xi^2}. \]
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 \( u \to 0 \), so that
\[ -\frac{x^2}{2} = -\frac{1}{u C_I} + \frac{\ln u}{C_I^2}, \]
To invert this expression, one can apply de Bruijn’s Theorem for slowly varying functions, see [22]. Hence, one ends up with
\[ u = \frac{2}{C_I} \left[ 1 + \frac{\ln(C_I x^2)}{C_I^2 x^2} \right], \]
and, going back to the variable \( \xi \):
\[ \xi = \frac{2}{C_I} \left[ 1 + \frac{\ln(C_I x^2)}{C_I^2 x^2} \right] + C_I. \]
Finally, the asymptotic behavior (\(|x| \to +\infty\)) of the rate function for the CTRW has the form
\[ \lambda(x) \sim \frac{C_I}{2} x^2 + \frac{2}{C_I} \ln|x|, \quad (20) \]
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 \( \alpha \), Eq. (20) does not depend on the parameter \( \alpha \).

B. Pareto Type II WTD

Let us now consider the case of a Pareto II WTD (Eq. (17)), which in the Laplace domain is given by Eq. (18). Following the same procedure as for the Pareto I WTD, let us consider the asymptotic behavior \((s \to \infty)\) of Eq. (18) given by
\[ \widetilde{\psi}(s) \sim \alpha C_{II}^{-1} s^{-1}, \]
(the difference with Eq. (19) is the absence of the exponential cutoff for very large \( s \)), which allow us to obtain the asymptotics of the cumulant generating function:
\[ \mathcal{L}(q) = \ln \alpha - \ln C_{II} - \ln(-q). \quad (21) \]
Applying the Legendre transformation we get:
\[ t = \frac{d\mathcal{L}(q)}{dq} = -\frac{1}{q} \quad \rightarrow \quad q = -\frac{1}{t}, \]
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
\[ -\frac{x^2}{2} = \frac{df(\xi)}{d\xi} = \frac{-1 + \ln(\xi)}{\xi^2}. \]
The values of \(|x| \to \infty\) correspond to \( \xi \to 0 \), so that
\[ \frac{x^2}{2} = - \frac{\ln \xi}{\xi^2}. \]
To invert this expression, one applies again the de Bruijn’s Theorem. Hence, the inverse reads
\[ \xi = |x|^{-1} \sqrt{2 \ln|x|}. \]
Finally, the asymptotic behavior (\(|x| \to +\infty\)) of the rate function for the CTRW is
\[ \lambda(x) \sim |x| \sqrt{2 \ln|x|}, \quad (22) \]
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 \( \alpha \) 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 [12] [14]
\[ \psi(t) = \lambda^n t^{n-1} e^{-\lambda t}, \]
\[ \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, \]
with \( n \in \{1, 2, 3, \ldots \} \), and \( \lambda \in (0, \infty) \). The Laplace characteristic function of the Erlang distributions has the following form

\[
f(s) = \lambda^n (\lambda + s)^{-n}.
\]

Then the asymptotic form of \( \mathcal{L}(q) \) differs from Eq. (21) only by an additional proportionality factor in front of \( \ln(-q) \):

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

in the limit \( q \to -\infty \). 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 more 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 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.

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 analysing important particular examples like the one-sided Lévy and the Pareto-distributed waiting times. This discussion shows that the large deviation in displacement probe 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.

\[
\exp \left[ -t \mathcal{I} \left( \frac{\tau}{t} \right) \right] \sim -\frac{d}{dt} \int_0^t \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' = \int_0^t \left[ \mathcal{I} \left( \frac{\tau'}{\tau} \right) - \frac{\tau'}{\tau} \mathcal{I}' \left( \frac{\tau'}{\tau} \right) \right] \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' \quad (A2)
\]

or

\[
\exp \left[ -t \mathcal{I} \left( \frac{\tau}{t} \right) \right] \sim \frac{d}{dt} \int_t^\infty \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' = -\int_t^\infty \left[ \mathcal{I} \left( \frac{\tau'}{\tau} \right) - \frac{\tau'}{\tau} \mathcal{I}' \left( \frac{\tau'}{\tau} \right) \right] \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' \quad (A3)
\]

The expression in the non-exponential part of both Eqs. (A2) and (A3),

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

is already known.

ACKNOWLEDGMENTS

The work of APP was financially supported by “Doctoral Programmes in Germany” funded by the Deutscher Akademischer Austauschdienst (DAAD) (Programme ID 57440921).

Appendix A: 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 given on the known relation between the cumulative distribution functions (CDFs) for the leading and directing processes [13, 23]. Let \( p(\tau|t) = \mathcal{I}(\tau, t) \) be probability density to have exactly \( \tau \) steps up to time \( t \) (i.e., the probability density of \( \tau \) 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 non-decreasing function of \( \tau \), the integral relation between the two follows. In the continuous limit this relation takes the form

\[
\int_\tau^\infty p(\tau'|t) d\tau' = \int_0^t p(t'|\tau) dt'. \quad (A1)
\]

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 non-decreasing (in our case, essentially monotonically decaying) for \( z < z_0 \) with \( z_0 \) being equal to the mean waiting time, and monotonically non-decaying (in our case growing) for \( z > z_0 \). If the mean diverges, the function is always monotonically decaying. From Eq. (A1) it follows that

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

or, equivalently,

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

Substituting the large deviation forms we get

\[
\exp \left[ -t \mathcal{I} \left( \frac{\tau}{t} \right) \right] \sim -\frac{d}{dt} \int_0^t \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' = \int_0^t \left[ \mathcal{I} \left( \frac{\tau'}{\tau} \right) - \frac{\tau'}{\tau} \mathcal{I}' \left( \frac{\tau'}{\tau} \right) \right] \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' \quad (A2)
\]

or

\[
\exp \left[ -t \mathcal{I} \left( \frac{\tau}{t} \right) \right] \sim \frac{d}{dt} \int_t^\infty \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' = -\int_t^\infty \left[ \mathcal{I} \left( \frac{\tau'}{\tau} \right) - \frac{\tau'}{\tau} \mathcal{I}' \left( \frac{\tau'}{\tau} \right) \right] \exp \left[ -\tau \mathcal{I} \left( \frac{\tau'}{\tau} \right) \right] d\tau' \quad (A3)
\]

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

\[ g(z) \geq g(y) + g'(y)(z - y) \quad (A4) \]

for convex differentiable functions, which we rewrite as

\[ g(y) - yg'(y) \leq g(z) - g'(y)z. \]

Now one takes \( g(y) = I(y) \) and \( z = z_0 \), so that \( g(z_0) = I(z_0) \) vanishes, and notes that for \( y > z_0 \) the derivative \( g'(y) = I'(y) \) is non-negative. Therefore the prefactors of the exponentials in both integrals on the r.h.s. of Eqs. [A2] and [A3] are non-negative (essentially, positive for non-degenerated cases). For \( z < z_0 \) the function \( I(z) \) is monotonically decaying (\( I'(z) < 0 \)), and the argument of the exponential is therefore monotonically growing towards the upper integration boundary. For \( z > z_0 \) the function \( I(z) \) is monotonically growing (\( I'(z) > 0 \)), and the absolute maximum of the integrand is achieved on the lower integration boundary.

To get the relation between \( I \) and \( T \), 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:

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

where the limits of integration \( \{a, b\} \) are \( \{0, x\} \) in case of Eq. [A2] and \( \{x, \infty\} \) in case of Eq. [A3]. Assuming that in the vicinity of \( x \), \( I(x') \approx I(x) + I'(x)(x' - x) + o(x - x') \), both integrals can be estimated as

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

i.e. in the exponential order of magnitude

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

from which Eq. [9] follows.