Record statistics and persistence for a random walk with a drift

Satya N Majumdar, Grégory Schehr and Gregor Wergen

1 Laboratoire de Physique Théorique et Modèles Statistiques, UMR 8626, Université Paris Sud
11 and CNRS, Bât. 100, Orsay F-91405, France
2 Institut für Theoretische Physik, Universität zu Köln, D-50937 Köln, Germany

E-mail: satya.majumdar@u-psud.fr, gregory.schehr@u-psud.fr and gw@thp.uni-koeln.de

Received 29 June 2012
Published 15 August 2012
Online at stacks.iop.org/JPhysA/45/355002

Abstract
We study the statistics of records of a one-dimensional random walk of \( n \) steps, starting from the origin, and in the presence of a constant bias \( c \). At each time step, the walker makes a random jump of length \( \eta \) drawn from a continuous distribution \( f(\eta) \), which is symmetric around a constant drift \( c \). We focus in particular on the case where \( f(\eta) \) is a symmetric stable law with a Lévy index \( 0 < \mu \leq 2 \). The record statistics crucially depends on the persistence probability, which, as we show here, exhibits different behaviors depending on the sign of \( c \) and the value of the parameter \( \mu \). Hence, in the limit of a large number of steps \( n \), the record statistics is sensitive to these parameters (\( c \) and \( \mu \)) of the jump distribution. We compute the asymptotic mean record number \( \langle R_n \rangle \) after \( n \) steps as well as its full distribution \( P(R, n) \). We also compute the statistics of the ages of the longest and the shortest lasting record. Our exact computations show the existence of five distinct regions in the \((c, 0 < \mu \leq 2)\) strip where these quantities display qualitatively different behaviors. We also present numerical simulation results that verify our analytical predictions.

PACS numbers: 02.50.Ga, 05.40.Fb, 05.45.Tp

(Some figures may appear in colour only in the online journal)

1. Introduction

The statistical properties of record-breaking events in stochastic processes have been a popular subject of research in recent years. The theory of records has found many interesting applications. Record events are very important in sports [1, 2] and climatology [3–6], but have also been found relevant in biology [7], in the theory of spin glasses [8, 9] and in models of growing networks [10]. Also, in finance, record-breaking events, e.g., when the price of a stock breaks its previous records, can lead to increased financial activities [11, 12]. In all of these fields, researchers have recently made progress in understanding and modeling the statistics of records by comparing the records in observational data with various kinds of stochastic
processes. In this context, it has become increasingly important to improve our understanding of the record statistics of elementary stochastic processes. In this paper, we focus on one such elementary stochastic process namely a random walk in the presence of a constant bias. We show that even for such a simple process, its record statistics is considerably nontrivial and rich.

In general, one is interested in the record events of a discrete-time series of random variables \( x_0, x_1, \ldots, x_n \). An (upper) record is an entry \( x_k \), which exceeds all previous entries:

\[
    x_k > \max(x_0, x_1, \ldots, x_{k-1}).
\]

Until the end of the past century, record statistics was fully understood only in the case when the entries of the time series are independent and identically distributed (i.i.d.) RVs (see, e.g., [13–15]). For i.i.d. RVs from a continuous distribution \( p(x) \), the probability \( r_n \) of a record in the \( n \)th time step is given by [13]

\[
    r_n = \text{Prob}\{x_n > \max(x_0, x_1, \ldots, x_{n-1})\} = \frac{1}{n+1},
\]

which is universal, i.e. independent of the parent distribution \( p(x) \). This universality follows simply from the isotropy in ordering, i.e. any one of the \( (n+1) \) entries is equally probable to be a record. Let \( R_n \) denote the total number of records up to step \( n \). The mean record number is then simply \( \langle R_n \rangle = \sum_{m=0}^{n} r_m \), which grows asymptotically as \( \sim \ln n \) for large \( n \).

Due to the numerous applications of the theory of records, it became interesting to consider more general models. There has been a lot of interest in the record statistics of RVs which are uncorrelated but no longer identical. For instance, Ballerini et al considered uncorrelated RVs with a linear drift [16]. More recently, Franke et al also studied the same problem and found numerous new results [17–19] by also considering the correlations between individual record events. This model was then successfully applied to the statistics of temperature records in the context of global warming [5]. In 2006, Krug studied the statistics of records of uncorrelated RVs with a time-increasing standard deviation, a model with important biological implications [20].

Another important issue is the study of record statistics for correlated RVs. For weak correlations, with a finite correlation time, one would expect the record statistics for a large sequence to be asymptotically similar to the uncorrelated case. This is no longer true when there are strong correlations between the entries. Perhaps, one of the simplest and most natural time series with strong correlations between its entries corresponds to the positions of a one-dimensional random walk [21]. Despite the striking importance and abundance of random walk in various areas of research, the record statistics of a single, discrete-time random walk with a symmetric jump distribution was not computed and understood until only a few years ago. In 2008, Majumdar and Ziff [22] exactly computed the record statistics of a one-dimensional symmetric random walk model and showed that the record rate of such a process is completely universal for any continuous and symmetric jump distribution, thanks to the so-called Sparre Andersen theorem [23]. They considered a time series of RVs \( x_m \) given by

\[
    x_m = x_{m-1} + \eta_m,
\]

where \( \eta_m \) are i.i.d. RVs drawn from a symmetric and continuous jump distribution \( f(\eta) \) (it includes even Lévy flights where \( f(\eta) \sim 1/|\eta|^{\mu+1} \) with \( 0 < \mu < 2 \)). Then, the record rate \( r_n \) for such a process is given by the universal formula [22]

\[
    r_n = \left( \frac{2n}{n} \right) 2^{-2n} \frac{1}{\sqrt{n}},
\]

independently of the jump distribution \( f(\eta) \). They also computed exactly the mean record number \( \langle R_n \rangle \) and even its full distribution [22]. In addition, there exists nice connection between the record statistics and the extreme value statistics for the one-dimensional symmetric jump.
processes and many universal results can be subsequently derived using the Sparre Andersen theorem (see [24] for a review).

Following [22], there has been considerable interests in generalizing them to more general set of strongly correlated stochastic processes. For instance, Sabhapandit discussed symmetric random walks with a random, possibly heavy tailed, waiting time between the individual jumps (the so-called continuous time random walk model) [25]. Recently, the present authors considered the record statistics of an ensemble of $N$ independent and symmetric random walks [12]. There, in contrast to the case of a single random walker, the record statistics of $N$ Lévy flights with a heavy-tailed jump distribution was found to be different from the one of $N$ Gaussian random walkers with a jump distribution that has a finite second moment.

Another important generalization is to consider a single one-dimensional random walker but with asymmetric jump distribution, e.g., in the presence of a constant bias $c$. First steps toward this generalization were taken by Le Doussal and Wiese [26] who derived the exact record statistics for a biased random walker with a Cauchy jump distribution (a special case of Lévy flights with Lévy index $\mu = 1$). More recently, in 2011, Wergen et al showed that a biased random walk is useful to model record-breaking events in daily stock prices [11]. They were able to obtain results in some special limits of a biased random walker with a Gaussian jump distribution. Apart from these two special cases, namely the Cauchy and the Gaussian jump distribution, there are no other analytical results available, to our knowledge, for other jump distributions for a biased random walker. Recently, the record statistics for a biased random walker was also numerically studied in order to quantify the contamination spread in a porous medium via the particle-tracking simulations [27].

In this paper, we present a complete analysis of the record statistics for a biased random walker with arbitrary jump distributions. As we will see, the record statistics crucially depends on the persistence probability $Q(n)$ (see equation (17)), the probability that the biased walker stays to the left of its initial starting position up to $n$ steps. While the persistence probability for various stochastic processes has extensively been studied in the recent past [28], it seems that for this simple biased jump process, it has not been systematically studied in the literature to the best of our knowledge. Here, we provide exact results for the persistence probability $Q(n)$ for a biased random walk arbitrary jump distributions (see equation (67)), which subsequently leads to the exact record statistics for the same process.

The rest of the paper is organized as follows. Since the paper is long with many detailed results, we provide in section 2 a short review on the record statistics for random walks both with and without bias, followed by a summary of the main results of this paper. Readers not interested in the details of the calculations can skip the rest of the paper. In section 3, we will show how to use the renewal property of the random walk and a generalized version of the Sparre Andersen theorem [23] to compute the persistence of random walks in the presence of both the positive and the negative drift. The results for the persistence are interesting on their own and will be discussed in detail in section 4, but they will also allow us to compute the record statistics. In particular, we will show that, in the presence of drift, the complete universality found for the record statistics in the unbiased case [22] breaks down and there are five different types of asymptotic behaviors which emerge depending on the two parameters of the model, namely the drift $c$ and the index $0 < \mu \leq 2$ characterizing the tail of the jump distribution. This record statistics will be discussed in detail in section 5. Later, in section 6, we will also discuss the extreme value statistics of the ages of the longest (section 6.1) and the shortest lasting records (section 6.2) in each of the regimes. We will show that the asymptotic behavior of these quantities is also systematically different in the five regimes. Finally, in section 7, we will conclude with some open problems.
2. Record statistics for random walks: a short review and a summary of new results

In this section, we provide a short review on the record statistics of a one-dimensional random walk model, with and without the external drift. This will also serve to set up our notations for the rest of the paper. At the end of this section, we summarize the main new results obtained in this work.

Let us first start with the driftless case. Consider a sequence of RVs \( x_0 = 0, x_1, x_2, \ldots, x_n \), where \( x_m \) represents the position of a discrete-time unbiased random walker at step \( m \). The walker starts at the origin and its position evolves via the Markov rule \( x_m = x_{m-1} + \eta_m \), where \( \eta_m \) represents the stochastic jump at the \( m \)th step. The jump variables \( \eta_m \) are i.i.d. RVs, each drawn from a common probability distribution function (pdf) \( f(\eta) \), normalized to unity. The pdf \( f(\eta) \) is continuous and symmetric with zero mean. Let \( \hat{f}(k) = \int_{-\infty}^{\infty} f(\eta) e^{ik\eta} d\eta \) denote the Fourier transform of the jump distribution. We will henceforth focus on jump distributions \( f(\eta) \) whose Fourier transform has the following small \( k \) behavior

\[
\hat{f}(k) = 1 - (l_\mu |k|)^\mu + \cdots,
\]

where \( 0 < \mu \leq 2 \) and \( l_\mu \) represents a typical length scale associated with the jump. The exponent \( 0 < \mu \leq 2 \) denotes the large \( |\eta| \) tail of \( f(\eta) \). For jump densities with a finite second moment \( \sigma^2 = \int_{-\infty}^{\infty} \eta^2 f(\eta) d\eta \), such as Gaussian, exponential, uniform etc, one evidently has \( \mu = 2 \) and \( l_2 = \sigma / \sqrt{2} \). In contrast, \( 0 < \mu < 2 \) corresponds to jump densities with fat tails \( f(\eta) \sim |\eta|^{-1-\mu} \) as \( |\eta| \to \infty \). A typical example is \( f(k) = \exp[-|k|^\mu] \), where \( \mu = 2 \) corresponds to the Gaussian jump distribution, while \( 0 < \mu < 2 \) corresponds to Lévy flights (for reviews on these jump processes see [29, 30]).

A quantity that will play a crucial role later is \( P_n(x) \) which denotes the probability density of the position of the symmetric random walk at step \( n \). Using the Markov rule in equation (2), it is easy to see that \( P_n(x) \) satisfies the recursion relation

\[
P_n(x) = \int_{-\infty}^{\infty} P_{n-1}(x') f(x - x') dx',
\]

starting from \( P_0(x) = \delta(x) \). This recurrence relation can be trivially solved by taking Fourier transform and using the convolution structure. Inverting the Fourier transform, one obtains

\[
P_n(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{f}(k)^n e^{-ikx}.
\]

In the limit of large \( n \), the small \( k \) behavior of \( \hat{f}(k) \) dominates the integral on the right-hand side (rhs) of equation (6). Substituting the small \( k \) behavior from equation (4), one easily finds that for \( 0 < \mu < 2 \), typically \( x \sim l_\mu n^{1/\mu} \) and \( P_n(x) \) approaches the scaling form [29]

\[
P_n(x) \to \frac{1}{l_\mu n^{1/\mu}} \mathcal{L}_\mu \left( \frac{x}{l_\mu n^{1/\mu}} \right), \quad \text{where} \quad \mathcal{L}_\mu(y) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-|k|^\mu} e^{-iky}.
\]

For \( 0 < \mu < 2 \), the scaling function \( \mathcal{L}_\mu(y) \) decays as a power law for large \( |y| \) [29]

\[
\mathcal{L}_\mu(y) \sim A_\mu |y|^{\mu+1}, \quad \text{where} \quad A_\mu = \frac{1}{\pi} \sin(\mu\pi/2) \Gamma(1+\mu).
\]

In particular, for \( \mu = 1 \), the scaling function \( \mathcal{L}_1(y) \) is precisely the Cauchy density itself:

\[
\mathcal{L}_1(y) = \frac{1}{\pi} \frac{1}{1 + y^2}.
\]
In contrast, for \( \mu = 2 \), the central limit theorem holds, \( x \sim \sigma_{n}^{1/2} \), and \( P_{n}(x) \) approaches a Gaussian scaling form

\[
P_{n}(x) \rightarrow \frac{1}{\sigma_{n}^{1/2}} L_{2} \left( \frac{x}{\sigma_{n}^{1/2}} \right), \quad \text{where} \quad L_{2}(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^{2}/2).
\]

From the sequence of symmetric RVs representing the position of a discrete-time unbiased random walker, we next construct a new sequence of RVs \( \{y_{0} = 0, y_{1}, y_{2}, \ldots, y_{m}\} \), where

\[
y_{m} = x_{m} + cm \quad \text{implying} \quad y_{m} = y_{m-1} + c + \eta_{m}.
\]

where \( \eta_{m} \) are symmetric i.i.d. jump variables each drawn from the pdf \( f(\eta) \). Clearly, \( y_{m} \) then represents the position of a discrete-time random walker at step \( m \) in the presence of a constant bias \( c \).

In this paper, we are interested in the record statistics of this biased sequence \( \{y_{0} = 0, y_{1}, y_{2}, \ldots, y_{n}\} \). A record happens at step \( m \) if \( y_{m} > \max(y_{0} = 0, y_{1}, y_{2}, \ldots, y_{m-1}) \), i.e. if the position of the biased walker \( y_{m} \) at step \( m \) is bigger than all previous positions, with the convention that the initial position \( y_{0} = 0 \) is counted as a record. Let \( R_{n} \) denote the number of records up to step \( n \). Clearly, \( R_{n} \) is a RV and we denote its distribution by

\[
P(R, n) = \text{Proba.} [R_{n} = R].
\]

We would like to compute the asymptotic properties of this record number distribution \( P(R, n) \) for large \( n \), for arbitrary drift \( c \) and for arbitrary symmetric and continuous jump density \( f(\eta) \) whose Fourier transform \( \hat{f}(k) \) has the small \( k \) behavior as in equation (4) with the index \( 0 < \mu \leq 2 \).

In the absence of a drift, i.e. for \( c = 0 \), the distribution \( P(R, n) \) was computed exactly in [22] for all \( R \) and \( n \), using a renewal property of the record process. Amazingly, the distribution was found to be completely universal, i.e. independent of the jump distribution \( f(\eta) \) (as long as it is symmetric and continuous) for all \( R \) and \( n \) [22]. In particular, for large \( n \), it was shown that \( P(R, n) \) has a scaling form [22]

\[
P(R, n) \approx \frac{1}{\sqrt{n}} g_{0} \left( \frac{R}{\sqrt{n}} \right),
\]

where the universal scaling function

\[
g_{0}(x) = \frac{1}{\sqrt{\pi}} \exp(-x^{2}/4), \quad \text{for} \quad x \geq 0,
\]

is half-Gaussian. Consequently, the mean and the variance of the number of records grow asymptotically as [22]

\[
\langle R_{n} \rangle \approx \frac{2}{\sqrt{\pi}} n^{1/2}, \quad \langle R_{n}^{2} \rangle - \langle R_{n} \rangle^{2} \approx 2 \left( 1 - \frac{2}{\pi} \right) n.
\]

The renewal property of the record process derived originally for the unbiased random walker in [22] was then generalized to the case with a nonzero drift \( c \) in [26]. In particular, the authors of [26] studied in detail the special case of the Cauchy jump distribution \( f_{\text{Cauchy}}(\eta) = 1/\pi (1 + \eta^{2}) \) (which belongs to the \( \mu = 1 \) family of jump densities in equation (4)) and found that the mean number of records \( \langle R_{n} \rangle \) grows algebraically with \( n \) for large \( n \) with an exponent that depends continuously on \( c \) [26]:

\[
\langle R_{n} \rangle \approx \frac{1}{\Gamma(1 + \theta(c))} n^{\theta(c)}, \quad \text{where} \quad \theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c).
\]

In addition, the asymptotic distribution \( P(R, n) \) for large \( n \) was found [26] to have a scaling distribution \( P(R, n) \sim n^{-\theta(c) c_{0}} g_{c}(R_{n}^{\theta(c)}) \) with a nontrivial scaling function \( g_{c}(x) \) which reduces, for \( c = 0 \), to the half-Gaussian in equation (14).
Figure 1. Phase diagram in the $(c, 0 < \mu \leq 2)$ strip depicting five regimes: (I) $0 < \mu < 1$ and $c$ arbitrary, (II) the line $\mu = 1$ and $c$ arbitrary, (III) $1 < \mu < 2$ and $c > 0$, (IV) the semi-infinite line $\mu = 2$ and $c > 0$, and (V) $1 < \mu \leq 2$ and $c < 0$. The persistence $Q(n)$, the record number distribution $P(R, n)$ and the mean ages of the longest and the shortest lasting records exhibit different asymptotic behaviors in these five regimes (see the text).

For jump densities with a finite second moment $\sigma^2$ and in the presence of a nonzero positive drift $c > 0$, the mean number of records $\langle R_n \rangle$ was analyzed in [11] and found to grow linearly with $n$ for large $n$: $\langle R_n \rangle \approx a_2(c) n$, where the prefactor $a_2(c)$ was computed approximately for the Gaussian jump distribution. However, an exact expression of the prefactor for arbitrary jump densities with a finite $\sigma^2$ is missing. In addition, these results were then applied [11] to analyze the record statistics of stock prices from the Standard and Poors 500. The distribution of the record number $P(R, n)$ for large $n$ has not been studied for jump densities with a finite second moment.

In this paper, we present detailed exact results for the asymptotic record number distribution $P(R, n)$ for large $n$, for arbitrary drift $c$ (both positive and negative) and for arbitrary symmetric and continuous jump densities $f(\eta)$ with the Fourier transform $\hat{f}(k)$ having a small $k$ behavior as in equation (4) parametrized by the exponent $0 < \mu \leq 2$. We find a variety of rather rich behaviors for $P(R, n)$ depending on the value of $c$ and the exponent $\mu$. On the strip $(c, 0 < \mu \leq 2)$ (see figure 1), we find five distinct regimes: (I) when $0 < \mu < 1$ with $c$ arbitrary, (II) when $\mu = 1$ and $c$ arbitrary, (III) when $1 < \mu < 2$ and $c > 0$, (IV) when $\mu = 2$ and $c > 0$, and (V) when $1 < \mu \leq 2$ and $c < 0$. In these five regimes, the record statistics behave differently, resulting in different asymptotic forms for the record number distribution $P(R, n)$. The line $\mu = 1$ (regime II) is a critical line on which the record statistics exhibits marginal behavior. These five regimes are summarized in the phase diagram in the $(c, 0 < \mu \leq 2)$ strip in figure 1.

As we will see later, a quantity that plays a crucial role in the study of record statistics is the persistence $Q(n)$ which denotes the probability that the process $y_m$ in equation (11) stays below its initial value $y_0$ up to step $n$:

$$Q(n) = \text{Proba.}[y_i < y_0, \text{for all } i = 1, 2, \ldots, n].$$

Due to the translational invariance of the process, $Q(n)$ does not depend on $y_0$. The persistence probability has been studied quite extensively in recent years in a variety of theoretical and experimental systems [28]. We will see that even for the simple stochastic process $y_m$ representing the position of a discrete-time random walker in the presence of a drift, the persistence $Q(n)$ has a rather rich asymptotic behavior depending on the parameters $\mu$ and $c$. 
Hence, even though here our main interest is in the record statistics, we include the results for the persistence $Q(n)$ as a byproduct.

We also analyze the statistics of waiting times between individual record events. In particular, we are interested in the expected ages of the longest and the shortest lasting records. The age of the longest lasting record is defined as

$$l_{\text{max},n} = \max (l_1, l_2, \ldots, l_R),$$

where $l_i$ is the length of the time interval between the the $i$th and the $(i+1)$th record. Similarly, one defines the age of the shortest lasting record as

$$l_{\text{min},n} = \min (l_1, l_2, \ldots, l_R).$$

In [22], the mean values of $l_{\text{max},n}$ and $l_{\text{min},n}$ were computed exactly for the symmetric random walk with arbitrary jump distribution. It was found that

$$\langle l_{\text{max},n} \rangle \sim C_0 n,$$

where $C_0 \approx 0.626508 \ldots$ is a universal constant independent of the jump distribution. Interestingly, the same constant $C_0$ also appears in other related problems [33, 34]. In contrast, the shortest record exhibits different behavior for large $n$ [22]:

$$\langle l_{\text{min},n} \rangle \sim \sqrt{n/\pi}.$$

In this paper, we generalize these results to the case of a biased random walk and as in the case of record number distribution, we find five different asymptotic behaviors depending on $c$ and $\mu$.

**Summary of the new results.** Let us then summarize the main new results in this paper for the asymptotic behavior of the persistence $Q(n)$, the record number distribution $P(R, n)$ and the extremal ages of records in the five regimes in the $(c, \mu)$ strip mentioned above.

**Regime I** ($0 < \mu < 1$ and $c$ arbitrary). In this regime, we find that the persistence $Q(n)$ decays algebraically for large $n$:

$$Q(n) \approx \frac{B_I}{\sqrt{n}},$$

where the prefactor $B_I$ depends on the details of the jump distribution $f(\eta)$ and the drift $c$ and can explicitly be computed (see equation (78)). The mean record number up to $n$ steps grows asymptotically for large $n$ as

$$\langle R_n \rangle \approx A_I \sqrt{n}.$$

While the growth exponent $1/2$ is universal, i.e. independent of $c$ and the precise form of the jump distribution $f(\eta)$, the prefactor $A_I$ depends on $c$ and on the details of the density $f(\eta)$. In addition, the two prefactors $A_I$ and $B_I$ are related simply via $B_I = 2/(\pi A_I)$. We find the following exact expression for the prefactor $A_I$:

$$A_I = \frac{2}{\sqrt{\pi}} \exp \left[ \frac{1}{\pi} \int_0^\infty \frac{dk}{k} \arctan \left( \frac{\hat{f}(k) \sin (kc)}{1 - \hat{f}(k) \cos (kc)} \right) \right].$$

In the scaling limit when $n \to \infty$ and $R \to \infty$, but with the ratio $R/\sqrt{n}$ being fixed, we find that the distribution $P(R, n)$ approaches the scaling form

$$P(R, n) \approx \frac{2}{A_I \sqrt{n}} g_0 \left( \frac{2R}{A_I \sqrt{n}} \right),$$

where $g_0(x) = \frac{1}{\sqrt{\pi}} \exp(-x^2/4)$.
Averaging over \( R \) evidently reproduces the result in equation (23). Thus, the record number, rescaled by the nonuniversal scale factor \( R \rightarrow R/A_1 \), asymptotically approaches the same universal half-Gaussian scaling distribution as in the driftless case \( c = 0 \) in equation (14).

The statistics of the longest lasting record is completely unaffected by the drift \( c \) in this regime. For the mean value \( \langle l_{\text{max},n} \rangle \), we find that

\[
\langle l_{\text{max},n} \rangle \sim C_1 n,
\]

where the same constant \( C_1 = C_0 \approx 0.626 \) was also found in the unbiased case (see equation (20)). The age of the shortest lasting record is given by

\[
\langle l_{\text{min},n} \rangle \sim D_1 \sqrt{n},
\]

with a prefactor \( D_1 = B_1 \). Therefore, in contrast to \( \langle l_{\text{max},n} \rangle \), \( \langle l_{\text{min},n} \rangle \) slightly differs from the unbiased case and has a prefactor that depends non-trivially on \( c \).

**Regime II (the line \( \mu = 1 \) and \( c \) arbitrary).** On this line, we find that the persistence \( Q(n) \) decays algebraically for large \( n \) but with an exponent that depends continuously on \( c \):

\[
Q(n) \approx \frac{B_{II}}{n^{\theta(c)}},
\]

where the exponent \( 0 \leq \theta(c) \leq 1 \) is given in equation (16). In this sense, the behavior is marginal. The mean record number also grows marginally for large \( n \):

\[
\langle R_n \rangle \approx \frac{A_{II}}{\Gamma[1 + \theta(c)]} n^{\theta(c)},
\]

where the prefactor \( A_{II} = 1/\Gamma[1 - \theta(c)]B_{II} \). The record number distribution exhibits an asymptotic scaling form

\[
P(R, n) \sim \frac{1}{A_{II} n^{\theta(c)}} g_c \left( \frac{R}{A_{II} n^{\theta(c)}} \right),
\]

where one can obtain a formal exact expression (109) and explicit tails of the scaling function \( g_c(x) \) that also exhibits marginal behavior, i.e. depends continuously on \( c \).

Like in regime I we find that the mean age of the longest lasting record grows linearly in \( n \), but this time with a non-trivial \( c \)-dependent prefactor. We find that

\[
\langle l_{\text{max},n} \rangle \sim C_{II} n,
\]

where \( C_{II} \) is given in equation (147). The mean age of the shortest lasting record is more strongly affected by the drift. Here, we find that \( \langle l_{\text{min},n} \rangle \) grows algebraically with \( n \) with an exponent that depends continuously on \( c \):

\[
\langle l_{\text{min},n} \rangle \sim D_{II} n^{1 - \theta(c)},
\]

with \( D_{II} = B_{II} \) as in equation (28) and \( \theta(c) \) as defined in equation (16).

**Regime III (\( 1 < \mu < 2 \) and \( c > 0 \)).** In this regime, the persistence \( Q(n) \) decays for large \( n \) as

\[
Q(n) \approx \frac{B_{III}}{n^{\mu}},
\]

where the prefactor \( B_{III} \) depends on the details of the jump distribution and can be computed (see equation (90)). The mean number of records grows linearly with increasing \( n \)

\[
\langle R_n \rangle \approx a_{\mu}(c) n,
\]

\[\text{Page 8}\]
where the prefactor $a_\mu(c)$ can explicitly be computed (see equation (115)). The record number distribution $P(R, n)$, for large $n$, behaves as

$$P(R, n) \approx \frac{1}{a_\mu(c)n^{1/\mu}} V_\mu \left( \frac{R - a_\mu(c)n}{a_\mu(c)n^{1/\mu}} \right),$$

(35)

where the scaling function $V_\mu(u)$ can exactly be computed and it has a non-Gaussian form with highly asymmetric tails

$$V_\mu(u) \approx A_\mu |u|^{-\mu - 1} \quad \text{as} \quad u \to -\infty$$

(36)

$$\approx c_1 u^{(2-\mu)/2(\mu-1)} \exp[-c_2 u^{\mu/(\mu-1)}] \quad \text{as} \quad u \to \infty,$$

(37)

where $A_\mu$ is the same constant as in equation (8) and the constants $c_1$ and $c_2$ are given explicitly by

$$c_1 = \left[ 2\pi (\mu - 1) (\mu B_\mu)^{1/(\mu-1)} \right]^{-1/2},$$

(38)

$$c_2 = (1 - 1/\mu) (\mu B_\mu)^{-1/(\mu-1)},$$

(39)

where

$$B_\mu = \frac{1}{2 \cos(\mu \pi/2)} > 0 \quad \text{for} \quad 1 < \mu < 2.$$ (40)

Thus, in this regime, while the mean record number grows linearly with $n$, the fluctuations around the mean are anomalous $\sim n^{1/\mu}$ and described by a non-Gaussian distribution.

Also the extremal ages of records have an interesting behavior in this regime. In particular, we find that the average age of the longest lasting record grows like

$$\langle l_{\text{max},n} \rangle \sim C_{\text{III}} n^{\frac{1}{\mu}},$$

(41)

where the constant $C_{\text{III}}$ can explicitly be computed (see equation (150)). On the other hand and in contrast to the results of regime I and II, the mean age of the shortest lasting record converges to a finite value:

$$\langle l_{\text{min},n} \rangle \sim D_{\text{III}} \equiv 1 - a_\mu(c),$$

(42)

which thus depends continuously on $c$. The strongly different $n$ dependence of $\langle l_{\text{max},n} \rangle$ and $\langle l_{\text{min},n} \rangle$ in regimes I and III is a consequence of the fact that while in regime I the asymptotic behavior is dominated by the fluctuations, in regime III the effect of the drift is stronger in the large $n$ limit.

**Regime IV (the semi-infinite line $\mu = 2$ and $c > 0$).** On this semi-infinite line, the variance $\sigma^2$ of the jump pdf is finite. This leads to an exponential tail of the persistence $Q(n)$ for large $n$. More precisely, we show that

$$Q(n) \approx B_{\text{IV}} \frac{1}{n^{3/2}} \exp[-(c^2/2\sigma^2) n],$$

(43)

where the nonuniversal prefactor $B_{\text{IV}}$ can exactly be computed (see equation (96)). We also show that the mean and the variance of the record number both grow linearly for large $n$

$$\langle R_n \rangle \approx a_2(c) n \quad \text{and} \quad (\langle R_n^2 \rangle - \langle R_n \rangle^2) \approx b_2(c) n,$$

(44)

where the amplitudes $a_2(c)$ and $b_2(c)$ are nonuniversal and depend on the details of the jump distribution $f(\eta)$. We provide exact expressions for these amplitudes, respectively, in equations (125) and (128) as well as in appendix B. The distribution of the record number $P(R, n)$ approaches a Gaussian form asymptotically for large $n$

$$P(R, n) \approx \frac{1}{\sqrt{2\pi b_2(c) n}} \exp \left[ -\frac{1}{2b_2(c)n} (R - a_2(c) n)^2 \right].$$

(45)
Thus, in this regime, the mean record number grows linearly with $n$ with normal Gaussian fluctuations $\sim n^{1/2}$ around the mean.

It is interesting to see that the asymptotic behavior of $\langle l_{\text{max},n} \rangle$ in regime IV is qualitatively different from regime III. Here, we find that $\langle l_{\text{max},n} \rangle$ grows only logarithmically with $n$ for $n \to \infty$:

$$\langle l_{\text{max},n} \rangle \sim C_{\text{IV}} \ln n,$$

with an $n$-independent constant $C_{\text{IV}} = \frac{2\sigma^2}{c^2}$. Like in regime III, the average age of the shortest lasting record approaches a (different) constant value depending on $c$:

$$\langle l_{\text{min},n} \rangle \sim D_{\text{IV}} = 1 - a_2(c),$$

which depends continuously on $c$.

**Regime V** ($1 < \mu \leq 2$ and $c < 0$). In this case, the walker predominantly moves toward the negative axis due to the drift. Consequently, the events where the walker crosses the origin from the negative to the positive side become extremely rare. As a result, with a finite probability, the walker stays forever on the negative side. Thus, the persistence $Q(n)$ approaches a constant for large $n$:

$$Q(n) \to \alpha_\mu(c).$$

Similarly, the occurrence of the records (with positive record values) are also rare. Subsequently, we find that the mean record number also approaches a constant for large $n$:

$$\langle R_n \rangle \to \frac{1}{\alpha_\mu(c)},$$

where the constant $\alpha_\mu(c)$ is given by

$$\alpha_\mu(c) = \alpha_\mu(|c|) \text{ for } 1 < \mu < 2,$$

$$= a_2(|c|) \text{ for } \mu = 2,$$

where $\alpha_\mu(c)$ and $a_2(c)$ are precisely the amplitudes of the linear growth of the mean record number, respectively, in regimes III and IV (respectively in equations (34) and (44)). An explicit expression for $\alpha_\mu(c)$ is given in equation (101). The record number distribution $P(R, n)$ also approaches a steady state, i.e. $n$-independent distribution as $n \to \infty$. This distribution has a purely geometric form

$$P(R, n \to \infty) = \alpha_\mu(c) \{1 - \alpha_\mu(c)\}^{R-1}.$$

Physically, this result is easy to understand because for $c < 0$ and $\mu > 1$, the walker typically moves away from the origin on the negative side with very rare and occasional excursions to the positive side caused by rare large jumps. As a result, the occurrence of a record is like a Poisson process which eventually leads to a geometric distribution as in equation (52).

In this regime, the statistics of the longest and the shortest lasting records are particularly simple. Since the record number is finite, the longest lasting record will grow linearly in $n$:

$$\langle l_{\text{max},n} \rangle \sim C_V n, \quad C_V = 1.$$

For the shortest lasting record, we find a similar behavior:

$$\langle l_{\text{min},n} \rangle \sim \alpha_\mu(c) n,$$

with the same $c$-dependent constant $\alpha_\mu(c)$ as in equation (48). Here, the main contributions to these quantities come from trajectories that never cross the origin and stay negative for all $n$. 
The five regimes in the $(c, 0 < \mu \leq 2)$ strip are depicted in figure 1. As mentioned above, the line $\mu = 1$ is a special ‘critical’ line with marginal exponents that depend continuously on the drift $c$. It is not difficult to understand physically why $\mu = 1$ plays a special role. Indeed, writing $y_n = x_n + cn$, where $x_n$ represents a symmetric random walk, we see that the two terms $x_n$ and $cn$ compete with each other for large $n$. Since $x_n \sim n^{1/\mu}$ for $0 < \mu < 2$ (see equation (7)), it is clear that for $0 < \mu < 1$, the term $x_n$ dominates over the drift and the presence of a nonzero drift only leads to subleading asymptotic effect. In contrast, for $\mu > 1$, the drift term starts to play an important role in governing the asymptotic record statistics. In the region $1 < \mu < 2$ and $c > 0$ (regime III), while the mean record number increases linearly with $n$ due to the dominance of the drift term, the typical fluctuation around the mean is still dominated by the $x_n \sim n^{1/\mu}$ term (see equation (35)). However, when $\mu = 2$ and $c > 0$ (regime IV), the drift term completely dominates over the $x_n$ term leading to Gaussian fluctuations around the mean. This competition between $x_n$ and $cn$ thus leads to (i) a ‘phase transition’ in the asymptotic behavior of record statistics of $y_n$ at the critical value $\mu = 1$ and (ii) an anomalous region with non-Gaussian fluctuations around the mean in the region $1 < \mu < 2$ and $c > 0$.

3. Record number distribution via a renewal property and the generalized Sparre Andersen theorem

The idea of using the renewal property of random walks to compute the distribution of record number was first used in [22] for symmetric random walks and was subsequently generalized to biased random walks [26]. We briefly summarize below the main idea.

Consider the random sequence $\{y_0, y_1, y_2, \ldots, \}$ representing the successive positions of a discrete-time biased random walker evolving via equation (11), starting from an arbitrary initial position $y_0$. Consider first the persistence $Q(n)$ defined in equation (17). Let us also define

$$F(n) = \text{Proba.} \{y_1 < y_0, y_2 < y_0, \ldots, y_{n-1} < y_0, y_n > y_0\},$$

which denotes the probability that the walker crosses its initial position $y_0$ from below for the first time at step $n$. Clearly

$$F(n) = Q(n-1) - Q(n).$$

It is also useful to define the generating functions (GFs)

$$\tilde{Q}(z) = \sum_{n=0}^{\infty} Q(n) z^n, \quad \tilde{F}(z) = \sum_{n=1}^{\infty} F(n) z^n.$$

Using the relation in equation (56), it follows that

$$\tilde{F}(z) = 1 - (1-z)\tilde{Q}(z).$$

Consider now any realization of the sequence $\{y_0 = 0, y_1, y_2, \ldots, y_n\}$ up to $n$ steps and let $R_n$ be the number of records in this realization. Let $\tilde{I} = \{l_1, l_2, \ldots, l_N\}$ denote the time intervals between successive records in this sequence (see figure 2). Clearly, $l_i$ denotes the age of the $i$th record, i.e. the time up to which the $i$th record survives. The last record, i.e. the $N$th record, stays a record till step $n$. Let $P(\tilde{I}, R|n)$ denote the joint distribution of the ages and the number of records up to step $n$. Using the two probabilities $Q(n)$ and $F(n)$ defined earlier and the fact that the successive intervals between records are statistically independent due to the Markov nature of the process, it follows immediately that

$$P(\tilde{I}, R|n) = F(l_1)F(l_2) \cdots F(l_N) Q(l_N) \delta_{\sum l_i = 1, l_i, N}.$$
where the Kronecker delta enforces the global constraint that the sum of the time intervals equals \( n \). The fact that the last record, i.e. the \( R \)th record, is still surviving as a record at step \( n \) indicates that the distribution \( Q(l_R) \) of \( l_R \) is different from the preceding ones. It is easy to check that \( P(\bar{l}, R|n) \) is normalized to unity when summed over \( \bar{l} \) and \( R \). The record number distribution \( P(R, n) = \sum_{\bar{l}} P(\bar{l}, R|n) \) is just the marginal of the joint distribution when one sums over the interval lengths. Due to the presence of the delta function, this sum is most easily carried out by considering the GF with respect to \( n \). Thus, the basic object is the GF \( \bar{Q}(z) \).

The record number distribution \( P(R, n) = \sum_{\bar{l}} P(\bar{l}, R|n) \) where we used the relation in equation (58). Note that, by definition, \( R \leq (n + 1) \), i.e. \( P(R, n) = 0 \) if \( n < R - 1 \). Hence, the sum in equation (60) actually runs from \( n = R - 1 \) to \( \infty \).

Fig. 2. A typical realization of the biased random walk sequence \( \{y_0 = 0, y_1, y_2, \ldots, y_n\} \) of \( n \) steps with \( R \) records. Each record is represented by a filled circle. The set \( \{l_1, l_2, \ldots, l_{R-1}\} \) represents the time intervals between the successive records and \( l_R \) is the age of the last record which is still a record till step \( n \).

Thus, the basic object is the GF \( \bar{Q}(z) \). Once this is determined, one can, at least in principle, compute other quantities such as the statistics of records or their ages using the fundamental renewal equation (60). Fortunately, there exists a beautiful combinatorial identity first derived by Sparre Andersen [23] that allows one to compute \( \bar{Q}(z) \):

\[
\bar{Q}(z) = \sum_{n=0}^{\infty} Q(n) z^n = \exp \left[ \sum_{n=1}^{\infty} \frac{z^n}{n} p(n) \right],
\]

where \( p(n) = \text{Prob}. \{y_n < 0\} \). Using the relation \( y_n = x_n + cn \), where \( x_n \) represents the symmetric random walk at step \( n \) in equation (2), one obtains \( p(n) = \text{Prob}. \{x_n < -cn\} \). Then, using the pdf \( P_n(x) \) of the symmetric walk \( x_n \) at step \( n \) in equation (6), one obtains

\[
p(n) = \text{Prob}. \{x_n < -cn\} = \int_{-\infty}^{-cn} P_n(x) \, dx = \int_{-cn}^{\infty} P_n(x) \, dx,
\]

where, in obtaining the last equality, we used the symmetry \( P_n(x) = P_n(-x) \). Substituting this expression of \( p(n) \) into equation (61) results

\[
\bar{Q}(z) = \sum_{n=0}^{\infty} Q(n) z^n = \exp \left[ \sum_{n=1}^{\infty} \frac{z^n}{n} \int_{-cn}^{\infty} P_n(x) \, dx \right].
\]
Equation (63), with \( P_n(x) \) given in equation (6), determines \( \hat{Q}(z) \) in terms of the Fourier transform \( \hat{f}(k) \) of the jump distribution \( f(\eta) \). Subsequently, equation (60) then determines, in principle, the record number distribution \( P(R, n) \). In the driftless case \( c = 0 \), great simplification occurs, since by symmetry \( \int_0^\infty P_n(x)dx = 1/2 \). This results, from equation (63), \( \hat{Q}(z) = 1/\sqrt{1 - z} \). This is completely universal as all the dependence on the jump distribution \( f(\eta) \) drops out. Subsequently, equation (60) provides, for \( c = 0 \), the universal result for the record number distribution \cite{22}

\[
\sum_{n=0}^{\infty} P(R, n) z^n = \frac{(1 - \sqrt{1 - z})^{\mu-1}}{\sqrt{1 - z}},
\]

which, when inverted, yields \cite{22} for large \( n \) the scaling behavior in equation (13) with the scaling function given by the half-Gaussian form in equation (14).

However, in the presence of a nonzero bias \( c \), extraction of the precise large \( n \) behavior of \( P(R, n) \) from the set of equations (60), (63) and (6) is more complicated. For the special case of the Cauchy distribution, this was performed in \cite{26} which showed nontrivial behavior. The rest of this paper is devoted precisely to this technical task of extracting the large \( n \) behavior of \( P(R, n) \) for a general jump distribution \( f(\eta) \) and we will see that a variety of rather rich asymptotic behavior emerges depending on the value of the drift \( c \) and the exponent \( \mu \) characterizing the small \( k \) behavior of \( f(k) \) in equation (4).

Before finishing this section, let us remark that from equation (60) one can also compute the GFs of the moments of the number of records. For example, multiplying equation (60) by \( R \), summing over \( R \) and using the identity \( \sum_{n=0}^{\infty} n x^{\mu-1} = 1/(1-x)^2 \), we obtain for the first moment

\[
\sum_{n=0}^{\infty} \langle R_n \rangle z^n = \frac{1}{(1-z)^2 \hat{Q}(z)}.
\]

Similarly, multiplying equation (60) by \( R^2 \) and summing over \( R \), one obtains for the second moment

\[
\sum_{n=0}^{\infty} \langle R_n^2 \rangle z^n = \frac{2 - (1-z)\hat{Q}(z)}{(1-z)^3 \hat{Q}^2(z)}.
\]

We will use these two results later in section 4.2.

4. Asymptotic behavior of persistence \( Q(n) \) for large \( n \)

The persistence \( Q(n) \), i.e. the probability that the process \( y_n \) stays below its initial value \( y_0 \) up to step \( n \), and its GF \( \hat{Q}(z) \) are the key ingredients to determine the record number distribution \( P(R, n) \) via equation (60). Apart from the key role as an input for the record statistics, the persistence \( Q(n) \) for this process is, by itself, an interesting quantity to study. We will see in this section that even for the simple stochastic process \( y_n \), representing the position of a discrete-time random walker in the presence of a drift, the persistence \( Q(n) \) has a rather rich asymptotic behavior depending on the parameters \( \mu \) and \( c \). Before getting into the details of the derivation, it is useful to summarize these asymptotic results. We find that for large \( n \), the persistence \( Q(n) \) has the following asymptotic tails depending on \( \mu > 0 \) and \( c \):

\[
Q(n) \sim B_I n^{-1/2} \quad \text{for} \quad 0 < \mu < 1 \quad \text{and} \quad c \text{ arbitrary (regime I)},
\sim B_{II} n^{-\mu} \quad \text{for} \quad \mu = 1 \quad \text{and} \quad c \text{ arbitrary (regime II)},
\sim B_{III} n^{-\mu} \quad \text{for} \quad 1 < \mu < 2 \quad \text{and} \quad c > 0 \quad \text{(regime III)},
\sim B_{IV} n^{-3/2} \exp[-(c^2/2\sigma^2) n] \quad \text{for} \quad \mu = 2 \quad \text{and} \quad c > 0 \quad \text{(regime IV)},
\sim \alpha_{\mu}(c) \quad \text{for} \quad 1 < \mu \leq 2 \quad \text{and} \quad c < 0 \quad \text{(regime V)},
\]

(Eq. 67)
Figure 3. Numerical simulations of the persistence $Q(n)$, i.e. the probability that a random walker with a bias $c$ stays below its initial position up to step $n$. We considered four different Lévy-stable jump distributions characterized, respectively, by the Lévy index $\mu = 0.5, 1, 1.5$ and $\mu = 2$ (in the last case it is just Gaussian jump distribution). In all cases, we had a constant positive bias $c = 1$ and the data were obtained by averaging over $10^7$ samples. For comparison, we also present the result for the unbiased case ($c = 0$) with a Gaussian jump distribution (the top curve). The thin dashed lines give our analytical predictions from equation (67) with fitted prefactors $B_I$, $B_{II}$, $B_{III}$ and $B_{IV}$. For the $\mu = 1$ case, we used $\theta (c = 1) \approx 0.7498\ldots$

where the prefactors $B_I$, $B_{II}$, $B_{III}$, $B_{IV}$ can be explicitly computed. In regime V, $\alpha_{\mu}(c)$ is a constant independent of $n$ that can also be computed explicitly (see equation (101) and by appendix B for $\alpha_2(\mu)$). The exponent $\theta(c)$ depends continuously on $c$ and is given in equation (16) (see also equation (80)). In figure 3, these results are numerically confirmed for regimes I–IV.

To derive these asymptotic behaviors of $Q(n)$ for large $n$, we start with the key result in equation (63). Using Cauchy’s inversion formula in the complex $z$-plane, one can write

$$Q(n) = \int_{C_0} \frac{dz}{2\pi i} \frac{1}{z^{n+1}} \tilde{Q}(z) \quad \text{with} \quad \tilde{Q}(z) = \exp \left[ \sum_{n=1}^{\infty} \frac{z^n}{n} \int_{cn}^{\infty} P_n(x) \, dx \right],$$

where the contour $C_0$ encircles the origin 0 and is free of any singularity of $\tilde{Q}(z)$ (see figure 4). Let $z^*$ denote the singularity of $\tilde{Q}(z)$ on the real axis closest to the origin. Then, one can deform the contour $C_0$ to $C_1$ (see figure 4) such that the vertical part of $C_1$ is located just on the left of $z^*$ and the circular part has a radius $r_1$. By taking the $r_1 \to \infty$ limit, it follows from equation (68) that for large $n$, the contribution from the circular part vanishes exponentially. Thus, for large $n$, the leading contribution comes from the vertical part of $C_1$, i.e. the imaginary axis located just on the left of $z^*$. Next, we make a change of variable $z = e^{-\mu s}$ and define

$$\tilde{q}(s) = \tilde{Q}(z = e^{-s}) = \sum_{n=0}^{\infty} Q(n) e^{-\mu s} = \exp[W_{c,\mu}(s)],$$

where

$$W_{c,\mu}(s) = \sum_{n=1}^{\infty} \frac{e^{-\mu s}}{n} \int_{cn}^{\infty} P_n(x) \, dx.$$
Using this expression in the integrand in equation (68) and retaining only the contribution from the vertical part of the contour $C_1$ for large $n$, we obtain

$$Q(n) \approx \int_{s^*-i\infty}^{s^*+i\infty} \frac{ds}{2\pi i} e^{sn} \exp[W_{c,\mu}(s)],$$

(71)

where $W_{c,\mu}(s)$ is given in equation (70) and $s^* = -\ln(z^*)$ is the singularity of $\tilde{q}(s) = \exp[W_{c,\mu}(s)]$ on the real axis closest to $s = 0$. Identifying the integral on the rhs of equation (71) as a standard Bromwich integral in the complex $s$ plane, we see that for large $n$, $Q(n)$ is essentially given by the inverse Laplace transform of the function $\tilde{q}(s)$.

To make further progress, we need to first identify the position of the singularity $s^*$ of $W_{c,\mu}(s)$ and then analyze the dominant contribution in the Bromwich integral coming from the neighborhood of $s^*$ for large $n$. We see below that the singular behavior of $W_{c,\mu}(s)$ as a function of $s$ depends on the parameters $c$ and $\mu > 0$ and there are essentially five regimes in the $(c, 0 < \mu \leq 2)$ strip as shown in figure 1. Below we discuss these regimes separately.

4.1. Regime I: $0 < \mu < 1$ and $c$ arbitrary

To analyze the leading singularity of $W_{c,\mu}(s)$ as a function of $s$ in this regime, it is first convenient to use the normalization condition $\int_{-\infty}^{\infty} P_n(x) \, dx = 1$ and the symmetry $P_n(x) = P_n(-x)$ to rewrite

$$\int_{-\infty}^{\infty} P_n(x) \, dx = \frac{1}{2} - \int_{0}^{\infty} P_n(x) \, dx.$$  (72)

Substituting this into equation (70) results

$$W_{c,\mu}(s) = -\frac{1}{2} \ln(1 - e^{-s}) - \sum_{n=1}^{\infty} \frac{e^{-sn}}{n} \int_{0}^{\infty} P_n(x) \, dx.$$  (73)

Now, as $s \to 0$, the sum in equation (73) converges to a constant for $0 < \mu < 1$:

$$S_0 = \sum_{n=1}^{\infty} \frac{1}{n} \int_{0}^{\infty} P_n(x) \, dx.$$  (74)

To see this, let us see how the integral $\int_{0}^{\infty} P_n(x) \, dx$ behaves for large $n$. For large $n$, we can use the scaling form for $P_n(x)$ in equation (7). One finds that $\int_{0}^{\infty} P_n(x) \, dx \to \int_{0}^{\infty} L_\mu(y) \, dy$ as $n \to \infty$. For $0 < \mu < 1$, clearly this integral decreases leading to the convergence of the series in equation (74). Thus, the leading singularity of $W_{c,\mu}(s)$ occurs near $s = s^* = 0$, where it behaves as

$$W_{c,\mu}(s) \approx -\frac{1}{2} \ln(s) - S_0.$$  (75)
Substituting this result into equation (70) results
\[ \tilde{q}(s) = \sum_{n=0}^{\infty} Q(n) e^{-sn} \xrightarrow{s \to 0} e^{-S_0} \sqrt{s}. \]  
(76)

We now substitute this singular behavior of the integrand into equation (71) after setting \( s^* = 0 \) and perform the standard Bromwich integral (one can use the fact that the inverse Laplace transform \( LT^{-1}[s^{-1/2}] = 1/\sqrt{\pi s} \))
\[ Q(n) \xrightarrow{n \to \infty} B_1 \sqrt{n}, \]  
(77)
where the prefactor \( B_1 \) is given by
\[ B_1 = e^{-S_0} \frac{1}{\sqrt{\pi}} \exp\left[ -\sum_{n=1}^{\infty} \frac{1}{n} \int_0^c P_n(x) \, dx \right]. \]  
(78)

4.2. Regime II: \( \mu = 1 \) and \( c \) arbitrary

The case \( \mu = 1 \) is rather special and marginal as we demonstrate now. Consider the sum \( W_{c,1}(s) \) in equation (70). In this case, it follows from equation (7) that \( P_n(x) \to (1/n) L_1(x/n) \) as \( n \to \infty \), where \( L_1(y) = 1/[\pi(1+y^2)] \) for all \( y \) and, hence, is integrable. Thus, the integral \( \int_0^\infty P_n(x) \, dx \) converges to a constant for large \( n \):
\[ \int_0^\infty P_n(x) \, dx \xrightarrow{n \to \infty} \int_0^c L_1(y) \, dy = 1 - \theta(c), \]  
(79)
where
\[ \theta(c) = \int_{-\infty}^c L_1(y) \, dy = \frac{1}{2} + \frac{1}{\pi} \arctan(c). \]  
(80)
Hence, the \( n \)th term of the sum in \( W_{c,1}(s) \) behaves, for large \( n \), as \( T_n \to (1 - \theta(c)) e^{-s_n/n} \).

Consequently, the sum \( W_{c,1}(s) = \sum_{n \geq 1} T_n \) has a singularity at \( s = s^* = 0 \). The leading asymptotic behavior near this singularity reads
\[ W_{c,1}(s) \xrightarrow{s \to 0} -(1 - \theta(c)) \ln(s) - \gamma_0, \]  
(81)
where \( \gamma_0 \) is a constant that depends on the details of \( P_n(x) \), in particular on the difference between \( P_n(x) \) and its large \( n \) asymptotic form \( (1/n)L_1(x/n) \) for finite \( n \):
\[ \gamma_0 = \sum_{n=1}^{\infty} \left[ 1 - \theta(c) - \int_0^c P_n(x) \, dx \right]. \]  
(82)
Using this result on the rhs of equation (70) results
\[ \tilde{q}(s) = \xrightarrow{s \to 0} e^{-\gamma_0} s^{1/\theta(c)}. \]  
(83)
Substituting this result into the Bromwich contour in equation (71) (after setting \( s^* = 0 \)) and performing the Bromwich integral results
\[ Q(n) \xrightarrow{n \to \infty} B_{II} n^{\theta(c)}, \]  
(84)
where
\[ B_{II} = \frac{e^{-\gamma_0}}{\Gamma[1 - \theta(c)]} \quad \text{and} \quad \theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c), \]  
(85)
and \( \gamma_0 \) is given in equation (82).
Thus, for \( \mu = 1 \), the persistence \( Q(n) \) decays algebraically for large \( n \) but with an exponent \( \theta(c) \) that continuously depends on the drift \( c \). In this sense, the line \( \mu = 1 \) is marginally critical. The exponent \( \theta(c) \) in equation (85) increases continuously with \( c \) from \( \theta(c \to -\infty) = 0 \) to \( \theta(c \to \infty) = 1 \).

Let us end this subsection with the following remark on the special case of pure Cauchy jump distribution: \( f_{\text{Cauchy}}(\eta) = 1/\pi (1 + \eta^2) \). As mentioned before, the record statistics for this case was studied in detail in [26]. For the Cauchy case, it is well known that \( P_n(x) = (1/n)f_{\text{Cauchy}}(x/n) = (1/n)\zeta_1(x/n) \) for all \( n \). As a result, it follows from equation (82) that the constant \( \gamma_0 = 0 \) in this case. However, in the general \( \mu = 1 \) case (not necessarily the Cauchy case), \( \gamma_0 \) is generically nonzero. Thus, while the persistence exponent \( \theta(c) = 1/2 + \frac{\pi}{2} \arctan(c) \) is universal for all jump densities belonging to the \( \mu = 1 \) case, the amplitude \( B_\Pi \) is nonuniversal and depends on the details of the jump density.

4.3. Regime III: \( 1 < \mu < 2 \) and \( c > 0 \)

To analyze the singular behavior of the sum \( W_{c,\mu}(s) \) in equation (70) in this regime, we consider the \( n \)th term of the sum \( T_n = (e^{-s\eta}/n) \int_0^\infty P_n(x) \text{d}x \) and substitute, for large \( n \), the scaling behavior of \( P_n(x) \) into equation (7). This results \( T_n \approx (e^{-s\eta}/n) \int_0^\infty \zeta_1(x) \zeta_\mu(y) \text{d}y \). For \( 1 < \mu < 2 \), the lower limit of the integral in \( T_n \) becomes large as \( n \to \infty \) and we can use the tail behavior in equation (8) to estimate: \( T_n \approx (A_\mu/\mu e^{\mu \gamma}) e^{-s\eta}/n^\mu \) for large \( n \). Hence, the sum \( W_{c,\mu}(s) = \sum_{n=1}^\infty T_n \) clearly converges to a constant \( W_{c,\mu}(0) \) as \( s \to 0 \). For small \( s \), one can replace the sum by an integral and estimate exactly the first singular correction to this constant. This results

\[
W_{c,\mu}(s) \rightarrow W_{c,\mu}(0) - B_\mu s^{\mu-1},
\]

where the constant \( B_\mu = A_\mu \Gamma(2-\mu)/[\mu(\mu-1)e^{\mu \gamma}] \). Using the exact expression of \( A_\mu \) from equation (8) and simplifying, one finds \( B_\mu = -1/[2\cos(\mu \pi/2)] > 0 \) as in equation (40). Note also that from the definition in (70)

\[
\tilde{q}(0) = \exp[W_{c,\mu}(0)] = \exp \left[ \sum_{n=1}^\infty \frac{1}{n} \int_0^\infty P_n(x) \text{d}x \right].
\]

Substituting the small \( s \) behavior from equation (86) into equation (70) results

\[
\tilde{q}(s) \rightarrow \tilde{q}(0)[1 - B_\mu s^{\mu-1} + \cdots].
\]

Substituting this singular behavior of \( \tilde{q}(s) = \exp[W_{c,\mu}(s)] \) into the Bromwich integral in equation (71) (upon setting \( s^* = 0 \)) and performing the integral by the standard method provides the following large \( n \) power-law tail for \( Q(n) \):

\[
Q(n) \rightarrow \frac{B_\Pi}{(\mu-1)} \text{as } n \to \infty,
\]

where the prefactor \( B_\Pi \) is given by

\[
B_\Pi = \frac{(\mu-1)B_\mu}{\Gamma(2-\mu)e^{\mu \gamma}} \tilde{q}(0) = -\frac{(\mu-1)}{2\cos(\mu \pi/2)\Gamma(2-\mu)e^{\mu \gamma}} \exp \left[ \sum_{n=1}^\infty \frac{1}{n} \int_0^\infty P_n(x) \text{d}x \right].
\]

4.4. Regime IV: \( \mu = 2 \) and \( c > 0 \)

In this regime, the leading singularity \( s^* \) of \( W_{c,\mu}(s) \) occurs not at \( s = 0 \), but at \( s = s^* = -s_1 \), where \( s_1 = c^2/2\sigma^2 \). To see this, let us again consider the \( n \)th term of the sum \( W_{c,\mu}(s) \), i.e. \( T_n = (e^{-s\eta}/n) \int_0^\infty P_n(x) \text{d}x \). For large \( n \), \( P_n(x) \) now has the Gaussian scaling form in
equation (10) due to the central limit theorem. Substituting this Gaussian form and carrying out the integration, one obtains

\[ T_n \to \frac{e^{-m}}{2n} \text{erfc} \left( \frac{c}{\sigma \sqrt{2}} \sqrt{n} \right), \quad \text{where} \quad \text{erfc}(y) = \frac{2}{\sqrt{\pi}} \int_y^{\infty} e^{-x^2} \, dx. \]  

(91)

Using the asymptotic behavior \( \text{erfc}(y) \approx e^{-y^2}/\sqrt{\pi} \) for large \( y \), one finds that

\[ T_n \to \frac{\sigma}{c \sqrt{2\pi}} \frac{e^{-(s+s_1)n}}{n^{3/2}}, \quad \text{where} \quad s_1 = \frac{c^2}{2\sigma^2}. \]  

(92)

Consequently, the sum \( W_{c,\mu}(s) = \sum_{n \geq 1} T_n \) actually, while perfectly analytic near \( s = 0 \), has a singularity near \( s = s^* = -s_1 \). Close to this singular value, by taking the limit \( s + c^2/2\sigma^2 \to 0 \) whereby replacing the sum by an integral over \( n \), one finds the following leading singular behavior of \( W_{c,\mu}(s) \) near \( s = -s_1 \):

\[ W_{c,\mu}(s) \to W_{c,\mu}(-s_1) - \sqrt{2} \frac{\sigma}{c} \sqrt{s + s_1}, \]  

(93)

where \( W_{c,\mu}(-s_1) \) is just a constant. Substituting this leading singular behavior into the rhs of equation (70) results

\[ \tilde{q}(s) \to e^{W_{c,\mu}(-s_1)} \left[ 1 - \sqrt{2} \frac{\sigma}{c} \sqrt{s + s_1} + \cdots \right]. \]  

(94)

We set \( s^* = -s_1 \) in the Bromwich contour in equation (71), substitute the singular behavior of \( \tilde{q}(s) \) into equation (94) and perform the Bromwich integral to obtain

\[ Q(n) \to \frac{B_{IV}}{n^{5/2}} e^{-s_1n}, \quad \text{where} \quad s_1 = \frac{c^2}{2\sigma^2}. \]  

(95)

and the prefactor

\[ B_{IV} = \frac{\sigma e^{W_{c,\mu}(-s_1)}}{c \sqrt{2\pi}} = \frac{\sigma}{c \sqrt{2\pi}} \exp \left[ \frac{c^2}{n} \int_{c^{2/n}}^{\infty} P_n(x) \, dx \right]. \]  

(96)

Thus, contrary to regimes I, II and III, here the persistence \( Q(n) \) has a leading exponential tail (modulated by a power law \( n^{-3/2} \)).

4.5. Regime V: \( 1 < \mu \leq 2 \) and \( c < 0 \)

In this regime, \( c = -|c| < 0 \) and \( \mu > 1 \). It is convenient, using the normalization condition \( \int_{-\infty}^{\infty} P_n(x) \, dx = 1 \), to first re-express the sum \( W_{c,\mu}(s) \) in equation (70) as

\[ W_{c,\mu}(s) = \sum_{n=1}^{\infty} \frac{e^{-m}}{n} \int_{-|c|}^{\infty} P_n(x) \, dx = \sum_{n=1}^{\infty} \frac{e^{-m}}{n} \left[ 1 - \int_{-|c|}^{\infty} P_n(x) \, dx \right]. \]  

(97)

Performing the sum, and using the definition of \( W_{c,\mu}(s) \) in equation (70), one obtains

\[ W_{c,\mu}(s) = -\ln(1 - e^{-s}) - W_{c,\mu}(0). \]  

(98)

For \( \mu > 1 \), \( W_{c,\mu}(0) \) is constant as was demonstrated in the previous two subsections. Hence, one obtains from equation (98) the leading singular behavior for small \( s \):

\[ W_{c,\mu}(s) \to -\ln(s) - W_{c,\mu}(0), \]  

(99)

which yields, via equation (70),

\[ \tilde{q}(s) \to \frac{\exp[-W_{c,\mu}(0)]}{s}. \]  

(100)
Thus, in this regime, the leading singularity of \( \tilde{q}(s) \) occurs at \( s = s^* = 0 \). Setting \( s^* = 0 \) and the result (100) into the Bromwich integral in equation (71) results in

\[
Q(n) \xrightarrow{n \to \infty} \alpha_n(c) = \exp[-W_{|c|,\mu}(0)] = \exp \left[ -\sum_{n=1}^{\infty} \frac{1}{n} \int_{\mid c \mid n}^{\infty} P_n(x) \, dx \right].
\] (101)

The fact that the persistence \( Q(n) \) approaches to a constant for large \( n \) in this regime can physically be understood because for \( c < 0 \) and \( \mu > 1 \), a finite fraction of trajectories escape to \(-\infty\) as \( n \to \infty \).

5. Asymptotic record number distribution \( P(R, n) \) for large \( n \)

In this section, we analyze the asymptotic large \( n \) properties of the mean record number \( \langle R_n \rangle \) and its full distribution \( P(R, n) \) for arbitrary \( c \) by analyzing the set of equations (6), (60), (63) and (65) with arbitrary jump distribution \( f(\eta) \). Consider first the mean record number. As in section 4, we invert equation (65) by using the Cauchy inversion formula, deform the contour (as in figure 4), keep only the vertical part of the contour \( C_1 \) for large \( n \) and finally make the substitution \( z = e^{-i} \) to obtain the following Bromwich formula:

\[
\langle R_n \rangle \approx \int_{s = -i \infty}^{s = +i \infty} ds \frac{e^{st}}{2\pi i} \frac{1}{(1 - e^{-s})^2 \tilde{q}(s)}.
\] (102)

where \( \tilde{q}(s) \) is given in equations (69) and (70) and its small \( s \) properties have already been analyzed in section 4 in different regimes in the \( c, 0 < \mu \leq 2 \) strip. As in section 4, \( s^* \) denotes the singularity of \( \tilde{q}(s) \) on the real line in the complex plane that is closest to the origin at \( s = 0 \).

Similarly, the record number distribution is obtained by inverting equation (60) in the same way

\[
P(R, n) \approx \int_{s = -i \infty}^{s = +i \infty} ds \frac{e^{st}}{2\pi i} \tilde{q}(s) \left[ 1 - (1 - e^{-s})\tilde{q}(s) \right]^{R-1}.
\] (103)

In this section, we use the already derived results for \( \tilde{q}(s) \) in section 4 and analyze the asymptotic behavior of \( \langle R_n \rangle \) and \( P(R, n) \), respectively, in equations (102) and (103) in different regimes of the \( (c, 0 < \mu \leq 2) \) strip and on the critical line \( \mu = 1 \).

5.1. Regime I: \( 0 < \mu < 1 \) and \( c \) arbitrary

Let us first consider the asymptotic behavior of the mean number of records \( \langle R_n \rangle \) for large \( n \) in this regime. Consider the Bromwich integral in equation (102). For large \( n \), this integral can be shown to be dominated by the small \( s \) region of the integrand. Taking the \( s \to 0 \) limit in the integrand, substituting the result (76) into the rhs of equation (102) and performing the Bromwich integral, we obtain the leading asymptotic behavior for large \( n \):

\[
\langle R_n \rangle \approx A_1 \sqrt{n}, \quad \text{where} \quad A_1 = \frac{2}{\sqrt{\pi}} e^{S_0} = \frac{2}{\sqrt{\pi}} \exp \left[ -\frac{1}{n} \int_{0}^{\infty} P_n(x) \, dx \right].
\] (104)

Comparing this to the amplitude of persistence in equation (78), we see that the two prefactors are related simply via \( B_1 = 2/(\pi A_1) \). The prefactor \( A_1 \) can further be expressed explicitly in terms of the Fourier transform of the jump distribution \( \hat{f}(k) \) as in equation (24). This is shown in appendix A where we also compute the asymptotic behavior of \( A_1 \) for large \( |c| \) (see equation (A.9)). In figure 5(a), we compare this result for \( \langle R_n \rangle \) to numerical simulations. The numerical results for \( n \gg 1 \), \( \langle R_n \rangle / \sqrt{n} \) agree nicely with our analytical values for \( A_1 (c) \).

Next, we turn to \( P(R, n) \) in the limit of large \( n \). To extract the scaling behavior of \( P(R, n) \) from equation (103), we substitute into the rhs the small \( s \) behavior of \( \tilde{q}(s) \) from equation (76)
and use the notation $e^{-S_0} = (2/\sqrt{n}) A_1$. The appropriate scaling limit is clearly $R \to \infty$, $s \to 0$ but keeping the product $\sqrt{s} R$ fixed. Taking this limit in equation (103) results

$$P(R, n) \approx \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} e^{sn} \frac{2}{A_1 \sqrt{n} \pi} \exp \left[ -\frac{2}{A_1 \sqrt{n} \pi} \sqrt{s} R \right].$$

One can simply evaluate the Bromwich integral by using the identity $LT_{-\partial}^{-\infty} [e^{-hR/\sqrt{s}}] = e^{-hR^2/4n} / \sqrt{s} \pi$. This leads to the asymptotic result announced in equation (25) in the scaling limit $n \to \infty, R \to \infty$ with the ratio $R/\sqrt{n}$ fixed. In figure 5(b), we numerically computed the rescaled distribution $A_1 \sqrt{n} P(R, n)/2$ as a function of $2R/A_1 \sqrt{n}$ and compared it with $g_0(x)$ given in equation (25). The figure confirms that in regime I, the record number has a half-Gaussian distribution with a width that depends non-trivially on the drift $c$ and the Lévy index $\mu$.

In summary, for $0 < \mu < 1$, the drift is not strong enough to change the $\sqrt{n}$ growth of the mean record number. The presence of drift just modifies the prefactor of the $\sqrt{n}$ growth. Similarly, the distribution of the record number in equation (25) in the presence of a drift, when appropriately scaled, remains unchanged from the universal half-Gaussian form in the driftless case.

### 5.2. Regime II: $\mu = 1$ and $c$ arbitrary

As mentioned in the introduction, on the critical line $\mu = 1$, the record statistics was investigated in detail in [26] for the special case of Cauchy jump distribution $f_{\text{Cauchy}}(y) = 1/[\pi (1 + y^2)]$. For a general jump distribution with $\mu = 1$ (not necessarily of the Cauchy form), the record statistics has a very similar mathematical structure that can be derived from the general framework developed in this paper.

Let us first consider the growth of the mean record number $\langle R_n \rangle$ in equation (102). Substituting the small $s$ behavior of $\bar{q}(s)$ from equation (83) and performing the Bromwich integral upon setting $s^* = 0$, we obtain for large $n$

$$\langle R_n \rangle \approx \frac{A_{II}}{\Gamma(1 + \theta(c))} n^{\theta(c)} \text{, where } A_{II} = e^{\theta}.$$
Figure 6. $\ln\langle R_n \rangle / \ln n$ as a function of the drift $c$ for the Cauchy distribution with the Lévy index $\mu = 1$ and for different values of $n = 10^2, 10^3, 10^4$ and $10^5$. For each $n$ and $c$, the average was performed over $10^3$ samples. The results from the numerical simulations collapse and agree with the predicted analytical behavior of $\ln\langle R_n \rangle / \ln n = \theta(c)$ and $\theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c)$ as in equation (85).

Note that $\gamma_0$ is a distribution-dependent constant, while the exponent $\theta(c) = \int_{-\infty}^{c} \mathcal{L}_1(y) dy = 1/2 + \frac{1}{\pi} \arctan(c)$ is universal. In figure 6, this exponent is plotted and compared with numerical simulations of random walks with a Cauchy jump distribution ($\mu = 1$).

Turning now to the distribution $P(R, n)$ in equation (103), as before, we substitute the small $s$ expansion of $\tilde{q}(s)$ from equation (83). It turns out that the appropriate scaling limit for $P(R, n)$ is $n \to \infty, R \to \infty$ but keeping the ratio $R/n^{\theta(c)}$ fixed. To see this, we first set $s^* = 0$, set $R$ large but fixed and keep the leading terms for small $s$ to obtain

$$P(R, n) \approx e^{-\gamma_0} \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} e^{s n} 1 \exp\left[-e^{-\gamma_0} s^{\theta(c)} R\right].$$

(107)

Rescaling $sn \to s$ and keeping the scaled variable $R/n^{\theta(c)}$ fixed results the asymptotic scaling distribution

$$P(R, n) \approx \frac{1}{A_{II} n^{\theta(c)}} g_c \left( \frac{R}{A_{II} n^{\theta(c)}} \right),$$

(108)

where the scaling function $g_c(u)$, which depends continuously on $c$, is given by the formal Bromwich integral:

$$g_c(u) = \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} e^{s u} \cdot e^{s^{\theta(c)}-1} \text{ with } u \geq 0,$$

(109)

where we recall that $0 \leq \theta(c) \leq 1$.

One can easily extract the tail behavior of the scaling function $g_c(u)$ by analyzing the integral in equation (109). For instance, when $u \to 0$, $g_c(u)$ approaches a constant

$$g_c(0) = \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} s^{\theta(c)-1} e^{s} = \frac{1}{\Gamma[\theta(c)]} \sin[\pi \theta(c)] = \frac{1}{\Gamma[1-\theta(c)]}.$$  

(110)

The integral in equation (110) can be performed by wrapping the contour around the branch cut on the negative real $s$-axis.
In the opposite limit, when $u \to \infty$, the integral in equation (109) can be performed using the standard steepest descent method. Skipping details and using the shorthand notation $\theta = \theta(c)$, we obtain
\[
g_c(u \to \infty) \approx [2\pi (1 - \theta) \theta^{1 - 2\theta} (1 - \theta)^{-1/2}] u^{(1 - 2\theta)/2(1 - \theta)} \exp[-(1 - \theta) \theta^{1/2} u^{1/2}].
\]

Thus, the distribution has a non-Gaussian tail. The function $g_c(u)$ can be expressed in terms of the one-sided Lévy distribution, which was discussed, e.g., in [32]. In some particular cases, the Bromwich integral in equation (109) can explicitly be evaluated. For rational values of $\theta(c)$, $g_c(u)$ can be expressed as a finite sum of hypergeometric functions. A very special case corresponds to $c = -1/\sqrt{3}$ where one has $\theta = 1/3$, such that
\[
g_c(u) = -\frac{1}{\sqrt{3}}(u)^{3/4} Ai\left(\frac{u}{3^{1/3}}\right), \quad u \geq 0,
\]

where $Ai(x)$ is the Airy function. Its asymptotic behaviors are then given by
\[
g_{c=-1/\sqrt{3}}(u) \sim 1/\Gamma(2/3), \quad u \to 0,
\]
\[
g_{c=-1/\sqrt{3}}(u) \sim \frac{3}{2\sqrt{\pi}} u^{-1/4} \exp\left(-\frac{2}{3\sqrt{3}} u^{3/2}\right),
\]

which agree with the above-presented general analysis (see equations (110) and (111)). In figure 7, we show a plot of the rescaled probability $A_{II} n^\theta(c) P(R, n)$ as a function of $R/A_{II} n^\theta(c)$ computed numerically for $c = -1/\sqrt{3}$, which agrees reasonably well with our exact analytical result in equation (112).

5.3. Regime III: $1 < \mu < 2$ and $c > 0$

We first compute the asymptotic growth of the mean number of records in this regime. Substituting the leading singular behavior of $\tilde{q}(s)$ from equation (88) on the rhs of equation (102) and performing the Bromwich integral results
\[
\langle R_n \rangle \approx a_{\mu}(c) n, \quad \text{where} \quad a_{\mu}(c) = \frac{1}{\tilde{q}(0)} = \exp\left[-\sum_{n=1}^{\infty} \frac{1}{n} \int_{cn}^{\infty} P_n(x) \, dx\right].
\]
Note that above we used the expression of $\tilde{q}(0)$ in equation (87). We have numerically checked this linear behavior, and in figure 10, the bottom curve shows a plot of $\langle R_n \rangle / n$ as a function of $c$, although we have not tried to numerically evaluate $a_\mu(c)$.

We next consider the distribution $P(R_n)$ in equation (103). We substitute the small $s$ behavior of $\tilde{q}(s)$ from equation (88) into the rhs of equation (103), set $s^* = 0, R$ large and keep only leading small $s$ terms to obtain

$$ P(R_n) \approx \tilde{q}(0) \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} \exp[-s(\tilde{q}(0)R - n) + B_\mu \tilde{q}(0)Rs^*]. \quad (116) $$

Next, we set

$$ R = a_\mu(c)n + a_\mu(c)n^{1/u}, \quad (117) $$

where $a_\mu(c) = 1/\tilde{q}(0)$, and take the limit $R \to \infty$, $n \to \infty$ but keeping the scaled variable $u$ above fixed. We substitute equation (117) into the rhs of equation (116). Keeping only the two leading terms for large $n$ and fixed $u$ results

$$ P(R, n) \approx \tilde{q}(0) \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} \exp[-sn^{1/u} + B_\mu n s^*]. \quad (118) $$

Note that for fixed $u$ both terms in the exponential are of the same order. In fact, the scaling in equation (117) is chosen so as to make the two leading terms precisely of the same order for large $n$. Rescaling $s$ by $n^{1/u}$, i.e. $s n^{1/u} \to s$, and using $a_\mu(c) = 1/\tilde{q}(0)$ reduces equation (118) to a nicer scaling form announced in equation (35):

$$ P(R, n) \approx \frac{1}{a_\mu(c)n^{1/u}} V_\mu(u), \quad \text{where} \quad u = \frac{R - a_\mu(c)n}{a_\mu(c)n^{1/u}}, \quad (119) $$

and the scaling function $V_\mu(u)$ is formally given by the Bromwich integral

$$ V_\mu(u) = \int_{-i\infty}^{+i\infty} \frac{ds}{2\pi i} e^{-us+B_\mu s^*}, \quad (120) $$

where the constant $B_\mu > 0$ is given in equation (40).

Interestingly, the same scaling function $V_\mu(u)$ also appeared in [31] in the context of the partition function of the zero range process on a ring. The asymptotic tails of the function $V_\mu(u)$ were analyzed in great detail in [31] (see equations (78)–(83) and figure 5 in [31], and note that in [31] the index $\mu$ was denoted by $\gamma - 1$). We do not repeat the computations here, but just quote the results. It was found that $V_\mu(u)$ has highly asymmetric tails. For $u \to -\infty$, it decays as a power law: $V_\mu(u) \to K_\mu|u|^{-\mu-1}$, where the prefactor $K_\mu = B_\mu \Gamma(1+\mu) \sin[\pi(\mu+1)/\pi]$. Using our expression $B_\mu = -1/2\cos(\mu\pi/2)$ from equation (40), it is easy to show that $K_\mu = A_\mu$, where the constant $A_\mu$ is defined in equation (8). This leads to equation (36). In contrast, when $u \to \infty$, $V_\mu(u)$ has a faster than Gaussian tail as precisely described in equation (37). To plot this scaling function, a convenient real space representation can be used from [31]. Replacing $\gamma - 1$ by $\mu$ in equation (84) of [31] and using $B_\mu = -1/2\cos(\mu\pi/2)$, we obtain

$$ V_\mu(u) = \frac{1}{\pi} \int_0^\infty dy e^{-y^2/2} \cos\left[\frac{1}{2} \tan(\mu\pi/2) y^\mu + y u\right]. \quad (121) $$

We compared this result for a Lévy index of $\mu = 1.5$ with numerical simulations in figure 8. Even though the convergence of the numerically obtained distributions is slow, it is clear that the asymptotic distribution $V_\mu(u)$ is approached for $n \to \infty$. In figure 9, we plotted numerical simulations of the rescaled record number distribution for different values of $\mu$. One finds both numerically and by taking the limit in equation (134) that, for $\mu \to 2$, this rescaled distribution approaches a Gaussian form (see regime IV).
Figure 8. Rescaled distribution \( a_{\mu}(c)n^{1/\mu}P(R, n) \) of the record number \( R_n \) after \( n \) steps for a random walk with a Lévy-stable jump distribution of Lévy index \( \mu = 1.5 \). The data are plotted as a function of the shifted and scaled variable \( u = (R - a_{\mu}(c)n^{1/\mu})/(a_{\mu}(c)n^{1/\mu}) \). The different curves correspond to different values of \( n = 10^3, 10^4, 10^5 \) and \( 10^6 \) and for a drift \( c = 1 \). They were obtained by averaging over \( 10^6 \) samples. For \( n = 10^5 \) and \( n = 10^6 \) the numerical results were binned for technical reasons. We also plotted our analytical results for the scaling function \( V_{\mu}(u) \) given by equation (121). While for smaller values of \( n \), there is still a significant difference between the simulations and our analytical result, it converges to the behavior in equation (121) when \( n \) increases.

Figure 9. Rescaled distribution \( a_{\mu}(c)n^{1/\mu}P(R, n) \) of the record number \( R_n \) after \( n = 10^4 \) steps for a random walk with a Lévy-stable jump distribution with different Lévy indices \( \mu = 1.25, 1.5, 1.75 \) and \( 2 \). The data are plotted as a function of the shifted and scaled variable \( u = (R - a_{\mu}(c)n^{1/\mu})/(a_{\mu}(c)n^{1/\mu}) \). For all these data, the value of the drift is \( c = 1 \) and they have been obtained by averaging over \( 10^6 \) samples. For \( \mu \to 2 \), this rescaled distribution approaches the Gaussian form given in equation (134).

To summarize, in this regime, the mean record number increases linearly with increasing \( n \), but the typical fluctuations around the mean are anomalously large of \( O(n^{1/\mu}) \) (superdiffusive) as described in equation (117). In addition, the probability distribution of these typical fluctuations around the mean are described by a highly non-Gaussian form precisely described in equation (119).
5.4. Regime IV: $\mu = 2$ and $c > 0$

In this regime, as explained in section 4.3, $\tilde{q}(s) = \exp[W_{\epsilon,\mu}(s)]$ in equations (69) and (70) is analytic at $s = 0$. This can be seen by expanding the sum $W_{\epsilon,\mu}(s)$ in equation (70) in a Taylor series in $s$:

$$W_{\epsilon,\mu}(s) = \sum_{m=0}^{\infty} d_m s^m, \quad \text{where} \quad d_m = \frac{(-1)^m}{m!} \sum_{n=1}^{\infty} n^{m-1} \int_{c_n}^{\infty} P_n(x) \, dx. \quad (122)$$

The coefficient $d_m$, for each $m$, is finite as the sum over $n$ is convergent since the integral $\int_{c_n}^{\infty} P_n(x) \, dx$ decreases with $n$ faster than exponentially for large $n$ (see section 4.3), as long as $\mu = 2$ and $c > 0$. Consequently, for small $s$, $\tilde{q}(s)$ also has a Taylor series expansion:

$$\tilde{q}(s) = \tilde{q}(0) + \tilde{q}'(0) \, s + \frac{1}{2} \tilde{q}''(0) \, s^2 + \cdots. \quad (123)$$

Let us start with the asymptotic behavior of the mean record number $\langle R_n \rangle$ in equation (102).

Once again, the dominant contribution to the integral in equation (102) for large $n$ comes from the small $s$ region. Taking the $s \to 0$ limit in the integrand and using the small $s$ expansion in equation (123), keeping only the leading terms and performing the Bromwich integral term by term, one obtains for large $n$

$$\langle R_n \rangle \approx \int_{s^{-1} \to 0} \frac{dx}{2\pi i} \frac{e^{i\pi n}}{\tilde{q}(0)s^2} \left[ 1 + \left( 1 - \frac{\tilde{q}'(0)}{\tilde{q}(0)} \right) s + O(s^2) \right] \approx a_2(c)n + \kappa_2(c) + O(1/n), \quad (124)$$

where

$$a_2(c) = \frac{1}{\tilde{q}(0)} = \exp \left[ -\sum_{n=1}^{\infty} \frac{1}{n} \int_{c_n}^{\infty} P_n(x) \, dx \right] \quad (125)$$

and $\kappa_2(c) = [1 - \tilde{q}'(0)/\tilde{q}(0)]/\tilde{q}(0)$. 

---

**Figure 10.** Numerical simulations of $\langle R_n \rangle/n$ for random walks with a Gaussian (with variance $\sigma = 1$), an exponential (with parameter $b = 1$, see its definition below equation (126)), both in regime IV, and a Lévy-stable jump distribution with $\mu = 1.5$, in regime III, with the positive drift $c > 0$. For each distribution we show data for $n = 10^4$ which were obtained by averaging over $10^4$ samples. For the Gaussian and the exponential distribution, we also plotted a numerical evaluation of our exact formula for $a_2(c)$ using equation (126) for the Gaussian case and equation (B.12) for the exponential case. These curves perfectly agree with the numerical simulations.
For example, for a Gaussian jump distribution \( f(\eta) = (2\pi \sigma^2)^{-1/2} e^{-\eta^2/2\sigma^2} \), we have \( P_n(x) = (2\pi \sigma^2)^{-1/2} e^{-x^2/2\sigma^2 n} \), and hence, \( a_2(c) \) in equation (125) is given by the explicit formula

\[
a_2(c) = \exp \left[ -\sum_{n=1}^{\infty} \frac{1}{2n} \text{erfc} \left( \frac{c \sqrt{n}}{\sigma \sqrt{2}} \right) \right]. \tag{126}
\]

For instance, for \( c = 1 \) and \( \sigma = 1 \), one obtains \( a_2(c = 1) = 0.800543 \ldots \). Another example is the exponential jump distribution \( f(\eta) = (2b)^{-1} \exp(-|\eta|/b) \). In this case, one can also compute (see appendix B) the constant \( a_2(c) = \lambda \), where \( \lambda \) is given by the solution of the transcendental equation \( \exp(-\lambda c/b) = 1 - \lambda^2 \). For example, for \( c = 1, b = 1 \), one obtains \( \lambda = 0.714556 \ldots \). For these two examples, we have confirmed the leading asymptotic result for the mean record number in equation (124) with the exactly computed prefactors \( a_2(c) \) (as discussed above) in our numerical simulations (see figure 10).

In a similar way, one can also analyze equation (66) for the large \( n \) behavior of the second moment \( \langle R^2 \rangle \). Skipping details, we obtain the following leading large \( n \) behavior:

\[
\langle R^2 \rangle \approx a_2^+(c) n^2 + a_2^-(c) n + O(1), \quad \text{where} \quad a_2^+(c) = \frac{1}{\tilde{q}^2(0)} \left[ 3 - \tilde{q}(0) - 4 \tilde{q}'(0) \tilde{q}(0) \right]. \tag{127}
\]

Consequently, the variance of the record number grows for large \( n \) as

\[
\langle R^2 \rangle - \langle R \rangle^2 \approx b_2(c) n, \quad \text{where} \quad b_2(c) = \frac{1}{\tilde{q}^2(0)} \left[ 1 - \tilde{q}(0) - 2 \tilde{q}'(0) \tilde{q}(0) \right]. \tag{128}
\]

Thus, in this regime, the mean record number grows linearly with \( n \) for large \( n \), while the size of typical fluctuations around this mean grows as \( \sim \sqrt{n} \).

How are these typical fluctuations around the mean distributed? To answer this, we need to analyze \( P(R, n) \) in equation (103) in the scaling limit where both \( n \) and \( R \) are large, but the ratio \( (R - a_2(c) n)/\sqrt{n} \) is fixed. To proceed, we set \( s^* = 0 \) and substitute the small \( s \) expansion of \( \tilde{q}(s) \) in equation (123) into the rhs of equation (103), take \( R \) large but fixed, to obtain

\[
P(R, n) \approx \tilde{q}(0) \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} \exp \left[ -s (\tilde{q}(0) R - n) + (1/2) b_2(c) \tilde{q}^2(0) R^2 s^2 \right], \tag{129}
\]

where \( b_2(c) \) is given in equation (128). Next, we set

\[
R = a_2(c) n + \sqrt{b_2(c)} \sqrt{n} u, \tag{130}
\]

where \( a_2(c) = 1/\tilde{q}(0) \) is given in equation (125), and take the scaling limit where \( R \to \infty \), \( n \to \infty \) but keeping the scaled variable \( u \) above fixed. Substituting \( R \) from equation (130) into equation (129) and keeping only the two leading terms for large \( n \) results

\[
P(R, n) \approx \tilde{q}(0) \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} \exp \left[ -\sqrt{b_2(c)} \tilde{q}(0) \sqrt{n} s u + (1/2) b_2(c) \tilde{q}^2(0) n s^2 \right]. \tag{131}
\]

Note that for fixed \( u \), both terms inside the exponential are of the same order. Indeed, as in the section 5.2, the scaling in equation (130) is chosen so as to make the two leading terms precisely of the same order for large \( n \). Rescaling \( \sqrt{b_2(c)} \tilde{q}(0) \sqrt{n} s \to s \) simplifies to

\[
P(R, n) \approx \frac{1}{\sqrt{b_2(c) n}} V_2(u), \quad \text{where} \quad u = \frac{R - a_2(c) n}{\sqrt{b_2(c) n}}, \tag{132}
\]

and the scaling function \( V_2(u) \) is given by the Bromwich integral

\[
V_2(u) = \int_{-\infty}^{+\infty} \frac{ds}{2\pi i} e^{-us + s^2/2}. \tag{133}
\]
Figure 11. Plot of the cumulative distribution of record numbers $P_{\leq}(R, n) = \Pr[R_n \leq R]$ as a function of the shifted and scaled variable $u = (R - a_2(c)n)/(\sqrt{b_2(c)n})$ for a random walk with Gaussian jump distribution (with $\sigma = 1$) of $n = 10^4$ steps. The different curves correspond to different values of positive drift $c = 1/16, 1/4, 1$ and $2$. For each $c$ the data were obtained by averaging over $10^6$ samples. We compared the numerical results to the cumulative distribution of $V_2(\mu)$, which we obtained analytically (equation (134)). All curves collapse nicely, confirming that the asymptotic record number of a biased Gaussian random walk with a positive drift has the Gaussian distribution given by equation (132).

which can exactly be computed (since it is a Gaussian integral) to result

$$V_2(u) = \frac{1}{\sqrt{2\pi}} \exp\left[-u^2/2\right]. \tag{134}$$

This then proves that $P(R, n)$ is asymptotically Gaussian as announced in equation (45). Figure 11 numerically confirms this result. We plotted the cumulative distribution of record numbers $P_{\leq}(R, n) = \Pr[R_n \leq R]$ as a function of the shifted and scaled variable $u = (R - a_2(c)n)/(\sqrt{b_2(c)n})$ after $n = 10^4$ steps for different values of positive drift $c$ and compared them to a Gaussian cdf (cumulative distribution function). All numerical results collapsed perfectly on the analytical curve.

5.5. Regime V: $1 < \mu \leq 2$ and $c < 0$

In this regime, we set $s^* = 0$ in equation (103) and substitute into its rhs the small $s$ expansion of $\tilde{q}(s)$ from equation (100). Keeping only leading order behavior for small $s$ results, for large $n$,

$$P(R, n) \approx \alpha_{\mu}(c)[1 - \alpha_{\mu}(c)]^{R-1} \int_{s^*}^{s^* + i\infty} \frac{dx}{2\pi i} e^{x n} \frac{1}{x}, \tag{135}$$

where the constant $\alpha_{\mu}(c) = \exp[-W_{\mu,\mu}(0)] = \exp\left[-\sum_{n=1}^{\infty} \frac{1}{n} \int_{\mu_{\mu}}^{\infty} P_n(x) \, dx\right]$ as given in equation (101).

Using the fact that $LT^{-1}[1/s] = 1$ results the large $n$ (but $R$ fixed) behavior of $P(R, n):$

$$P(R, n) \xrightarrow{n \to \infty} \alpha_{\mu}(c)[1 - \alpha_{\mu}(c)]^{R-1}. \tag{136}$$

Thus, the distribution becomes independent of $n$ for large $n$ and has a simple geometric form with mean $\langle R_n \rangle \to 1/\alpha_{\mu}(c)$. Comparing the expression of $\alpha_{\mu}(c)$ as given in equation (101) and those of $\alpha_{\mu}(c)$ in equation (115) and $a_2(c)$ in equation (125) for $c > 0$, one immediately
Figure 12. Rescaled distribution $a_2(\mid c\mid) P(R, n)$ of the record number $R_n$ after $n = 10^4$ steps for a random walk with a Gaussian jump distribution, of variance $\sigma = 1$, with different negative values of the drift $c = -0.01$, $c = -0.05$, $-0.1$ and $-0.25$. The data are plotted as a function of the rescaled variable $u = R a_2(\mid c\mid)$. For each value of $c$, the data were obtained by averaging over $10^4$ samples. We compared the numerical results with a simple geometric distribution. The good agreement confirms our analytical findings given by equation (136).

finds that $a_\mu(c) = a_\mu(\mid c\mid)$ for $1 < \mu < 2$, while $a_2(c) = a_2(\mid c\mid)$, the results mentioned, respectively, in equations (50) and (51).

In figure 12, we compared equation (136) to numerical simulations of negatively biased Gaussian random walks with different values of $c$. For large $n$, the rescaled distribution of $u = R a_2(\mid c\mid)$ approaches the geometric (exponential) distribution $e^{-u}$.

6. Extreme statistics of the age of a record

From the previous study of the mean number of records $\langle R_n \rangle$, one deduces that the typical age (see figure 2) of a record is given by $l_{typ} \sim n/\langle R_n \rangle$. However, following [22] for the unbiased case, it turns out that the extreme ages of records do not share the typical behavior. In this section, we probe such atypical extremal statistics by considering the longest and shortest lasting records characterized by their respective ages (durations) $l_{max,n}$ and $l_{min,n}$. We focus on their mean values $\langle l_{max,n} \rangle, \langle l_{min,n} \rangle$ and find rather different asymptotic behaviors in the five regimes in the $(c, 0 < \mu \leq 2)$ strip mentioned before (figure 1).

6.1. Age of the longest lasting record $l_{max,n}$

We first consider the longest lasting record whose age $l_{max,n}$ is given by (see figure 2)

$$l_{max} = \max(l_1, l_2, \ldots, l_R).$$

(137)

The cumulative distribution $F_n(m) = \text{Proba.}(l_{max,n} \leq m)$ was studied in [22], where an explicit formula for its GF was obtained:

$$\sum_{m=0}^{\infty} F_n(m) z^n = \frac{\sum_{j=1}^{m} Q(j) z^j}{1 - \sum_{j=1}^{m} F(j) z^j},$$

(138)
where $F(l) = Q(l - 1) - Q(l)$, from which one deduces the GF of the mean $\langle l_{\text{max},n} \rangle = \sum_{m=1}^{\infty} [1 - F(m)]$

$$\sum_{n=0}^{\infty} z^n \langle l_{\text{max},n} \rangle = \sum_{m=1}^{\infty} \left[ \frac{1}{1-z} - \frac{\sum_{l=1}^{m} Q(l)z^l}{1 - \sum_{l=1}^{m} F(l)z^l} \right]$$

$$= \frac{1}{1-z} \left[ \sum_{m=1}^{\infty} \sum_{l=0}^{\infty} F(l)z^l + \sum_{l=0}^{\infty} (1-z)^l \sum_{m=1}^{\infty} Q(l)z^l \right]$$

$$(1 - z)Q(z) + \sum_{l=0}^{\infty} F(l)z^l,$$  \hspace{1cm} (139)

$$(140)$$

where we have used that $\tilde{F}(z) = 1 - (1 - z)\tilde{Q}(z)$; see equation (58).

In the absence of drift, $c = 0$, it was shown in [22] that $\langle l_{\text{max},n} \rangle$ behaves, for large $n$, linearly with $n$ with a non-trivial coefficient, independently of the jump distribution $f(\eta)$:

$$\langle l_{\text{max},n} \rangle \sim C_0 n, \quad C_0 = \int_0^{\infty} \frac{1}{1 + y^{3/2}e^y} \frac{dy}{\sqrt{\pi}} \frac{dx}{x^{-1/2}e^{-x}} = 0.626508 \ldots.$$  \hspace{1cm} (141)

Interestingly, this constant $C_0$ appears also in the study of the longest excursion of Brownian motion [33, 34]. Note that to obtain the large $n$ behavior of $\langle l_{\text{max},n} \rangle$ from equation (139), one has to analyze formula (139) in the limit $z \to 1$. We will see that in this limit the above sum over $m$ is dominated by the large values of $m$, which thus crucially depends on the large $m$ behavior of the persistence probability $Q(m)$. Consequently, $\langle l_{\text{max},n} \rangle$ behaves quite differently in the five regimes in the $(c, 0 < \mu \leq 2)$ strip in figure 1 and are summarized as follows:

$$\langle l_{\text{max},n} \rangle \sim n$$ for $0 < \mu < 1$ and $c$ arbitrary (regime I),

$$\sim n \mu$$ for $\mu = 1$ and $c$ arbitrary (regime II),

$$\sim n^{3/2}$$ for $1 < \mu < 2$ and $c > 0$ (regime III),

$$\sim \ln n$$ for $\mu = 2$ and $c > 0$ (regime IV),

$$\sim n$$ for $1 < \mu < 2$ and $c < 0$ (regime V).  \hspace{1cm} (142)

In the following, we will discuss the behavior of $\langle l_{\text{max},n} \rangle$ separately for the five regimes.

6.1.1. Regime I: $0 < \mu < 1$, $c$ arbitrary. In this regime, we recall that $Q(m)$ behaves, for large $m$, as

$$Q(m) \sim \frac{B_1}{\sqrt{m}}, \quad F(m) \sim \frac{B_1}{2m^{3/2}}.$$  \hspace{1cm} (143)

where $B_1$ is given in equation (78).

Setting $z = e^{-s}$ we are interested in the limit $s \to 0$ in the formula given in equation (139), where one can replace $F(m)$ and $Q(m)$ by their asymptotic behaviors

$$\sum_{n=0}^{\infty} \langle l_{\text{max},n} \rangle e^{-sn} \sim \frac{1}{s} \sum_{m=1}^{\infty} \frac{1}{2} \sum_{l=0}^{\infty} l^{-3/2}e^{- sl} + s \sum_{m=1}^{\infty} l^{-1/2}e^{- sl}.$$  \hspace{1cm} (144)

Here, we have used $\tilde{q}(s) \sim \sqrt{\pi}B_1/\sqrt{s}$ when $s \to 0$; see equations (76) and (78). In the limit $s \to 0$, the discrete sums over $l$ and $m$ can be replaced by integrals and one finds that the right-hand side in equation (144) behaves like $1/s^2$ when $s \to 0$ with a prefactor that we can compute to obtain the large $n$ behavior of $\langle l_{\text{max},n} \rangle$ as

$$\langle l_{\text{max},n} \rangle \sim C_1 n, \quad C_1 = \int_0^{\infty} \frac{dy}{\sqrt{\pi} + \frac{1}{2} \int_0^{\infty} dx x^{-3/2}e^{-x}} = C_0,$$  \hspace{1cm} (145)

where $C_0$ is the same as given in (141) and where the last equality is simply obtained by performing integration by part in the integral over $x$ in the denominator. In figure 13, we have plotted the results of our numerical estimate of $\langle l_{\text{max},n} \rangle$ (obtained by averaging over $10^4$
Figure 13. Plot of $\langle l_{\text{max},n} \rangle / n$ in the different regimes I, II and V: the points are the results of our numerical simulations. For regime II ($\mu = 1$), we present two curves, one with a positive drift ($c = -1$) (the second curve from top) and one with a negative drift ($c = 1$) (the bottom curve). These data indicate that in all these cases $\langle l_{\text{max},n} \rangle \propto n$, for large $n$, with an amplitude that agrees quite well with our analytical results, which are represented by solid lines for each of these cases and correspond to the formula given in equations (145), (147) and (158).

different realizations of random walks) for $\mu = 0.5$ and two different values of $c = \pm 1.0$. This plot shows that $\langle l_{\text{max},n} \rangle / n$ saturates rather quickly to the constant $C_0$, independently of $c$, in agreement with equation (145).

Thus, in this regime, the large $n$ behavior of $\langle l_{\text{max},n} \rangle$ is unaffected by the presence of the drift $c$. This result could have been anticipated as $l_{\text{max},n}$ can be considered as the longest excursion between two consecutive zeros of a renewal process with a persistence exponent $1/2$. This quantity was studied in [34] and its average was computed, yielding the large $n$ behavior obtained in equation (145).

6.1.2. Regime II: $\mu = 1$ and $c$ arbitrary. In this regime, we recall that the persistence probability $Q(m)$ behaves algebraically for large $m$ with an exponent $\theta(c)$, which depends continuously on $c$:

$$Q(m) \sim B_{\text{II}} m^\theta(c), \quad \theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c),$$

where the amplitude $B_{\text{II}}$ is the same as given in equation (85). Here, again we can use the result obtained in [34] for the longest excursion between consecutive zeros of a renewal process with a persistence exponent $\theta(c)$ to obtain

$$\langle l_{\text{max},n} \rangle \sim C_{\text{II}} n, \quad C_{\text{II}} = \int_0^\infty dy \frac{1}{1 + y^\theta(c)} \int_0^y dx x^{-\theta(c)} e^{-x},$$

which continuously depends on $c$ and is independent of the non-universal amplitude $B_{\text{II}}$ (see equation (146)). In figure 14, we show a comparison of $C_{\text{II}}$ numerically obtained (the squares symbols) and from our exact formula (solid line), which shows a very good agreement between both.

6.1.3. Regime III: $1 < \mu < 2$ and $c > 0$. In this regime, the persistence probability $Q(m)$ behaves for large $m$ as

$$Q(m) \sim B_{\text{III}} m^\mu,$$
where the amplitude $B_{III}$ is given in equation (90). Using again the results obtained in [34], one obtains that

$$\langle l_{\text{max}}, n \rangle \sim C_{III} n^{1/\mu},$$

where, however, the amplitude $C_{III}$ was not given in [34]. A careful analysis of formula (139) allows us to obtain the amplitude $C_{III}$ as

$$C_{III} = \frac{1}{c} \left[ \Gamma(1 - 1/\mu) \left[ \frac{1}{\pi} \sin \left( \frac{\mu \pi}{2} \right) \Gamma(\mu) \right]^{1/\mu} \right],$$

which diverges as $C_{III} \sim (\pi(\mu - 1))^{-1}$ when $\mu \to 1$ and vanishes as $C_{III} \sim \sqrt{\pi(2 - \mu)/2}$ when $\mu \to 2$. In figure 15, we show a plot of our numerical data for $\langle l_{\text{max}}, n \rangle$ (averaged again over $10^4$ different realizations) for different values of $\mu = 1.4, 1.5, 1.7, 1.9$ and for a fixed value of the drift $c = 5.0$. The solid lines indicate the corresponding exact asymptotic behaviors in equations (149) and (150): the agreement between the two is quite good although the convergence to the asymptotic behavior gets slower as $\mu$ decreases to 1.

6.1.4. Regime IV: $\mu = 2$ and $c > 0$. In this case, the persistence $Q(m)$ behaves quite differently as it vanishes exponentially for large $m$ as

$$Q(m) \sim \frac{B_{IV}}{n^{3/2}} e^{-s_1 n}, \quad \text{where} \quad s_1 = \frac{c^2}{2\sigma^2},$$

where the amplitude $B_{IV}$ is given in equation (96). This case was not analyzed in [34]. From equation (139), one has in this case

$$\sum_{n=0}^{\infty} \langle l_{\text{max}}, n \rangle e^{-sn} \sim \frac{1}{s} \sum_{m=1}^{\infty} \left[ s\hat{q}(0) + \sum_{l=m}^{\infty} F(l) \right] = \frac{1}{s} \sum_{m=1}^{\infty} Q(m) s\hat{q}(0) + Q(m).$$

Therefore, in the limit when $s \to 0$, one can estimate the leading behavior of the sum over $m$ as

$$\sum_{n=0}^{\infty} \langle l_{\text{max}}, n \rangle e^{-sn} \sim \frac{m^*}{s},$$

![Figure 14. Plot of $C_m$ as a function of $c$. The red squares correspond to numerical data, while the solid line corresponds to our analytical result in equation (147) together with equation (146).](image-url)
Figure 15. Plot, in a ln–ln scale, of $\langle l_{\text{max}, n} \rangle$ as a function of $n$ in regime III: the different curves correspond to different values of $\mu = 1.4, 1.5, 1.7, 1.9$ with a fixed value of $c = 5.0$. The solid line are the exact results given in equations (149) and (150), without any fitting parameter.

where $m^*$ is such that

$$Q(m^*) \sim s\tilde{q}(0).$$  \hfill (154)

From the asymptotic behavior given in (151), one finds that $m^* \sim -\frac{c^2}{2\sigma^2} \ln s$ so that finally

$$\langle l_{\text{max}, n} \rangle \sim C_{IV} n, \quad C_{IV} = \frac{2\sigma^2}{c^2},$$  \hfill (155)

which is in sharp contrast with the algebraic growth obtained in equation (149) for $1 < \mu < 2$ and $c > 0$. In figure 16, we show a plot of $\langle l_{\text{max}, n} \rangle$ as a function of $\ln n$: the straight line suggests indeed a logarithmic growth, in agreement with our analytic result (155). However, a more precise comparison with this exact asymptotic result, as shown in the inset of figure 16, suggests that the leading corrections are proportional to $\ln \ln n$, and hence quite strong.

6.1.5. Regime V: $1 < \mu \leq 2$ and $c < 0$. In this case, the persistence probability $Q(m)$ tends asymptotically to a constant (101):

$$Q(m) \xrightarrow{m \to \infty} \alpha_\mu(c) = \exp[-W_{|c|, \mu}(0)] = \exp \left[ - \sum_{n=1}^{\infty} \delta \int_{|c|}^{\infty} P_n(x) \, dx \right].$$  \hfill (156)

In addition from (86), one has that $Q(m) - \alpha_\mu(c) \propto n^{1-\mu}$ so that $F(m) \propto m^{-\mu}$ for large $m$. Therefore, the terms entering into the sum in equation (139) are given to leading order when $1 - z = e^{-s} \to 0$ and large $m$ (which are terms that give the leading contribution to this sum over $m$):

$$\sum_{l=0}^{\infty} F(l) z^l + (1 - z) \sum_{l=0}^{\infty} Q(l) z^l \sim \frac{\alpha_\mu(c)}{\tilde{q}(0)} e^{-s m} = e^{-s m}.$$  \hfill (157)

Therefore this yields

$$\langle l_{\text{max}, n} \rangle \sim C_{V} n, \quad C_{V} = 1.$$  \hfill (158)

This result, which is corroborated by our numerical simulations (see figure 13), can be physically understood as in this regime, where $c < 0$ and $\mu > 1$, the number of records
is finite and these records typically occur during the first steps of the random walks, where the walker might stay positive for a short while before it escapes to negative values when $n \to \infty$, and record no longer happens.

6.2. Age of shortest lasting record $l_{\text{min}, n}$

We now consider the shortest lasting record whose age $l_{\text{min}, n}$ is given by (see figure 2)

$$l_{\text{min}, n} = \min(l_1, l_2, \ldots, l_R).$$

Note that, given that the final incomplete interval $l_R$ is taken into consideration above, $l_{\text{min}, n}$ can be zero: this happens when a record has been broken at the last step, such that $l_R = 0$.

The cumulative distribution $G_n(m) = \text{Proba.}(l_{\text{min}, n} \geq m)$ was studied in [22] and an explicit formula was obtained for its GF:

$$\sum_{n=0}^{\infty} G_n(m) z^n = \frac{\sum_{l=0}^{\infty} Q(l) z^l}{1 - \sum_{l=m}^{\infty} F(l) z^l},$$

from which one obtains the GF of the average value $\langle l_{\text{min}, n} \rangle$ as

$$\sum_{n=0}^{\infty} z^n \langle l_{\text{min}, n} \rangle = \sum_{m=1}^{\infty} \frac{\sum_{l=m}^{\infty} Q(l) z^l}{1 - \sum_{l=m}^{\infty} F(l) z^l}.$$ 

In the absence of drift, $c = 0$, it was shown in [22] that

$$\langle l_{\text{min}, n} \rangle \sim D \sqrt{n}, \quad D = \frac{1}{\sqrt{\pi}}.$$ 

As for $\langle l_{\text{max}, n} \rangle$ we will see that the behavior of $\langle l_{\text{min}, n} \rangle$, in the presence of non-zero drift $c \neq 0$, is quite different in the five different regimes discussed above. Again we start by giving a brief
summary of our results for \( \langle l_{\text{min}, n} \rangle \):

\[
\langle l_{\text{min}, n} \rangle \sim \sqrt{n} \quad \text{for} \quad 0 < \mu < 1 \text{ and } c \text{ arbitrary (regime I)},
\]

\[
\sim n^{1-\theta(c)} \quad \text{for} \quad \mu = 1 \text{ and } c \text{ arbitrary (regime II)},
\]

\[
\sim \text{const.} \quad \text{for} \quad 1 < \mu < 2 \text{ and } c > 0 \quad (\text{regime III}),
\]

\[
\sim \text{const.} \quad \text{for} \quad \mu = 2 \text{ and } c > 0 \quad (\text{regime IV}),
\]

\[
\sim n \quad \text{for} \quad 1 < \mu \leq 2 \text{ and } c < 0 \quad (\text{regime V}),
\]

(163)

again with \( \theta(c) \) being the same as defined in equation (16). In the following, we discuss the behavior of \( \langle l_{\text{min}, n} \rangle \) in more detail for each of the five regimes.

6.2.1. Regime I: \( 0 < \mu < 1 \) and \( c \text{ arbitrary} \). In this case, the persistence probability decays algebraically as given in equation (143) and the analysis of \( \langle l_{\text{min}, n} \rangle \) can be obtained by noting that, in the limit \( z \to 1 \), the denominator in equation (161) can be simply replaced by 1 while the remaining sums over \( l \) (in the numerator) and over \( m \) can be replaced by integrals. This straightforwardly yields

\[
\langle l_{\text{min}, n} \rangle \sim D_1 \sqrt{n},
\]

(164)

\[
D_1 \equiv B_1 \equiv \frac{1}{\sqrt{\pi}} \exp \left[ -\frac{1}{\pi} \int_0^\infty \frac{dk}{k} \arctan \left( \frac{\hat{f}(k) \sin (kc)}{1 - \hat{f}(k) \cos (kc)} \right) \right],
\]

(165)

where the expression of \( B_1 \) is the same as given in equation (78). In figure 17, we show the results of our numerical simulations which are in a rather good agreement with equation (164), although the corrections to this exact asymptotic behavior are clearly visible, in particular for \( \mu = 0.8, c = 1.0 \). In figure 18, we show a plot of the numerical computation of \( \langle l_{\text{min}, n} \rangle \) for \( \mu = 1 \) and different values of \( c = -1, 0.5 \) and \( c = 1 \): these data are in good agreement with the power-law growth in equation (164), although we have not attempted to numerically estimate the prefactor \( D_1 \).
6.2.2. Regime II: \( \mu = 1 \) and \( c \) arbitrary. In this regime where the persistence probability \( Q(m) \) decays algebraically as in equation (84), \( \langle l_{\text{min}, n} \rangle \) can be analyzed as in the regime I where in the limit \( z \to 1 \), the denominator in equation (161) can be simply replaced by 1 while the remaining sums over \( l \) (in the numerator) and over \( m \) can be replaced by integrals. This straightforwardly yields

\[
\sum_{n=1}^{\infty} e^{-\mu} \langle l_{\text{min}, n} \rangle \sim \frac{B_{\text{II}}}{2^{-\theta(c)}},
\]

which yields

\[
\langle l_{\text{min}, n} \rangle \sim D_{\text{II}} n^{1-\theta(c)}, \quad D_{\text{II}} = B_{\text{II}},
\]

where \( B_{\text{II}} \) is given in equation (85) and \( \theta(c) = 1/2 + 1/\pi \arctan(c) \).

6.2.3. Regime III: \( 1 \leq \mu < 2 \) and \( c > 0 \). In this case, we write formula (161) as

\[
\sum_{m=1}^{\infty} e^{-\mu} \langle l_{\text{min}, n} \rangle = \frac{1}{1-z} \left( 1 - \frac{1}{\tilde{Q}(0)} \right) + \sum_{m=2}^{\infty} \frac{\sum_{l=m}^{\infty} Q(l) z^l}{1 - \sum_{m=2}^{\infty} F(l) z^l},
\]

where we have simply isolated the term \( m = 1 \) and used \( 1 - \tilde{F}(0) = (1-z)\tilde{Q}(0) \) (see equation (58)). Now the sum (168), which starts with \( m = 2 \), is dominated by the large values of \( m \). Because of the algebraic decay of \( Q(m) \sim m^{-\mu} \) in this case (148) and \( \mu > 1 \) in this regime, one obtains that this second term behaves like \( (1-z)^{\mu/2} \), which is then subleading, compared to the first term, which behaves like \( (1-z)^{-1} \). Therefore, one obtains in this case

\[
\langle l_{\text{min}, n} \rangle \sim D_{\text{III}}, \quad D_{\text{III}} = 1 - \frac{1}{\tilde{q}(0)} = 1 - \exp \left[ -\sum_{n=1}^{\infty} \int_{cn}^{\infty} P_n(x) \, dx \right],
\]

where we have used the expression for \( 1/\tilde{q}(0) \) given in equation (115). In figure 19, we show a plot of the numerical computation \( \langle l_{\text{min}, n} \rangle \) for \( \mu = 1.5 \) and different values of \( c = 0.5 \) and \( c = 1 \), which is in very good agreement with equation (169). Note that we have extracted the value of \( 1/\tilde{q}(0) \), which enters into the expression of \( D_{\text{III}} \) from the linear growth of the mean record number \( \langle R_n \rangle \), according to (115).
6.2.4. Regime IV: $\mu = 2$ and $c > 0$. A similar analysis can be carried out in this case, starting from the same formula (168). In this case, in the sum (168), which starts with $m = 2$, one can safely put $z = 1$, because of the behavior of the exponential decay of $Q(m)$ in this case (see equation (151)). Therefore, one immediately obtains

$$\langle l_{\text{min}}, n \rangle \sim D_{IV}, \quad D_{IV} = 1 - \frac{1}{q(0)} = 1 - \exp \left[ - \sum_{n=1}^{\infty} \frac{1}{n} \int_{c_n}^{\infty} P_n(x) \, dx \right], \quad (170)$$

where we have used the expression for $1/q(0)$ given in equation (125). In figure 19, we show a plot of the numerical computation $\langle l_{\text{min}}, n \rangle$ for $\mu = 2$ and $c = 1$, which is in good agreement with equation (170). Note that we have extracted the value of $1/q(0)$ which enters into the expression of $D_{IV}$ from the linear growth of the mean record number $\langle R_n \rangle$, according to equation (125).

6.2.5. Regime V: $1 < \mu \leq 2$ and $c < 0$. In this regime where the persistence goes to a constant $Q(m) \to \alpha_{\mu}(c)$, for $m \gg 1$, one can simply replace $Q(l)$ by this constant value in the sum of the numerator in equation (161) while the denominator can be simply approximated by 1 in the limit $1 - z = e^{-s} \to 0$. This yields straightforwardly

$$\langle l_{\text{min}}, n \rangle \sim \alpha_{\mu}(c) \, n. \quad (171)$$

In figure 20, we show a plot of $\langle l_{\text{min}, n} \rangle / n$ that we have numerically computed for different values of $\mu = 1.7, 1.5$ and $\mu = 2$ and also for different values of the drift. These results are in very good agreement with our exact asymptotic result in equation (171), where the value of $\alpha_{\mu}(c)$ have been extracted from the mean record number $\langle R_n \rangle \sim 1/\alpha_{\mu}(c)$ (49). This result (171) can easily be understood by realizing that $l_{\text{min}, n} = n$ if the whole trajectory is on the negative side, which happens with probability $\alpha_{\mu}(c)$, while $l_{\text{min}, n}$ is of order $O(1)$ if the walker makes an excursion on the positive side. One also notes that, in this case, $l_{\text{typ}} \sim \langle l_{\text{min}, n} \rangle$. 

---

**Figure 19.** Plot of $\langle l_{\text{min}, n} \rangle$ as a function of $n$ for $\mu = 1.5$ and $\mu = 2$ and different values of $c > 0$, therefore corresponding to regimes III and IV. The solid line corresponds to the exact result, from equations (169) and (170).
7. Conclusion

In this paper, we considered a very simple model of a one-dimensional discrete-time random walk in the presence of a constant drift $c$. At each time step, the particle jumps by a random distance $c + \eta$, where the noise $\eta$ is drawn from a continuous and symmetric jump distribution $f(\eta)$, characterized by a Lévy index $0 < \mu \leq 2$. The jump has a finite second moment for $\mu = 2$, while for $0 < \mu < 2$ the second moment diverges. For this discrete-time series consisting of the successive positions of the biased walker, we presented complete analytical studies of the persistence and the record statistics. For the later, we studied the mean and the full distribution of the number of records up to step $n$ and also the statistics of the duration of records, in particular those for the longest and shortest lasting records. As a function of the two parameters $c$ and $0 < \mu \leq 2$, we found that it is necessary to distinguish between five different universal regimes, as shown in the basic phase diagram in figure 1. In these five regimes, the persistence and the record statistics exhibit very different asymptotic behaviors that are summarized in section 2 and we do not repeat them here. For instance, the growth of the mean record number with $n$ in all five regimes is summarized in the simulation results in figure 21, in complete agreement with our analytical predictions. The main conclusion is that even though this is a rather simple model, it exhibits very rich and varied universal behaviors for record statistics and persistence depending on the two parameters $c$ and $0 < \mu \leq 2$.

Our results provide a simple yet nontrivial, but fully solvable model for the record statistics, a subject that has gained considerable interest over the last few years. Our results provide one generalization of the previous results for record statistics for symmetric random walks [22]. However, it is important to note that this extension does not yet cover all possible kinds of discrete-time random walks. In principle, one could consider more complicated asymmetries of the jump distribution. It might be interesting to consider a jump distribution that has different tail exponents in the left and in the right tail. Also, a generalization of these results to an asymmetric lattice random walk is still missing. In [22], a symmetric lattice random was also considered. It should be possible to compute the record statistics of a lattice random walk that has a higher probability of jumping in one direction than in the other.
Figure 21. The figure shows numerical results for the mean record number $<R_n>$ for biased random walks from all five regimes. For regimes I–IV, we used a positive bias of $c = 1$, and in regime V, we simulated a Gaussian random walk (with $\sigma = 1$) with a negative bias of $c = -0.01$. For each jump distribution, we averaged over $10^4$ samples. In all these cases, as shown in detail in the previous figures, the asymptotic behavior agree very well with our analytical predictions (which are not shown in this figure for clarity).

It might be interesting to see if our results can be applied to financial data, similar to the analysis in [11, 12]. Daily stock data however proved not to be useful for comparison because the asymptotic limit is hardly achieved in the available observational data. An application to stock data with a higher temporal resolution however should be possible and might provide new insights. Such an analysis is definitely an interesting subject for future research. Also, the distribution of records in stock prices has not been analyzed in detail before and it would be interesting to see if such an analysis for available data can be fitted to our theoretical distributions.

Acknowledgments

SNM and GS acknowledge support by ANR grant 2011-BS04-013-01 WALKMAT and in part by the Indo-French Centre for the Promotion of Advanced Research under Project 4604-3. GW is grateful for the kind hospitality of the Laboratoire de Physique Théorique et Modèles Statistiques during the completion of this work and for the financial support provided by DFG within the Bonn Cologne Graduate School of Physics and Astronomy.

Appendix A. The constant $A_I$

The constant $A_I$ in equation (104) can be directly expressed in terms of $\tilde{f}(k)$ as announced in equation (24). To derive this, we use the explicit expression of $P_n(x)$ from equation (6) in the expression for $A_I$ and integrate over $x$ to obtain

$$A_1 = \frac{2}{\sqrt{\pi}} \exp \left[ \sum_{n=1}^{\infty} \frac{1}{n} \int_{-\infty}^{\infty} \frac{dk}{2\pi} |\tilde{f}(k)|^n \frac{1-e^{-ikcn}}{ik} \right]. \quad (A.1)$$
Next, we use the symmetry $\hat{f}(k) = \hat{f}(-k)$ that leads to
\[
A_1 = \frac{2}{\sqrt{\pi}} \exp \left[ \frac{1}{\pi} \int_0^\infty \frac{dk}{k} \sum_{n=1}^\infty \frac{\sin(kn)}{n} \left(\hat{f}(k)\right)^n \right].
\] (A.2)

The sum on the rhs can explicitly be evaluated using the identity
\[
\sum_{n=1}^\infty \frac{x^n}{n} \sin(an) = \arctan \left[ \frac{x \sin(a)}{1 - x \cos(a)} \right],
\] (A.3)

which then leads to the exact expression in equation (24).

We then analyze the behavior of $A_1$ when $|c|$ is large and in the case where $\hat{f}(k) = \exp(-|k|^\mu)$, with $\mu < 1$. In that case, one has $P_n(x) = n^{-1/\mu} L_\mu(x/n^{1/\mu})$ for all $n$ and it is easier to start from the formula given in the text in equation (104):
\[
A_1 = \frac{2}{\sqrt{\pi}} e^{S_0}, \quad S_0 \equiv S_0(c) = \sum_{n=1}^\infty \frac{1}{n} \int_0^\infty L_\mu(x/n^{1/\mu}) \, dx/n^{1/\mu}.
\] (A.4)

Note that, given that $P_n(x) = P_n(-x)$, one has $S_0(c) = S_0(-c)$ and we thus present the analysis for $c > 0$. Performing the change of variable $y = x/n^{1/\mu}$ in the integral (A.4), we write
\[
S_0(c) = \sum_{n=1}^\infty \frac{1}{n} \int_0^\infty \frac{c}{n^{1/\mu}} L_\mu(y) \, dy,
\] (A.5)

and take the derivative with respect to $c$
\[
S_0'(c) = \sum_{n=1}^\infty \frac{1}{n} n^{-\frac{1}{\mu}} L_\mu \left( \frac{c}{n^{1/\mu}} \right).
\] (A.6)

In this expression, one notes that $c/n^{1/\mu} = (n/c^{1/\mu})^{-1/\mu}$ so that when $c \to \infty$ the discrete sum over $n$ in equation (A.6) can be replaced by an integral (we recall that $\mu < 1$ here), which leads to
\[
S_0'(c) \sim \frac{1}{c} \int_0^\infty L_\mu \left( \frac{c^{1/\mu}}{y} \right) y^{-1/\mu} \, dy.
\] (A.7)

Finally, performing the change of variable $z = y^{1/\mu}$ in equation (A.7) yields
\[
S_0'(c) \sim \frac{1}{c} \frac{\mu}{1 - \mu} \int_0^\infty L_\mu(z) \, dz = \frac{1}{c} \frac{\mu}{2(1 - \mu)},
\] (A.8)

so that one obtains
\[
A_1 = \frac{2}{\sqrt{\pi}} e^{S_0} \propto c^{-\frac{\mu}{1 - \mu}}, \quad c \to \infty.
\] (A.9)

This power-law behavior (A.9) can be understood from the following scaling argument. We are indeed interested in the records statistics of the variables $y_n$, with $y_n = x_n + cn (11)$, where $x_n$ behaves for large $n$ as $x_n = \mathcal{O}(n^{1/\mu})$. Therefore, for small $n, n < n^*$ when $c$ is large, $y_n$ is dominated by the drift term and $n^*$ is such that $cn^* \sim n^{1/\mu}$, which yields
\[
n^* \sim c^{1/\mu}.
\] (A.10)

On the other hand, for small $n, n < n^*$, $y_n$ is dominated by the (positive) drift and hence is almost deterministic that yields $\langle R_n \rangle \sim n$, for $n < n^*$ while $\langle R_n \rangle \sim A_1 \sqrt{n}$ for $n > n^*$. By matching these two behaviors for $n = n^*$, one obtains
\[
A_1 \sim \sqrt{n^*} \propto c^{\frac{1}{2(1 - \mu)}},
\] (A.11)

which yields the result obtained above in equation (A.7).

Note finally that, by using $S_0(c) = -S_0(-c)$, one obtains
\[
A_1 \sim (-c)^{\frac{\mu}{1 - \mu}}, \quad c \to -\infty.
\] (A.12)
Appendix B. Computation of $\alpha_2(c) = a_2(|c|)$, $c < 0$, for an exponential jump distribution with $c < 0$

The expression for the amplitude $\alpha_2(c)$ in regime V (with $c < 0$) and for a general jump distribution is given in equation (101). By comparing with equation (125), we see that $\alpha_2(c < 0) = a_2(|c|)$, where $a_2(|c|)$ is the prefactor of the leading linear growth of mean record number in regime IV with the positive drift $|c|$. For a general jump distribution $f(\eta)$, we then have

$$\alpha_2(c) = \exp\left[-\sum_{n=1}^{\infty} \frac{1}{n} \int_{|c|}^{\infty} P_n(x) \, dx\right], \quad (B.1)$$

where we recall that $P_n(x) = \int_{-\infty}^{\infty} \frac{df}{\pi} \left[\hat{f}(k)\right]^n e^{-\eta k^2}$ and $\hat{f}(k) = \int_{-\infty}^{\infty} f(\eta) e^{\eta k} \, d\eta$ is the Fourier transform of the jump distribution. Thus, in general, computing the prefactor $\alpha_2(c) = a_2(|c|)$ explicitly is difficult for arbitrary $f(\eta)$. It can explicitly be computed for the Gaussian distribution where $P_n(x) = (2\pi \eta)^{-1/2} \exp[-x^2/(2\eta)]$ itself is Gaussian and $\alpha_2(c) = a_2(|c|)$ is then given by the formula given in equation (126). In this appendix, we show that $\alpha_2(c) = a_2(|c|)$ can also be explicitly computed for the symmetric exponential distribution $f(\eta) = (2b)^{-1} \exp(-|x|/b)$.

For this exponential jump distribution, the Fourier transform has the Lorentzian form: $\hat{f}(k) = 1/[\pi (b^2 k^2 + 1)]$. One can then substitute this into the expression for $P_n(x)$ and eventually into equation (B.1). After a quite convoluted computation involving contour integration in the complex plane, one can explicitly find $\alpha_2(c)$. However, as we show below, for the exponential case, there is an alternative simpler way to compute $\alpha_2(c)$ directly (without going through the formula given in equation (B.1)).

The first observation is that $\alpha_2(c)$ is just the limiting value of the persistence probability $Q(n)$ (the probability that the walker stays below 0 up to $n$ steps starting at 0) when $n \to \infty$ in the presence of a negative drift $c < 0$. By symmetry, $Q(n)$ is then also the probability that the walker, starting at the origin, stays above the origin up to $n$ steps, but in the presence of a positive drift $|c| > 0$. So, the idea is to compute this probability $Q(n)$ directly for the exponential jump distribution and then take the limit $n \to \infty$ to compute $\alpha_2(c) = Q(n \to \infty)$.

To compute $Q(n)$, we first define

$$q^n_{+}(y) = \text{Proba. that the random walker, starting at } y \geq 0 \text{ stays positive up to step } n. \quad (B.2)$$

If we can compute $q^n_{+}(y)$, then $Q(n)$ is simply obtained by putting the starting position to be 0, i.e. $Q(n) = q^n_{+}(0)$. To compute $q^n_{+}(y)$, we can write a backward recurrence relation for $q^n_{+}(y)$ by considering the jump that happens at the first step from $y$ to $y' \geq 0$:

$$q^n_{+}(y) = \int_{y}^{\infty} q^{n-1}_{+}(y') f(y + |c| - y') \, dy', \quad (B.3)$$

$$q^n_{+}(y) = 1 \text{ for } y \geq 0. \quad (B.4)$$

In the limit of large $n$, we expect that $q^n_{+}(y)$ approaches to an $n$-independent stationary value $q^+_{\infty}(y) \to q^+(y)$ that just denotes the eventual probability with which the walker escapes to infinity (starting from $y$) in the presence of a positive drift $|c|$. Taking $n \to \infty$ limit on both sides of equation (B.3) results the integral equation for $y \geq 0$

$$q^+(y) = \int_{y}^{\infty} q^+(y') f(y + |c| - y') \, dy'. \quad (B.5)$$

Note that this equation is valid for arbitrary jump distribution $f(\eta)$. This half-space Wiener–Hopf-type integral equation with an asymmetric kernel cannot be solved in general. However,
for the special case of the exponential distribution, \( f(\eta) = 1/(2b) \exp(-|\eta|/b) \), this integral equation (B.5) can be transformed into a differential equation using

\[
f''(\eta) = -\frac{1}{b^2} \delta(\eta) + \frac{1}{b^2} f(\eta).
\] (B.6)

By differentiating twice equation (B.5) with respect to \( y \), one then obtains (using equation (B.6))

\[
d^2q^+(y)/dy^2 = -\frac{1}{b^2} q^+(y + |c|) + \frac{1}{b^2} q^+(y).
\] (B.7)

Note that the solution \( q^+(y) \) must approach to 1 as \( y \to \infty \): \( q^+(y \to \infty) = 1 \). This follows from the fact that if the particle starts at the positive infinity, it escapes to positive infinity with the probability 1 in the presence of any positive drift \(|c| > 0\).

Note that the differential equation (B.7), though linear, is actually nonlocal in \( y \) due to the first term on the rhs and, hence, is still not completely trivial to solve. Fortunately, it turns out that it admits a solution of the form

\[
q^+(y) = 1 - F \exp(-\lambda y/b),
\] (B.8)

where \( F \) and \( \lambda \) are two dimensionless constants (independent of \( y \)) that are yet to be determined. Note that this ansatz manifestly satisfies the boundary condition \( q^+(y \to \infty) = 1 \). Substituting this ansatz into equation (B.7) we see that indeed equation (B.8) is a solution provided \( \lambda \) satisfies

\[
\exp(-\lambda |c|/b) = 1 - \lambda^2, \quad \text{with} \quad \lambda > 0.
\] (B.9)

The transcendental equation has a unique positive solution which then uniquely determines \( \lambda \). For example, for \( b/c = 1 \), we obtain using Mathematica the root \( \lambda = 0.714556 \ldots \). But we still need to determine the prefactor \( F \) in the ansatz in equation (B.8). The amplitude \( F \) in equation (B.8) is obtained by injecting this solution back into the integral equation (B.7) and performing the integral. Indeed, one finds that equation (B.8) is a solution of the integral equation provided

\[
F = 1 - \lambda.
\] (B.10)

This then uniquely determines the solution of the integral equation (B.7)

\[
q^+(y) = 1 - (1 - \lambda) \exp(-\lambda y/b),
\] (B.11)

where \( \lambda \) is the unique positive solution of the transcendental equation (B.9).

Noting finally that \( \alpha_2(c) = Q(n \to \infty) = q^+(0) \) results

\[
\alpha_2(c) = \alpha_2(|c|) = q^+(0) = \lambda,
\] (B.12)

where \( \lambda > 0 \) is the solution of equation (B.9). We have checked that we indeed obtain exactly the same expression by evaluating the original general expression in equation (B.1) for the exponential jump distribution, though this was not completely trivial to check (we do not give details of this check here).

References

[1] Gembris D, Taylor J G and Suter D 2002 Nature 417 506
[2] Gembris D, Taylor J G and Suter D 2007 J. Appl. Stat. 34 529
[3] Redner R and Petersen M R 2006 Phys. Rev. E 74 061114
[4] Meehl G A, Tebaldi C, Walton G, Easterling D and McDaniel L 2009 Geophys. Res. Lett. 36 L23701
[5] Wergen G and Krug J 2010 Europhys. Lett. 92 30008
[6] Anderson A and Kostinski A 2011 J. Appl. Meteorol. Climatol. 50 1859
[7] Krug J and Jain K 2005 Physica A 358 1
[8] Oliveira L P, Jensen H J, Nicodemi M and Sibani P 2005 Phys. Rev. B 71 104526
[9] Sibani P, Rodriguez G F and Kenning G G 2006 Phys. Rev. B 74 224407
[10] Godrèche C and Luck J M 2008 J. Stat. Mech. P11006
[11] Wergen G, Bogner M and Krug J 2011 Phys. Rev. E 83 051109
[12] Wergen G, Majumdar S N and Schehr G 2012 Phys. Rev. E 86 011119
[13] Foster F G and Stuart A 1954 J. R. Stat. Soc. 16 1
[14] Arnold B C, Balakrishnan N and Nagaraja H N 1998 Records (New York: Wiley)
[15] Nevzorov V B 2001 Records: Mathematical Theory (Providence, RI: American Mathematical Society)
[16] Ballerini R and Resnick S 1985 J. Appl. Probab. 22 487
[17] Franke J, Wergen G and Krug J 2010 J. Stat. Mech. P10013
[18] Wergen G, Franke J and Krug J 2011 J. Stat. Phys. 144 1206
[19] Franke J, Wergen G and Krug J 2012 Phys. Rev. Lett. 108 064101
[20] Krug J 2007 J. Stat. Mech. P07001
[21] Weiss G H 1994 Aspects and Applications of the Random Walk (Amsterdam: North-Holland)
[22] Majumdar S N and Ziff R M 2008 Phys. Rev. Lett. 101 050601
[23] Sparre Andersen E 1954 Mat. Scand. 2 195
[24] Majumdar S N 2010 Lecture Notes for the Summer School, Fundamental Problems in Statistical Physics: XII (Leuven, Belgium, 2009) Physica A 389 4299
[25] Sabhapandit S 2011 Europhys. Lett. 94 20003
[26] Le Doussal P and Wiese K J 2009 Phys. Rev. E 79 051105
[27] Edery Y, Kostinski A and Berkowitz B 2011 Geophys. Res. Lett. 38 L16403
[28] Majumdar S N 1999 Curc. Sci. 77 370
[29] Bouchaud J-P and Georges A 1990 Phys. Rep. 195 127
[30] Metzler R and Klafter J 2000 Phys. Rep. 339 1
[31] Evans M R, Majumdar S N and Zia R K P 2006 J. Stat. Phys. 123 357
[32] Schehr G and Le Doussal P 2010 J. Stat. Mech. P01009
[33] Pitman J and Yor M 1997 Ann. Probab. 25 855
[34] Godrèche C, Majumdar S N and Schehr G 2009 Phys. Rev. Lett. 102 240602