Record statistics for random walks and Lévy flights with resetting

Satya N Majumdar\(^1\), Philippe Mounaix\(^2\), Sanjib Sabhapandit\(^3\) and Grégory Schehr\(^4\),\(^*\)

\(^1\) LPTMS, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay, France
\(^2\) CPHT, CNRS, Ecole Polytechnique, IP Paris, F-91128 Palaiseau, France
\(^3\) Raman Research Institute, Bangalore 560080, India
\(^4\) Sorbonne Université, Laboratoire de Physique Théorique et Hautes Energies, CNRS, UMR 7589, 4 Place Jussieu, 75252 Paris Cedex 05, France

E-mail: satya.majumdar@lptms.u-psud.fr, philippe.mounaix@polytechnique.edu, sanjib.sabhapandit@gmail.com and schehr@lpthe.jussieu.fr

Received 8 October 2021, revised 26 November 2021
Accepted for publication 3 December 2021
Published 28 December 2021

Abstract

We compute exactly the mean number of records \(\langle R_N \rangle\) for a time-series of size \(N\) whose entries represent the positions of a discrete time random walker on the line with resetting. At each time step, the walker jumps by a length \(\eta\) drawn independently from a symmetric and continuous distribution \(f(\eta)\) with probability \(1 - r\) (with \(0 \leq r < 1\)) and with the complementary probability \(r\) it resets to its starting point \(x = 0\). This is an exactly solvable example of a weakly correlated time-series that interpolates between a strongly correlated random walk series (for \(r = 0\)) and an uncorrelated time-series (for \((1 - r) \ll 1\)). Remarkably, we found that for every fixed \(r \in [0, 1]\) and any \(N\), the mean number of records \(\langle R_N \rangle\) is completely universal, i.e. independent of the jump distribution \(f(\eta)\). In particular, for large \(N\), we show that \(\langle R_N \rangle\) grows very slowly with increasing \(N\) as \(\langle R_N \rangle \approx (1/\sqrt{r}) \ln N\) for \(0 < r < 1\). We also computed the exact universal crossover scaling functions for \(\langle R_N \rangle\) in the two limits \(r \rightarrow 0\) and \(r \rightarrow 1\). Our analytical predictions are in excellent agreement with numerical simulations.

Keywords: record statistics, resetting dynamics, random walks, extreme statistics

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

The study of record statistics in a time series has generated a lot of interest across disciplines. For instance, this time series may represent the stock price in finance or the daily temperature in a given city. Examples of such time series arise naturally in sports, climate data, finance, disordered systems, earthquake models etc, where record statistics provide a useful tool [1–22] to analyse the extremes and rare fluctuations present in the data—see references [23, 24] for recent reviews on record statistics. Our starting point is a generic time series with \( N \) entries \( \{x_1, x_2, \ldots, x_N\} \). A record occurs at step \( m \) if \( x_m > \{x_1, x_2, \ldots, x_{m-1}\} \), i.e. the \( m \)-th entry is bigger than all the previous entries. How many records \( R_N \) occur in this time series of size \( N \)? It is an easily measurable observable, that carries interesting informations about the correlations structure present in the time series. When the entries are random variables (either independent or correlated), \( R_N \) is also a random variable and studying its statistics is what is called ‘record statistics’. The study of record statistics started in the statistics and probability literature [25–29], but off late it has generated much interest in the statistical physics literature also, in various contexts such as in biological evolution models [30–34], random walks (RWs) [35–51], avalanches in disordered systems [20, 36, 52] etc. From the statistical physics point of view, one of the central issues is the universality of the statistics of \( R_N \) with respect to the underlying probability distribution of the entries [24].

The simplest and the most well studied example corresponds to the case where the entries are uncorrelated and each drawn independently from the same parent distribution \( p(x) \)—usually referred to as the independent and identically distributed (IID) model [25–28]. In this IID case, the statistics of \( R_N \) is completely universal for all \( N \) as long as \( p(x) \) is continuous. For example, the mean number of records in an IID time series of \( N \) entries is given by the universal formula (see for example the review [24])

\[
\langle R_N \rangle = 1 + \frac{1}{2} + \cdots + \frac{1}{N}.
\]

(1)

In particular, for large \( N \), the mean number of records grows very slowly as

\[
\langle R_N \rangle \approx \ln N, \quad \text{as } N \to \infty.
\]

(2)

This is because it gets harder and harder to break a record as \( N \) increases.

A completely contrasting case with strongly correlated entries corresponds to the RW model where the entry \( x_n \) corresponds to the position of a RW (in discrete time and continuous space) after \( n \) steps starting from \( x_0 = 0 \). The position of the random walker evolves via the Markov rule

\[
x_n = x_{n-1} + \eta_n,
\]

(3)

where the jump lengths \( \eta_n \)'s are IID, each drawn from a symmetric and continuous distribution \( f(\eta) \). Note that this model also includes Lévy flights where the jump distribution has a fat tail: \( f(\eta) \sim |\eta|^{-1-\mu} \) for large \( |\eta| \) and the Lévy index \( 0 < \mu \leq 2 \). Here, unlike the IID model, the entries/positions \( x_n = \sum_{k=1}^{n} \eta_k \) for different \( n \) are strongly correlated, even though the increments \( \eta_k \)'s are uncorrelated. In this case, the record statistics can be computed exactly and, interestingly, the statistics of \( R_N \) is again completely universal for all \( N \), i.e. independent of the jump distribution \( f(\eta) \), as long as \( f(\eta) \) is symmetric and continuous. The origin of the universality in the RW model turns out to be very different from that in the IID model [24, 35].
RW model, the mean number of records is given by the universal formula \[35\]
\[
\langle R_N \rangle = (2N + 1) \left( \frac{2N}{N} \right) 2^{-2N} - 1, \tag{4}
\]
where, by convention, the first entry \( x_0 = 0 \) is not counted as a record. For large \( N \), the mean number of records \( \langle R_N \rangle \) grows much faster than in the IID model, namely,
\[
\langle R_N \rangle \simeq \frac{2}{\sqrt{\pi}} \sqrt{N}, \quad \text{as } N \to \infty. \tag{5}
\]

It is natural then to ask what happens to the record statistics for a time series whose entries are neither IID nor strongly correlated as in the RW case, in other words when the entries are only ‘weakly correlated’. An ideal candidate for a model where this question can be addressed is in fact the random walk with resetting \([53, 54]\). RWs (or Brownian motion (BM) in continuous time) with resetting have been extensively studied in the recent past \([53–69]\)—see reference \([70]\) for a recent review. There are two aspects of BM with resetting dynamics that have attracted much attention. (i) The first interesting fact is that, in the presence of a resetting protocol, where a diffusing particle goes back to its initial position at a constant rate, the dynamics violates the detailed balance and the position distribution of the particle reaches a non-equilibrium stationary state at long times \([53, 54]\). The relaxation dynamics to this stationary state turns out to be rather unusual \([58]\). Furthermore, since the BM with resetting reaches a stationary state at long times, the two-time correlation function between positions \( \langle x(t_1) x(t_2) \rangle \) decays exponentially with the time difference \(|t_1 - t_2|\) \([71, 72]\). These ‘weak correlations’ differ strongly from the standard BM without resetting where the two-time correlation function \( \langle x(t_1) x(t_2) \rangle = 2D \min(t_1, t_2) \) (where \( D \) is the diffusion constant) does not decay and \( x(t_1) \) and \( x(t_2) \) remain strongly correlated even when \( t_2 \gg t_1 \). Thus the position \( x(t) \) as a function of time in the BM with a nonzero resetting is an ideal candidate for a time series with weak correlations. (ii) The second aspect is that a finite resetting rate renders the mean first-passage time (MFPT) of the Brownian particle to a fixed target finite—in sharp contrast to the BM without resetting where the MFPT is infinite. Moreover, there is usually an optimal resetting rate that minimises the MFPT, a fact that has now been verified in experiments on optical traps \([73–75]\).

The existence of an optimal resetting rate makes this model quite attractive in the context of random search strategies \([53, 54, 57, 61, 62, 64, 76–81]\).

In this paper, our goal is to show that a discrete-time version of the BM with resetting \([57]\) can also be used successfully to explore other aspects going beyond these two issues (i) and (ii) mentioned above, namely to understand the record statistics in a weakly correlated time series. Here, the entries of our time series \( \{x_0 = 0, x_1, x_2, \ldots, x_N\} \) correspond to the position of a random walker on the line in the presence of a constant resetting probability \( 0 \leq r < 1 \) to the origin. The position \( x_n \) now evolves by the stochastic rule
\[
x_n = \begin{cases} 
0, & \text{with proba. } r \\
x_{n-1} + \eta_n, & \text{with proba. } 1 - r,
\end{cases} \tag{6}
\]
where the jump length \( \eta_n \) is drawn independently at each step from a continuous and symmetric distribution \( f(\eta) \). In figure 1, we show schematically a typical trajectory of a RW with resetting, where the record values are marked by black dots.

For \( r = 0 \), this model \((6)\) reduces precisely to a standard RW/Lévy flight without resetting \((3)\). In the opposite limit \((1 - r) \ll 1 \), where the particle is reset to the origin with a high probability at all steps, only few entries, separated by an exponentially long time interval
Figure 1. A schematic trajectory of a RW with resetting. The walker starts at the origin and evolves by equation (6). With probability $r$, it resets to the origin, shown by the red arrows. A position $x_k$ at step $k$ is a record if $x_k$ exceeds all previous values. The record values in this trajectory are marked by black dots.

Figure 2. The resetting random walker model studied here (6) has a single parameter, namely the resetting probability $0 \leq r < 1$. This model reduces to the standard random walk (marked RW) in the $r = 0$ limit, while it corresponds to the IID model in the opposite limit $(1 - r) \ll 1$.

(distributed as $r^n$ where $n$ is the time interval), would be nonzero. Since only these nonzero entries are relevant for the record statistics (clearly an entry with value 0 can not be a record), they are effectively uncorrelated in the limit $r \to 1$. Hence, when $r$ is close to 1, one expects to recover the record statistics for the IID model. Thus this RW with resetting model, with parameter $r$, nicely interpolates between the two well studied models, namely the IID [$(1 - r) \ll 1$ limit] and the RW ($r = 0$) cases discussed before (see figure 2). We recall that in the IID case [$(1 - r) \ll 1$ limit], the mean number of records grows universally as $\langle R_N \rangle \sim \ln N$ for large $N$. In the opposite limit $r = 0$ (RW model), the mean number of records again grows universally, but much faster, as $\langle R_N \rangle \sim \sqrt{4N/\pi}$ for large $N$. It is then interesting to know how $\langle R_N \rangle$ behaves for intermediate values of the resetting probability $0 \leq r < 1$. Does it stay universal, i.e. independent of the jump distribution $f(\eta)$ as in the RW case?

Here, we compute exactly $\langle R_N \rangle$ for all $0 \leq r < 1$ and show that it depends continuously on $r$, nicely interpolating between the two known limits. However, to our complete surprise, we found that for every $r$, the average $\langle R_N \rangle$ is completely universal, i.e. independent of the jump distribution $f(\eta)$ as long as $f(\eta)$ is symmetric and continuous. Let us briefly summarize our main results. We show that the generating function (GF) of $\langle R_N \rangle$ is given by the exact formula

$$\tilde{R}(z) = \sum_{N=0}^{\infty} \langle R_N \rangle z^N = \frac{\sqrt{1 - (1-r)z}}{r z (1-z)} \ln \left[ \frac{1 - (1-r)z - rz \sqrt{1 - (1-r)z}}{1 - z} \right],$$

(7)

which is completely independent of the jump distribution $f(\eta)$. Thus this result holds not only for RW’s whose jump distributions have a finite variance, but also for Lévy flights where the jump distribution has a fat tail. By analysing this formula (7), we show that, asymptotically for
large \( N \) and fixed \( 0 < r < 1 \), the mean number of records grows as

\[
\langle R_N \rangle = \frac{1}{\sqrt{r}} \ln N + \frac{1}{\sqrt{r}} (\gamma_E + \ln(r(1 - \sqrt{r}))) + O(1/N),
\]

(8)

where \( \gamma_E = 0.57721 \ldots \) is the Euler constant. Note that this asymptotic large \( N \) formula holds strictly for \( r < 1 \). We then consider the two limiting cases \( r \to 0 \) and \( r \to 1 \) to see how \( \langle R_N \rangle \) converges respectively to the RW model and the IID model. We show that in the limit \( r \to 0 \), \( N \to \infty \) keeping the product \( rN \) fixed, \( \langle R_N \rangle \) takes the scaling form

\[
\langle R_N \rangle \approx \frac{1}{\sqrt{r}} f_1(rN),
\]

(9)

where \( f_1(u) \) is a universal scaling function that we compute explicitly (see equations (46) and (47)). It has the leading asymptotic behaviours

\[
f_1(u) \approx \begin{cases} 
\frac{2}{\sqrt{\pi}} \sqrt{u}, & u \to 0, \\
\ln u + \gamma_E, & u \to \infty.
\end{cases}
\]

(10)

Interestingly, the same scaling function \( f_1(u) \) also appears in the expected maximum up to time \( t \) of a continuous time BM with resetting \[69\]. In section 3, we will make a connection between the mean number of records and the expected maximum. From equation (9) and the small \( u \) behaviour of \( f_1(u) \) in equation (10), one recovers the RW result in equation (5). Similarly, in the opposite limit \( r \to 1 \), \( N \to \infty \) keeping the product \( (1 - r)N \) fixed, we show that \( \langle R_N \rangle \) takes the scaling form

\[
\langle R_N \rangle \approx f_2((1 - r)N),
\]

(11)

where \( f_2(v) \) is also a universal scaling function given by

\[
f_2(v) = -\text{Ei} \left(-\frac{v}{2}\right) + \gamma_E + \ln \left(\frac{v}{2}\right),
\]

(12)

where \( -\text{Ei}(z) = \int_z^\infty e^{-t}/t \, dt \). The scaling function has the leading asymptotic behaviours

\[
f_2(v) \approx \begin{cases} 
\frac{v}{2}, & v \to 0, \\
\ln \left(\frac{v}{2}\right) + \gamma_E, & v \to \infty.
\end{cases}
\]

(13)

From equation (11) and the asymptotic behaviour of \( f_2(v) \) as \( v \to \infty \) in equation (13), one recovers the leading behaviour \( \langle R_N \rangle \approx \ln N \) in the IID model in equation (2). Our numerical simulations are in perfect agreement with our analytical predictions.

The rest of the paper is organized as follows. In section 2, we present the main derivation of the result for the GF of \( \langle R_N \rangle \) in equation (7) and demonstrate precisely how the universality emerges. In section 3, we derive the asymptotic behaviour of \( \langle R_N \rangle \) for large \( N \), starting from equation (7). We also compute the two scaling functions \( f_1(u) \) and \( f_2(v) \) in the two limits \( r \to 0 \) and \( r \to 1 \) respectively. In that section, we also compare our analytical predictions with numerical simulations. Finally, we conclude in section 4 with some remarks and open questions.
2. Computing the exact GF of $\langle R_N \rangle$

We start with a discrete-time series with $N+1$ entries $\{x_0 = 0, x_1, x_2, \ldots, x_N\}$ where the positions $x_n$'s evolve via the stochastic rule in equation (6), i.e. $x_n$ corresponds to the position of a random walker at step $n$, with a resetting probability $r$ to the origin at each step. We recall that a record occurs at step $n$ if $x_n > \max\{x_0 = 0, x_1, \ldots, x_{n-1}\}$, i.e. it exceeds all the previous positions. We use the convention that $x_0 = 0$ is not a record. The reason for this convention will be clear later. We want to compute the average number of records $\langle R_N \rangle$ up to step $N$. To compute this quantity, it is convenient to introduce a binary variable $\sigma_n = \{1, 0\}$ at each step such that

$$\sigma_n = \begin{cases} 1 & \text{if a record occurs at step } n, \\ 0 & \text{otherwise.} \end{cases}$$  \hspace{1cm} (14)

Note that, by our convention, $\sigma_0 = 0$. For all $n \geq 1$, the indicator function $\sigma_n$ is a binary random variable that fluctuates from one realisation of the time series to another. For each realisation of the time series, the number of records $R_N$ (which is also a random variable) can be expressed as a sum

$$R_N = \sum_{n=1}^{N} \sigma_n.$$  \hspace{1cm} (15)

Taking the average over all realisations, we get

$$\langle R_N \rangle = \sum_{n=1}^{N} \langle \sigma_n \rangle.$$  \hspace{1cm} (16)

Since $\sigma_n = \{1, 0\}$ is a binary random variable, its average $\langle \sigma_n \rangle$ is just the probability that a record occurs at step $n$. Below, we show how to compute $\langle \sigma_n \rangle$ and from it, the mean number of records $\langle R_N \rangle$ using equation (16).

Suppose that a record occurs at step $n$, with a record value $x > 0$ (see figure 3). Let $P_r(x, n)$ denote the probability that the process reaches the level $x$ for the first time at step $n$, starting from $x = 0$. Here, the subscript $r$ refers to the presence of a nonzero resetting probability $r$. Hence $\langle \sigma_n \rangle$, the probability that a record occurs at step $n$ can be obtained by integrating $P_r(x, n)$ over all possible $x > 0$ at step $n$, i.e.

$$\langle \sigma_n \rangle = \lim_{\epsilon \to 0} \int_{\epsilon}^{\infty} P_r(x, n) \, dx.$$  \hspace{1cm} (17)

The reason for putting a cutoff $\epsilon$ is the following. For a record to happen at step $n$, the value of the process at step $n$ must be strictly bigger than 0, since the process starts at $x = 0$. This is true for any process. In the absence of resetting, one can set $\epsilon = 0$ in the outset, because the event that the process reaches exactly $x = 0$ at step $n$ is a set with measure zero. However, in the presence of a finite resetting probability $r$, the process resets to 0 with a nonzero probability and these events should not be counted as records. Hence, we need to impose a finite lower cutoff $\epsilon$ in the integral over $x$ in the integral in equation (17) to exclude these ‘spurious’ records with value 0 and eventually take the limit $\epsilon \to 0$.

Our next step is to estimate $P_r(x, n)$ using a renewal equation approach. For this, we need to first introduce a quantity $Q_r(x, n)$ that denotes the probability that the random walker (without resetting, hence the subscript ‘0’), starting at the origin, stays below the level $x > 0$ up to step
Figure 3. Schematic trajectory of a RW with resetting, starting from $x = 0$ up to $n$ steps (see equation (6)). The resetting events are marked by red arrows. If a record happens at step $n$ with value $x$, the trajectory reaches the level $x$ for the first time at step $n$.

$n$. We will refer to this as the survival probability of the walker without resetting. Then the renewal equation for $P_r(x, n)$ can be written as

$$P_r(x, n) = (1 - r)^n P_0(x, n) + r \sum_{m=0}^{n-1} (1 - r)^m Q_0(x, m) P_r(x, n - m - 1),$$

starting from $P_r(x, 0) = 0$, for all $x > 0$. This equation can be understood as follows. Suppose that there is no resetting up to $n$ steps. The probability for this event is $(1 - r)^n$. Hence, the first term on the right-hand side (rhs) of equation (18) corresponds to the event of ‘no resetting’. The second term on the rhs in equation (18) corresponds to the events where the number of resettings within $n$ steps is at least one. In this case, when there is at least one resetting event, suppose that the first resetting event occurs at step $m + 1$. The probability for this resetting event is $r(1 - r)^m$ since the first $m$ steps have no resetting [this happens with probability $(1 - r)^m$] followed by a resetting event exactly at step $m + 1$ (this occurs with probability $r$). During these first $m$ steps, the random walker evolves without resetting and the probability that this walk stays below the level $x$ up to step $m$ is just $Q_0(x, m)$. Finally, after the first resetting at step $m + 1$, the process renews itself, starting from the origin, and eventually arrives at $x$ for the first time at step $n$ of the original process, i.e. at step $n - m - 1$ for the renewed process starting at 0 at step $m + 1$. Thus the entire probability of the trajectory with first resetting at step $m + 1$ is given by the product $r(1 - r)^m Q_0(x, m) P_r(x, n - m - 1)$. Finally, noting that the first resetting can occur at all intermediate steps between 0 and $n$ we obtain the result in equation (18) by summing over all possible values of $m$. Notice that equation (18) holds for all $x \geq 0$ and all $n \geq 0$. For $n = 0$, we just interpret the sum in the second term in equation (18) as 0.

We next define the GFs

$$\tilde{P}_r(x, z) = \sum_{n=0}^{\infty} P_r(x, n) z^n$$

and

$$\tilde{Q}_0(x, z) = \sum_{m=0}^{\infty} Q_0(x, m) z^m.$$  

Multiplying both sides of equation (18) by $z^n$ and summing over $n$ from $n = 0$ to $n \to \infty$, we obtain, upon using the convolution structure of the second term in equation (18)

$$\tilde{P}_r(x, z) = \tilde{P}_0(x, (1 - r)z) + r z \tilde{Q}_0(x, (1 - r)z) \tilde{P}_r(x, z).$$

This allows us to express the GF $\tilde{P}_r(x, z)$ in terms of the GFs without resetting ($r = 0$), namely

$$\tilde{P}_r(x, z) = \frac{\tilde{P}_0(x, (1 - r)z)}{1 - rz \tilde{Q}_0(x, (1 - r)z)}.$$
For a standard RW/Levy flight without resetting, the first-passage probability $P_0(x, n)$ to level $x$ at step $n$ and the survival probability $Q_0(x, n)$ to stay below the level $x$ up to step $n$ are well studied [37, 82–87]. We first note that $Q_0(x, n)$ coincides with the probability that the random walker, starting at $x > 0$ initially, does not cross 0 up step $n$. This is easily seen by making a change of coordinates $y_n = x - x_n$. Similarly, $P_0(x, n)$ can be identified with the probability that the random walker, starting at the origin, reaches the level $x > 0$ at step $n$, without crossing the origin in between. This is seen first by making the transformation $y_n = x - x_n$, followed by the reversal of the time. The GF’s of these two probabilities for a RW can, in principle, be extracted from their Laplace transforms, which are known to satisfy [40, 85–88]

\[
\int_0^\infty P_0(x, z) e^{-\lambda x} \, dx = \phi(\lambda, z),
\]

(22)

\[
\int_0^\infty Q_0(x, z) e^{-\lambda x} \, dx = \frac{\phi(\lambda, z)}{\lambda \sqrt{1 - z}},
\]

(23)

where

\[
\phi(\lambda, z) = \exp\left(-\frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{\ln(1 - z f(k))}{\lambda^2 + k^2} \, dk \right),
\]

(24)

where $f(k) = \int_{-\infty}^{\infty} f(x) e^{ikx} \, dx$ is the Fourier transform of the jump distribution. Note that the dependence on the jump distribution $f(\eta)$ occurs only through the function $\phi(\lambda, z)$. Interestingly, the dependence on the jump distribution $f(\eta)$ of $Q_0(x, z)$ in equation (23) completely disappears in the limit $x \to 0$. To see this, one can perform the change of variable $\lambda x = u$ in equation (23) and take the limit $\lambda \to \infty$, using $\phi(\lambda \to \infty, z) = 1$. This gives

\[
\tilde{Q}_0(0, z) = \frac{1}{\sqrt{1 - z}} = \sum_{n=0}^{\infty} \left(\frac{2n}{n}\right) 2^{-2n} z^n.
\]

(25)

Thus $Q_0(0, n) = \left(\frac{2n}{n}\right) 2^{-2n}$ is universal for any $n$, i.e. independent of the jump distribution $f(\eta)$, as long as it is symmetric and continuous. This is known as the celebrated Sparre Andersen theorem [83]. Similarly, in the limit $x \to \infty$ one can extract $Q_0(x \to \infty, z)$ by taking the $\lambda \to 0$ limit in equation (23). Note that in the limit $\lambda \to 0$, one obtains $\phi(\lambda \to 0, z) = 1/\sqrt{1 - z}$ [83]. This gives

\[
\tilde{Q}_0(x \to \infty, z) = \frac{1}{1 - z} = \sum_{n=0}^{\infty} z^n.
\]

(26)

Hence $Q_0(x \to \infty, n) = 1$ for all $n \geq 0$, which of course is expected since, if the walker starts at $\infty$ it does not cross 0 up step $n$ with probability 1. Thus, in both limits $x \to 0$ and $x \to \infty$, the GF of the survival probability $Q_0(x, z)$ is universal. We will see shortly that these two limiting universal behaviours will play an important role in establishing the universality of the mean number of records.

There exists a nice relation between $P_0(x, z)$ and $Q_0(x, z)$ which will be useful shortly. To derive this relation, we integrate equation (23) by parts and we use the Sparre Andersen theorem (25), namely $Q_0(0, z) = 1/\sqrt{1 - z}$. This gives

\[
1 + \sqrt{1 - z} \int_0^\infty \frac{\partial Q_0(x, z)}{\partial x} \, e^{-\lambda x} \, dx = \phi(\lambda, z) = \int_0^\infty P_0(x, z) e^{-\lambda x} \, dx,
\]

(27)
where in the last equality we used the relation (22). Comparing the rhs and the left-hand side (lhs) of equation (27), we get

\[ \tilde{P}_0(x, z) = \delta(x) + \sqrt{1-z} \frac{\partial \tilde{Q}_0(x, z)}{\partial x}. \] (28)

Substituting this relation in equation (21), we obtain

\[ \tilde{P}_r(x, z) = \frac{\delta(x) + \sqrt{1-z} \partial_z \tilde{Q}_0(x, (1-r)z)}{1-z \tilde{Q}_0(x, (1-r)z)}. \] (29)

We now substitute this expression for \( \tilde{P}_r(x, z) \) on the rhs of equation (17). Due to a finite lower cut-off \( \epsilon \), the term containing \( \delta(x) \) in equation (29) does not contribute to the integral. Moreover, the second term can be integrated explicitly yielding

\[ \sum_{n=0}^{\infty} \langle \sigma_n \rangle z^n = \frac{\sqrt{1-(1-r)z}}{rz} \ln \left( \frac{1-z \tilde{Q}_0(0, (1-r)z)}{1-rz \tilde{Q}_0(x \to \infty, (1-r)z)} \right). \] (30)

Interestingly, the rhs of equation (30) involves only the limiting values of the GF \( \tilde{Q}_0(x, (1-r)z) \), namely the \( x \to 0 \) and \( x \to \infty \) limits. As shown before, \( \tilde{Q}_0(x, (1-r)z) \) is universal in both these limits. Consequently, the rhs of equation (30) is universal, i.e. independent of \( f(\eta) \) as long as it is symmetric and continuous. Indeed, using the results in equations (25) and (26), we get

\[ \sum_{n=0}^{\infty} \langle \sigma_n \rangle z^n = \frac{\sqrt{1-(1-r)z}}{rz} \ln \left( \frac{1-(1-r)z - rz \sqrt{1-(1-r)z}}{1-z} \right). \] (31)

Finally, using the relation (16), one gets the GF of the mean number of records \( \langle R_N \rangle \)

\[ \sum_{N=0}^{\infty} \langle R_N \rangle z^N = \frac{1}{1-z} \sum_{n=0}^{\infty} \langle \sigma_n \rangle z^n = \frac{\sqrt{1-(1-r)z}}{rz(1-z)} \ln \left( \frac{1-(1-r)z - rz \sqrt{1-(1-r)z}}{1-z} \right). \] (32)

Thus, remarkably, the dependence on \( \phi(\lambda, z) \), and hence on \( f(\eta) \), completely drops out and we get a universal result for any \( 0 \leq r < 1 \). Note that \( \langle R_N \rangle \) is universal, i.e. independent of \( f(\eta) \) for all \( N \), and not just for large \( N \). For example, by expanding the rhs of (32) in powers of \( z \), one can easily obtain \( \langle R_N \rangle \) for the first few values of \( N \) such as \( N = 1, 2 \) and 3

\[ \langle R_1 \rangle = \frac{1-r}{2}, \quad \langle R_2 \rangle = \frac{1}{8} (-r^2 - 6r + 7), \quad \langle R_3 \rangle = \frac{1}{16} (-r^3 - 3r^2 - 15r + 19). \] (33)

For general \( N \), it is hard to extract \( \langle R_N \rangle \) explicitly from its GF in equation (32). However, as shown in the next section, one can easily extract \( \langle R_N \rangle \) for large \( N \), by analysing the limit \( z \to 1 \) of the GF in equation (32).

In figure 4, we test the universality of \( \langle R_N \rangle \) vs \( N \) for all \( N \) by computing \( \langle R_N \rangle \) numerically for three different values of \( r \) and, for each \( r \), for three different jump distributions: (i) exponential, \( f(\eta) = (1/2)e^{-|\eta|} \) (for simplicity we call it ‘exponential’, even though it is actually a double-sided exponential distribution—sometimes called also the Laplace distribution), (ii) Cauchy, \( f(\eta) = (1/\pi)1/(1 + \eta^2) \) and (iii) Gaussian, \( f(\eta) = (1/\sqrt{2\pi})e^{-\eta^2/2} \). For each \( r \), we see that the
Figure 4. Plot of the average number of records $\langle R_N \rangle$ vs $N$ for three different values of $r = 0.01, 0.1$ and $r = 0.5$. For each value of $r$, the numerical data for three different jump distributions coincide for all $N$. In panel (a), we also show the exact results for $\langle R_N \rangle$ for smaller values of $N$ extracted from equation (32) by expanding the GF in powers of $z$ using Mathematica. The first three values ($N = 1, 2$ and 3) are given in equation (33). We find a perfect agreement between the numerical results and the analytical predictions. This also confirms the universality for even smaller values of $N$. In panel (b), we plot the analytical and the numerical results for $\langle R_N \rangle$ vs $N$ for a wider range of $N$ (plotted on a log-linear scale) and for three different values of $r$. For each $r$, we see that the three different jump distributions give identical results for all $N$, thus confirming the universality.

results for the three distributions coincide with each other and also with the exact analytical answer extracted from the GF in equation (32).

3. Asymptotic expansion for large $N$

In this section, we first analyse $\langle R_N \rangle$ for large $N$ with $r$ fixed. Then, we investigate the scaling limit $N \to \infty$ and $r \to 0$ with the product $u = rN$ fixed, to describe the crossover from the RW limit when $r$ increases from 0. In this limit, we also show how the result for $\langle R_N \rangle$ can be translated to the expected maximum of a continuous time BM with resetting. Finally, we also investigate the opposite scaling limit $N \to \infty$ and $r \to 1$ with the product $u = (1 - r)N$ fixed, to describe the crossover from the IID limit, when $r$ decreases from 1.

3.1. The limit $N \to \infty$ for fixed $r$

Here, we set $z = 1 - s$ with $s \to 0$ in equation (32). The lhs then converges to the Laplace transform

$$\sum_{N=0}^{\infty} \langle R_N \rangle z^N \approx \int_{0}^{\infty} \langle R_N \rangle e^{-sN} dN. \quad (34)$$

On the other hand, the rhs of equation (32), to leading orders in $s \to 0$ behaves as

$$\text{rhs} \approx -\frac{1}{\sqrt{r}} \ln \frac{s}{r} + \left[ \frac{1}{\sqrt{r}} \ln(r - r^{3/2}) \right] \frac{1}{s}. \quad (35)$$
Figure 5. The numerically obtained $\langle RN \rangle$ vs $N$ for large $N$ for the exponential jump distribution $f(\eta) = (1/2)e^{-|\eta|}$ for $r = 0.01, r = 0.1$ and $r = 0.5$ compared with the asymptotic formula in equation (38), where we kept only the two leading order terms for large $N$. The asymptotic result in equation (38) is expected to hold only for very large $N$, i.e. $N > N^*(r)$ where $N^*(r)$ is a crossover value. From the data for $r = 0.01$ we see that $N^*(r = 0.01) \approx 1000$, below which one needs to keep higher order terms in equation (38).

Equating the lhs and the rhs gives

$$\int_0^\infty \langle RN \rangle e^{-sN} dN \approx -\frac{1}{\sqrt{r}} \ln \frac{s}{s} + \left[ \frac{1}{\sqrt{r}} \ln(r - r^{3/2}) \right] \frac{1}{s}. \quad (36)$$

Using the inverse Laplace transform

$$\mathcal{L}^{-1}_{s \rightarrow N} \left( -\frac{\ln s}{s} \right) = \ln N + \gamma_E, \quad (37)$$

one can easily invert the Laplace transform in equation (36) to get the leading asymptotic behaviour of $\langle RN \rangle$ for large $N$

$$\langle RN \rangle = \frac{1}{\sqrt{r}} \ln N + \frac{1}{\sqrt{r}} \left( \gamma_E + \ln(r(1 - \sqrt{r})) \right) + O(1/N). \quad (38)$$

In figure 5, we compare this asymptotic result to the numerically obtained $\langle RN \rangle$ for three different values of $r$ and for the exponential jump distribution $f(\eta) = (1/2)e^{-|\eta|}$. We have also checked (not shown in the figure to avoid crowding) that, for a fixed $r$, exactly the same results are obtained for each $N$, for a Cauchy and a Gaussian jump distribution, as expected from the universality.

3.2. The limit $N \to \infty$, $r \to 0$ keeping $rN$ fixed

To study the large $N$ limit, we set, as before, $z = 1 - s$ with $s \to 0$. The lhs of equation (32) converges, as before to the Laplace transform in equation (34). We anticipate a scaling form
as in equation (9)—verified \textit{a posteriori}

\[ \langle R_N \rangle \approx \frac{1}{\sqrt{r}} f_1(rN). \] (39)

Substituting this scaling form in the Laplace transform in equation (34) and making a change of variable \( rN = u \), we get

\[ \sum_{N=0}^{\infty} \langle R_N \rangle z^N \approx \int_0^\infty \langle R_N \rangle e^{-N} dN \approx \frac{1}{r^{3/2}} \int_0^\infty f_1(u) e^{-\frac{u}{r}} du. \] (40)

Since we are considering the \( r \to 0 \) limit, and \( s \to 0 \) limit simultaneously, we see that we need to keep the ratio \( s/r = p \) fixed in equation (40). This gives

\[ \sum_{N=0}^{\infty} \langle R_N \rangle z^N \approx \int_0^\infty f_1(u) e^{-\frac{pu}{s}} du. \] (41)

On the rhs of equation (32), we set \( r = s/p \) with \( p \) fixed to be of \( O(1) \) and expand for small \( s \). This gives the small \( s \) behaviour of the rhs of equation (32) for fixed \( p \) as

\[ \text{rhs} \approx \frac{1}{s^{3/2}} \left[ \sqrt{\frac{p}{1+p}} \ln \left( \frac{1+p}{p} \right) \right]. \] (42)

Equating equations (41) and (42) gives us the exact Laplace transform of the scaling function \( f_1(u) \)

\[ \int_0^\infty f_1(u) e^{-\frac{pu}{s}} du = \sqrt{\frac{1+p}{p}} \ln \left( \frac{1+p}{p} \right). \] (43)

To invert this Laplace transform, we can use the following two identities

\[ \mathcal{L}^{-1}_{p=m} \left( \frac{\sqrt{p} + 1}{p} \right) = \frac{1}{\sqrt{\pi t}} e^{-t} + \text{erf}(\sqrt{t}), \] (44)

\[ \mathcal{L}^{-1}_{p=m} \left( \frac{1+p}{p} \right) = \frac{1}{t}(1 - e^{-t}), \] (45)

where \( \text{erf}(z) = (2/\sqrt{\pi}) \int_0^z e^{-t^2} dt \) is the error function. Now using the convolution theorem, we can explicitly write the scaling function (which is universal) as

\[ f_1(u) = \int_0^u dt \left( \frac{1}{\sqrt{\pi t}} e^{-t^2} + \text{erf}(\sqrt{t}) \right) \left( \frac{1 - e^{-\frac{(u-t)}{u-t}}}{u-t} \right). \] (46)

This integral seems hard to perform explicitly. However, its derivative \( f'_1(u) \) can be expressed in terms of a hypergeometric function as shown in the appendix A,

\[ f'_1(u) = \frac{e^{-u}}{\sqrt{\pi u}} {}_2F_2 \left( \begin{array}{c} 1, 1 \end{array} \right| \left( \frac{1}{2}, 2 \right), u \right). \] (47)

The asymptotic behaviour of \( f'_1(u) \) can be easily obtained from the hypergeometric function and one gets \( f'_1(u) \approx 1/\sqrt{u} \) as \( u \to 0 \) and \( f'_1(u) \approx 1/u \) as \( u \to \infty \). Integrating over \( u \), and using
Figure 6. Numerical verification of the scaling form in equation (39): $\sqrt{\langle R_N \rangle}$ plotted vs $u = rN$ for three different small values of $r$ and different values of large $N$. All the three curves collapse onto the same master curve, demonstrating the validity of the scaling form in equation (39). Furthermore, the numerically obtained scaling function matches perfectly with the analytical scaling function $f_1(u)$ in equation (46). The orange and the green dashed curves indicate the asymptotic behaviours for $u \to 0$ and $u \to \infty$ respectively in equation (48).

$f_1(0) = 0$, one gets the following asymptotic behaviour of the scaling function

$$f_1(u) \approx \begin{cases} 2 \sqrt{\frac{u}{\pi}}, & u \to 0, \\ \ln u + \gamma_E, & u \to \infty. \end{cases} \quad (48)$$

Note that the constant correction term $\gamma_E$ for large $u$ is not easy to obtain from the large $u$ expansion of $f_1'(u)$. However, it can be obtained by analysing the small $p$ behaviour of the Laplace transform in equation (43), where it behaves as $-(\ln p)/p$ for small $p$. In figure 6, we compare the analytical scaling function $f_1(u)$ with the numerically obtained scaling function, finding an excellent agreement.

If one now substitutes the small $u$ asymptotics in the scaling form in equation (39), one gets, as $r \to 0$ (for large $N$), $\langle R_N \rangle \approx (2/\sqrt{\pi}) \sqrt{N}$, thus recovering the leading asymptotic result for RW without resetting in equation (5). In contrast, in the limit $u \to \infty$ (i.e. $r$ small but fixed and $N \to \infty$), substituting the large $u$ asymptotic behaviour of $f_1(u)$ in the scaling form in equation (39) gives

$$\langle R_N \rangle \approx \frac{1}{\sqrt{r}} \ln(rN) \approx \frac{1}{\sqrt{r}} \ln N, \quad (49)$$

which matches smoothly with the leading order for large $N$ (for fixed $r$) in equation (38).

Note that this $\ln N$ asymptotics of $\langle R_N \rangle$ for large $N$ in equation (49) is reminiscent of the IID model in equation (2), however with an important difference that the prefactor $1/\sqrt{r}$ still carries information about the correlations between the entries in the $r \to 0$ limit. In fact, this
This consists of the following steps. It is convenient to compute first the renormalisation group argument discussed before in the context of extreme value statistics [22]. This consists of the following steps. It is convenient to compute first the renormalisation group argument discussed before in the context of extreme value statistics [22].

(a) We first divide the $N$ entries (i.e. the positions of the walker) into blocks of size of the order of the correlation time $\tau = 1/r$. So, the number of blocks is $N_b \approx N/(1/r) = rN$.

In general, $N_b$ is a random variable but for large $N$ one can replace it by its average value, which is indeed given by $rN$ for small $r$.

(b) The entries from block to block are independent. Hence, the probability that the global maximum belongs to the last block is $1/N_b = 1/(rN)$.

(c) Given that the global maximum belongs to the last block: what is the probability $p(r)$ that the global maximum is actually the last entry (at step $N$ of the full time series) of the last block, so that it is a record? One can easily compute this probability $p(r)$ as follows. In the last block, we have a pure RW of say $m$ steps, i.e. $(m + 1)$ entries. The probability that the last entry (within the last block of $m + 1$ steps) is a record is known to be given by the survival probability $Q_d(x = 0,m)$ [40], which is given by the Sparre Andersen theorem $Q_d(x = 0,m) = \left(\frac{2m}{m}\right) 2^{-2m}$ (see equation (25)). However, the size $m$ of the last block is a random variable distributed via $r(1 - r)^m$, where $m = 0, 1, 2, \ldots$ (i.e. resetting at the first entry followed by no resetting during $m$ consecutive steps). Hence, the probability that the last entry of the last block is a record within the block is given by

\[
p(r) = \sum_{m=0}^{\infty} r(1 - r)^m \left(\frac{2m}{m}\right) 2^{-2m} = \frac{r}{\sqrt{1 - (1 - r)}} = \sqrt{\tau}, \quad (50)
\]

where we used the result that $\sum_{m=0}^{\infty} \left(\frac{2m}{m}\right) 2^{-2m} = 1/\sqrt{1 - \varepsilon}$ as in equation (25).

(d) Hence, the record rate, i.e. $\langle \sigma_N \rangle$, at step $N$ (which is the probability that a record happens at step $N$) is therefore

\[
\langle \sigma_N \rangle = \text{Prob.} \left[ \text{the global maximum belongs to the last block} \right] \times \text{Prob.} \left[ \text{the maximum in the last block occurs at the last step} \right] \approx \frac{1}{N_b} \sqrt{\tau} \approx \frac{1}{\sqrt{\tau N}}, \quad (51)
\]

where we used $N_b \approx rN$. Finally, using $\langle R_N \rangle - \langle R_{N-1} \rangle = (\sigma_N)$, we get the result $\langle R_N \rangle \approx (1/\sqrt{\tau}) \ln N$, as obtained exactly in equation (49). Thus this argument clearly shows that the prefactor $1/\sqrt{\tau}$ comes from the correlation time $1/r$.

Connection between the mean number of records and the average maximum of a RW with resetting. We have seen above that in the limit $r \to 0$, $N \to \infty$, keeping $rN$ fixed, $\langle R_N \rangle \approx (1/\sqrt{\tau}) f_1(rN)$ where $f_1(u)$ is given in equation (46). It is convenient to rewrite it as

\[
\langle R_N \rangle \approx \sqrt{NF} \, f_1(rN), \quad \text{where } F_1(u) = \frac{1}{\sqrt{u}} f_1(u). \quad (52)
\]
For the same RW with resetting (6), another interesting observable is \( M_N = \max \{ x_0 = 0, x_1, \ldots, x_N \} \). One can ask: what is the average maximum \( \langle M_N \rangle \)? Intuitively, one may expect that this may be related to the mean number of records \( \langle R_N \rangle \). For a discrete time lattice RWs with jumps \( \pm 1 \), it is easy to see that \( \langle R_N \rangle = \langle M_N \rangle \), since whenever a new record happens, both \( R_N \) and \( M_N \) increase by 1 [40]. In contrast, for a RW with continuous jump distribution \( f(\eta) \) with a finite variance \( \sigma^2 \), it was shown in reference [40] that these two observables are proportional, but not exactly equal, namely

\[
\langle M_N \rangle \approx \frac{\sigma}{\sqrt{2}} \langle R_N \rangle.
\]  

(53)

On general grounds, one would conjecture this relation to be still valid in the presence of resetting. Below we show that this conjecture is certainly true in the scaling limit \( r \to 0, N \to \infty \), keeping \( rN \) fixed.

Indeed, assuming that this conjecture (53) is true, substituting the scaling form (52) in equation (53), one would predict that

\[
\langle M_N \rangle \approx \sqrt{\frac{N}{2}} \sigma F_1(rN).
\]  

(54)

In this scaling limit, the RW with resetting converges to the continuous time BM with resetting once we identify

\[
r = \tilde{r} \Delta t, \quad N = \frac{t}{\Delta t} \quad \text{and} \quad \sigma^2 = 2D \Delta t,
\]  

(55)

where \( \tilde{r} \) is the resetting rate and \( t \) denotes the continuous time in units of \( \Delta t \to 0 \). In addition, one needs to scale the variance as above where \( D \) denotes the diffusion constant of the BM. Substituting these relations (55) in equation (54) one would get the prediction that, for a continuous time BM with resetting rate \( \tilde{r} \), the expected maximum \( \langle M(t) \rangle \) up to time \( t \) would be given by

\[
\langle M(t) \rangle \approx \sqrt{Dt} F_1(\tilde{r}t) = \sqrt{\frac{D}{\tilde{r}}} f_1(\tilde{r}t),
\]  

(56)

with \( f_1(u) \) given in equation (46). Note that for a continuous time BM, this relation is expected to hold for all \( t \) and all \( \tilde{r} > 0 \). In fact, the expected maximum for a continuous time BM was recently computed in reference [69] in the context of computing the mean perimeter of the convex hull of a 2d resetting BM and the result in (56) coincides exactly with that in reference [69]. This validates, \textit{a posteriori}, the conjecture (53) for discrete-time random walker, at least in this scaling limit \( r \to 0, N \to \infty \), keeping \( rN \) fixed, where it converges to continuous time BM. It will be interesting to prove this conjecture (53) for a discrete-time random walker (with a finite variance of the continuous and symmetric jump distribution) for arbitrary \( 0 < r < 1 \).

### 3.3. \textit{The limit } \( N \to \infty, r \to 1 \ \text{keeping } (1 - r)N \text{ fixed} \)

As before, we start from the exact expression for the GF in equation (32) and set \( z = 1 - s \) with \( s \to 0 \). In the \( r \to 1 \) limit, and \( N \to \infty \) limit, we now anticipate a scaling form, again justified \textit{a posteriori},

\[
\langle R_N \rangle \approx f_2((1 - r)N).
\]  

(57)
Figure 7. Numerical verification of the scaling form in equation (57): $\langle R_N \rangle$ plotted vs $v = (1 - r)N$ for three different small values of $r$ and different values of large $N$. All the three curves collapse onto the same master curve, demonstrating the validity of the scaling form in equation (57). Furthermore, the numerically obtained scaling function matches perfectly with the analytical scaling function $f_2(v)$ in equation (12). The orange and the green dashed curves indicate the asymptotic behaviours $v \to 0$ and $v \to \infty$ respectively in equation (13).

Substituting this scaling form on the rhs of equation (34), we get

$$\sum_{N=0}^{\infty} \langle R_N \rangle z^N \approx \int_0^\infty \langle R_N \rangle e^{-sN} dN \approx \frac{1}{1 - r} \int_0^\infty f_2(v) e^{-p \frac{v}{1 - r}} dv.$$  (58)

Since we are considering the $r \to 1$ limit, and $s \to 0$ limit simultaneously, we see that we need to keep the ratio $s/(1 - r) = p$ fixed in equation (40). This gives

$$\sum_{N=0}^{\infty} \langle R_N \rangle z^N \approx \frac{1}{s} \left[ p \int_0^\infty f_2(v) e^{-pv} dv \right].$$  (59)

On the rhs of equation (32), we set $(1 - r) = s/p$ with $p$ fixed to be of $O(1)$ and expand for small $s$. This gives the small $s$ behaviour of the rhs of equation (32) for fixed $p$ as

$$\text{rhs} \approx \frac{1}{s} \left[ \ln \left( \frac{1 + 2p}{2p} \right) \right].$$  (60)

Equating equations (59) and (60) gives us the exact Laplace transform of the scaling function $f_2(v)$

$$\int_0^\infty f_2(v) e^{-pv} dv = \frac{1}{p} \ln \left( \frac{1 + 2p}{2p} \right).$$  (61)

This Laplace transform can be explicitly inverted to give the result in equation (12), with the asymptotic behaviours given in equation (13) (figure 7).
4. Summary and conclusion

In this paper, we have studied the record statistics of a RW of $N$ steps with a resetting probability $r \in [0, 1]$. This model interpolates between two well studied models, namely the standard RW model in the limit $r \to 0$ and the IID model in the limit $(1 - r) \ll 1$. In the presence of a nonzero resetting probability $r$, the positions of the walker are weakly correlated in time and hence it interpolates between the RW model where the positions are strongly correlated and the IID model where the positions are uncorrelated. For arbitrary $r \in [0, 1]$, we have computed exactly the mean number of records $\langle R_N \rangle$ and shown that it is universal for any $N$, i.e. independent of the jump distribution $f(\eta)$, as long as $f(\eta)$ is symmetric and continuous. For example, in the presence of resetting, both RWs with, say a Gaussian jump distribution, as well as Lévy flights share the same $\langle R_N \rangle$ for every $N$ and fixed $r$. We also computed the exact universal crossover scaling functions for $\langle R_N \rangle$ in the two limits $r \to 0$ and $r \to 1$. Finally, we have performed numerical simulations which are in excellent agreement with our analytical predictions.

It is well known that for both $r = 0$ (the RW problem without resetting) and $r$ close to 1 (IID model), not only the mean number of records $\langle R_N \rangle$ but also the higher moments of $R_N$, and its full probability distribution, are universal. In this paper, we have seen that the first moment is universal for any given $r \in [0, 1]$. It is then natural to ask whether the higher moments of $R_N$ are also universal for intermediate values of $r$ between the two known limits $r = 0$ and $r$ close to 1. Our preliminary computations and the simulations for the variance seem to indicate that the universality does not extend to higher moments beyond the average, for $0 < r < 1$. One reason behind this could be the following. It is clear from our exact results that the presence of $0 < r < 1$ does affect the average value of the number of records, but still keeps it universal. However, the higher moments involve $n$-point correlation functions of the $\sigma_i$ variables (with $n \geq 2$), which seem to be more sensitive to the jump distribution compared to the case $n = 1$—thus making the higher moments non-universal. However, one may expect that there will be a vestige of universality still left in the full distribution of $R_N$ at least in the two limiting cases $r$ close to 0 and $r$ close to 1. Hence it will be interesting to compute the full distribution of $R_N$ in these two limits—this is left for future investigations.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Computation of $f_1(u)$

In this appendix we give the details of the computations of $f_1(u)$ leading to the expression given in equation (47) in the main text. Our starting point is the expression for the Laplace transform of $f_1(u)$ given in equation (43). By performing an integration by parts and using $f_1(0) = 0$ from (46), we find, from equation (43)

$$\int_0^\infty f_1(u) e^{-ru} du = \sqrt{1 + p} \ln \left( \frac{1 + p}{p} \right).$$  

(A1)

To proceed, we use an integral representation of $\ln((1 + p)/p)$ to write the rhs of (A1) as

$$\sqrt{1 + p} \ln \left( \frac{1 + p}{p} \right) = \int_0^1 \frac{\sqrt{1 + p}}{z + p} dz.$$  

(A2)
The inverse Laplace transform of the integrand in (A2) can then be explicitly computed, yielding
\[ \mathcal{L}^{-1}_{p \to u} \left( \frac{\sqrt{1 + p}}{z + p} \right) = e^{-u} \sqrt{\frac{\pi}{u}} + e^{-z u} \sqrt{1 - z} \text{erf} \left( \sqrt{u(1 - z)} \right). \] (A3)

We now insert this result in the integral over \( z \) in equation (A2) to obtain, using equation (A1)
\[ f'_1(u) = \int_{0}^{1} \mathcal{L}^{-1}_{p \to u} \left( \frac{\sqrt{1 + p}}{z + p} \right) \, dz = \int_{0}^{1} \left[ e^{-u} \sqrt{\frac{\pi}{u}} + e^{-z u} \sqrt{1 - z} \text{erf} \left( \sqrt{u(1 - z)} \right) \right] \, dz. \] (A4)

The integral of the first term in equation (A4) is trivial. For the second one, we perform the change of variable \( x = 1 - z \) and use the identity (obtained using Mathematica)
\[ \int_{0}^{1} e^{-z u} \sqrt{1 - z} \text{erf} \left( \sqrt{u(1 - z)} \right) \, dz = \int_{0}^{1} e^{-u(1-x)} \sqrt{x} \text{erf}(\sqrt{ux}) \, dx \]
\[ = e^{-u} \left[ -\frac{1}{\sqrt{\pi u}} + \frac{1}{\sqrt{\pi u}} z F_2 \left( \{1, 1\}, \left\{ \frac{1}{2}, 2 \right\}, u \right) \right]. \] (A5)

Finally, inserting this result (A5) in equation (A4), yields the result given in equation (47) in the main text.

**ORCID iDs**

Satya N Majumdar  
https://orcid.org/0000-0001-9844-6980

Philippe Mounaix  
https://orcid.org/0000-0002-0186-9475

Sanjib Sabhapandit  
https://orcid.org/0000-0002-7417-1350

Grégory Schehr  
https://orcid.org/0000-0001-6648-5213

**References**

[1] Chandler K N 1952 *J. R. Stat. Soc.* B 14 220
[2] Foster F G and Stuart A 1954 *J. R. Stat. Soc.* B 16 1
[3] Hoyt D V 1981 *Clim. Change* 3 243
[4] Bassett G W 1992 *Clim. Change* 21 303
[5] Schmittmann B and Zia R K P 1999 *Am. J. Phys.* 67 1269
[6] Benestad R E 2003 *Clim. Rev.* 25 3–13
[7] Redner S and Petersen M R 2006 *Phys. Rev. E* 74 061114
[8] Krug J 2007 *J. Stat. Mech.* P07001
[9] Wergen G and Krug J 2010 *Europhys. Lett.* 92 30008
[10] Anderson A and Kostinski A 2010 *J. Appl. Meteorol. Clim.* 49 1681
[11] Wergen G, Hense A and Krug J 2013 *Clim Dyn* 42 1275
[12] Barlevy G and Nagaraja H N 2006 *J. Appl. Probab.* 43 1119
[13] Wergen G, Bogner M and Krug J 2011 *Phys. Rev. E* 83 051109
[14] Sabir B and Lanthanum M S 2014 *Phys. Rev. E* 90 032126
[15] Matalas N C 1997 *Clim. Change* 37 89
[16] Gembris D, Taylor J G and Suter D 2002 *Nature* 417 506
[17] Ben-Naim E, Redner S and Vazquez F 2007 *Europhys. Lett.* 77 30005
[18] Franke J, Wergen G and Krug J 2012 *Phys. Rev. Lett.* 108 064101
[76] Evans M R, Majumdar S N and Mallick K 2013 J. Phys. A: Math. Theor. 46 185001
[77] Christou C and Schadschneider A 2015 J. Phys. A: Math. Theor. 48 285003
[78] Bhat U, De Bacco C and Redner S 2016 J. Stat. Mech. 083401
[79] De Bruyne B, Randon-Furling J and Redner S 2020 Phys. Rev. Lett. 125 050602
[80] Bressloff P C 2020 J. Phys. A: Math. Theor. 53 275003
[81] Bressloff P C 2021 J. Stat. Mech. 063206
[82] Pollaczek F 1952 C. R. Acad. Sci. 234 2334
[83] Sparre Andersen E 1954 Math. Scand. 2 195
[84] Spitzer F 1956 Trans. Am. Math. Soc. 82 323
[85] Ivanov V V 1994 Astron. Astrophys. 286 328
[86] Comtet A and Majumdar S N 2005 J. Stat. Mech. P06013
[87] Majumdar S N, Mounaix Ph and Schehr G 2017 J. Phys. A: Math. Theor. 50 465002
[88] Majumdar S N, Mounaix Ph and Schehr G 2014 J. Stat. Mech. P09013